Heritage Science

RESEARCH

Open Access

Machine learning based prediction for bulk porosity and static elastic modulus of Yungang Grottoes sandstone

Ruoyu Zhang[1,2,3], Jizhong Huang[2,3], Yuan Cheng[2,3*] and Yue Zhang[2,3]

Abstract

In this work, four mainstream machine learning (ML) techniques are used to evaluate the bulk porosity and static elastic modulus of weathered Yungang Grottoes sandstone. Datasets are gathered from the experiments, which includes 432 groups effective experimental data including 8 inputs features. bulk porosity and static elastic modulus were considered as outputs to determine the weathering degrees of Yungang Grottoes sandstone. The 4 performance criteria were used to evaluate the ML models. Results demonstrate that the Artificial Neural Network (ANN) is the best-fitted models for estimating the bulk porosity and static elastic modulus compared to Multiple Linear Regression (MLR), Support Vector Regression (SVR), Gaussian Process Regression (GPR). The accuracy of the trained model for static elastic modulus is slightly higher than that of bulk porosity. The GPR and ANN model can accurately predict the bulk porosity and static elastic modulus in training stages. The ANN with multi-hidden layers developed is competent with high degree of precision and generalization ability for bulk porosity and static elastic modulus compared to other selected regression-based ML models (MLR, SVR, and GPR). The coefficient of determinations of ANN in the range of (0.9537–0.9641) during the testing stages is more stable and higher than that of (0.8883–0.9453) other built ML models. The prediction efficiency of pretrained ANN model was well adjusted for the actual and forecast datasets at the training and testing stages, and the error range was no more than 0.7% and 0.15 GPa at both stages of prediction for bulk porosity and static elastic modulus respectively. And the ANN based static elastic modulus prediction model's error proportions significantly decreased and were confined to a modest range between + 10% and − 10%. The proposed surrogate models are valid for the bulk porosity ranging from 7 to 14% and the static elastic modulus ranging from 0.7 to 1.4 Gpa, which can be utilized for the accurate and fast prediction of the weathering degrees of Yungang Grottoes sandstone.

Keywords Machine learning based prediction, Grottoes sandstone, Bulk porosity, Static elastic modulus, Artificial neural network

*Correspondence:
Yuan Cheng
chengyuan@shu.edu.cn
Full list of author information is available at the end of the article

Introduction

Stone cultural heritages with extremely high historical and cultural value have a long history and are distributed around the world [1]. Grotto like stone heritages is the most representative type [2–4]. Many immovable stone heritages have undergone thousands of years of natural or artificial weathering, resulting in significant deterioration of the material properties on its surface or internal [5, 6]. The weathering degree of large exposed stone heritages, is an important indicator for their preservation and a significant parameter that should be determined before the heritages are restored [6–8]. Many researchers have carried out restoration and protection work on it, but before that, it is necessary to evaluate or quantify the weathering degrees of stone cultural heritages.

Weathering diseases are caused by multiple factors coupled and long-term effects, mainly induced by freeze–thaw [9], chemical reaction from pollutants or precipitations [10, 11], temperature [12], moisture [13] and coupling effect of this factors. In general, weathering diseases can be classified into different types. In 1984, Huang and his colleagues [14] first classified weathering diseases into four types based on the weathering characteristics of some famous sandstone grottoes in China (e.g., Yungang Grottoes, Longmen Grottoes, Dazu Caves, Leshan Giant Buddha, etc.): powdery weathering, sheet-like weathering, strip-like weathering, and plate-like weathering. Other studies classified weathering types of sandstone grottoes from different perspectives [15–17], but described weathering characteristics were consistent basically. Four types typical weathering diseases are as shown in Fig. 1. The products formed by powdery weathering [18] are mostly white powdery or flocculent as shown in Fig. 1a, mainly occurring on the surface of medium to coarse-grained sandstone with accumulations formed. The structure is extreme loose and the bonding ability between powders is poor (or disappeared). The products formed by sheet-like weathering [19] are mostly thin sheets as shown in Fig. 1b, and the thickness of the thin sheets are usually determined by the particle size of the sandstone in the rock layer. The thickness of the thin sheets formed by coarse-grained sandstone is generally 3–4 mm, and the thickness of the thin sheets in fine-grained sandstone is generally 0.5–1 mm. Strip-like weathering, also known as cave-like weathering, is

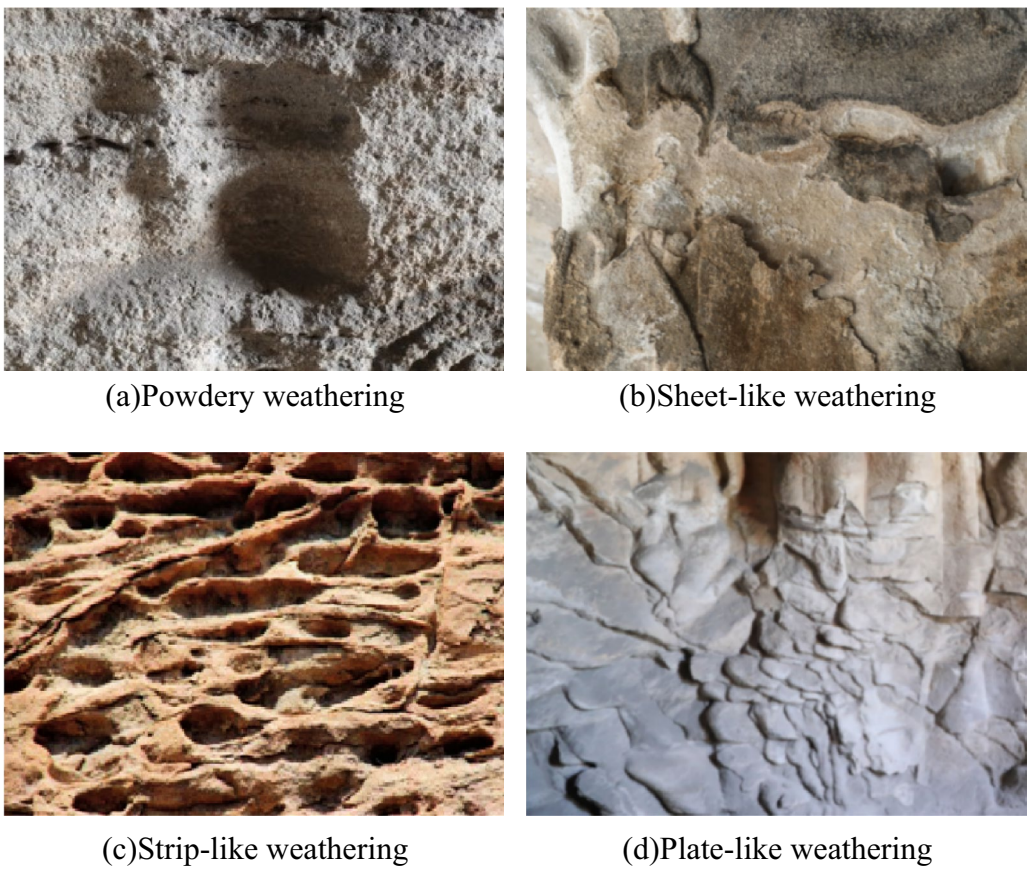


Fig.1 Typical weathering diseases of sandstone grottoes

parallel to the sandstone bedding in a strip-like [20] or cave-like way as shown in Fig. 1c. Strip-like weathering usually occurs in rock layers containing mud and siltstone, as well as medium to coarse-grained sandstone. The characteristic of plate-like weathering is that its products form plate-like peeling [21] as shown in Fig. 1d, generally parallel to the wall surface. The thickness of the plate is generally 2–4 cm, mainly distributed below 2 m from the ground.

The above four rough weathering classifications are the most common and mainly suitable for weathering in Yungang Grottoes. And there are more detailed weathering classifications for stone heritage (see ICOMOS-ISCS: Illustrated glossary on stone deterioration patterns [22]). Actually, the climate conditions, sandstone lithology, and architectural forms (such as churches) of other regions' sandstone cultural heritages may not be same or similar. André and her colleagues [23, 24] propose that external factors to control the weatherability of sandstone monuments and churches, suah as air pollution, the impact of wetting–drying cycles, frost action and salt concentration. And Oguchi et al. [7] interpret 8 patterns of disintegration of sandstone heritages due to weathering: cracking and disintegration, crumbling, scaling, surface pitting, thick scaling, delamination, cracking, and expansive dilatation patterns. For more detailed surface degradation of sandstone heritage, 12 types of deterioration are described and studied by Zhang et al. [25], including dust deposits, salt crystallization, black smudges of smoke, yellow stains, Residual resin, biological growth, cement-blocks, paint stain, grease, scrawl, water spots, plaster stain. Mammoliti et al. [26] assess various forms of weathering from a sandstone column of monument via a non-destructive approach, such as discolouration, scaling and loss of stone volume. In fact, there are many effective methods for evaluating the weathering degrees of grotto heritages [27] include high-density electrical methods, spectroscopic techniques, hardness tester, and ultrasonic methods, etc. [28–33]. The indicators obtained by these non-destructive or micro-destructive methods can define the degree of weathering of the sandstone, but these indicators can not describe the properties of the sandstone (e.g., elastic modulus, pore structure, etc.) directly. The properties of the sandstone itself are generally tested by cumbersome (non-portable) or destructive methods, such as using uniaxial compression test (UCT) to determine the elastic modulus of the sandstone, and using nuclear magnetic resonance (NMR) or mercury intrusion porosimetry (MIP) techniques to determine the pore parameters of the sandstone. The methods mentioned are generally difficult to implement at grottoes cultural heritages in-situ. Therefore, it is necessary to propose a nondestructive and ready-to-use method to obtain the bulk porosity and elastic modulus approximately for determining weathering degree of grotto sandstone.

To accurately evaluate weathered grottoes sandstone, it is critical to use an accurate and efficient model for forecasting the typical micro and macro indicators, such as bulk porosity and static elastic modulus. There is no computational model available in the previous studies to forecast the bulk porosity and static elastic modulus of weathered grottoes sandstone. Very limited analytical models (theoretical models) are available in the previous studies. The computational cost is very high and the efficiency is intolerable slow by using the refined numerical models. The artificial intelligence (AI) or machine learning (ML) techniques may give researchers a surrogate model to estimate bulk porosity and static elastic modulus to determine weathering degree non-destructively and quickly, allowing them to plan less experimental work with a higher level of accuracy. In the past 2 decades, ML based surrogate models have been investigated in stone cultural heritage conservation for the disease features detection [34–37]. Bewes et al. [38] predicted sex from skeletal remains using deep learning (DL) model; Cintas et al. [39] used the pottery profile classification method. There are also damage or disease features assessment in the limited components of historic monuments [40, 41]. Hatır et.al [42–45] use vision-based DL methods to classify the types of deterioration with high accuracy. Currently, many studies categorize weathering of grottoes without quantifying a specific indicator. But to predict and evaluate weathered grottoes sandstone accurately, it is critical to use an accurate and efficient model for forecasting the typical micro indicators (e,g, bulk porosity and pore size distribution, etc.). There is no computational model available in the previous studies to forecast the pore parameters and modulus information of weathered grottoes sandstone. In addition, the human errors that experts make in weathering recognition will accelerate the weathering process in building stones. There are also some factors that can lead to misjudgment. Taking the acoustic emission detection of stone cultural heritage as an example, if there are not enough testing frequencies, the values of compression wave velocity will undergo significant variability. This variability may be attributed to the inherent heterogeneity of the sandstone itself and potential inaccuracies in portable device, etc. Thus, it is necessary to propose a surrogate model for quickly and non-destructively obtaining indicators that can directly describe the weathering of sandstone in grottoes (e.g., bulk porosity and static elastic modulus, etc.).

Although some studies have conducted predictions on the degree of weathering of grotto sandstone, less input features are used such as compression wave velocity, hardness, etc. But that using single and limited detection approaches is difficult to consider multiple weathering factors (e.g., acid, alkali, salt and freeze–thaw, etc.). Meanwhile, the types of machine learning based surrogate models used are relatively limited and monotonous. There is no comparison of performance for predictions between different ML models. Hence, as a supplement, the sandstone from the Yungang Grottoes are selected as typical object, and the feasibility and reliability of machine learning models including GPR, SVR, ANN and MLR in predicting bulk porosity and static elastic modulus of weathered sandstone were studied in this work. Firstly, simulation tests on the weathering of grotto sandstone were carried out under acid, alkali, salt and freeze–thaw conditions in Sect. "Dataset description and analysis of weathering experiments of sandstone". An experimental database containing 432 unique records of weathering tests, as well as 8 characteristics, was gathered and processed for this purpose. Then, the framework of the data processing approach and machine learning modelling technique used in this study has been described in Sect. "Experimental database collection and machine learning approach". The prediction performance of some typical ML models have been evaluated and analysed in Sect. "Predicted performances and comparative analyses". This study can provide reference for the prediction and evaluation of weathering degrees of stone cultural heritages.

Dataset description and analysis of weathering experiments of sandstone

Characterization of weathered sandstone

The stratigraphic structure of the area where the Yungang Grottoes are located is relatively simple. It belongs to the upper Middle Jurassic formations of the Mesozoic era and the middle to upper Quaternary formations. These include the Middle Pleistocene (residual-alluvial), Upper Pleistocene (alluvial-pluvial), Holocene (alluvial-colluvial), and the Middle Jurassic Yungang Formation. The grotto area is a typical continental semi-arid climate, and average temperature of the coldest month and the hottest month, annual average temperature and annual average precipitation are − 13.3 ℃, 23.1 ℃, 9.1 ℃ and 423.8 mm. The rock samples used in this study were taken from the exposed strata in the eastern part of the Yungang Grottoes area, which date back to the same period as the Yungang Grottoes. And Yungang grotto has been subjected to air pollution since the last century, field investigations have indicated that acid gases, particularly sulfur dioxide (SO_2) or nitrogen dioxide (NO_2), have accumulated on the surface of the sculptures [10]. Especially, the concentration of SO_2 continued to rise, reaching to a historical high value of 767.01 μg/m^3 in 1991 (the limitation in Chinese national standards is 20 μg/m^3). These gases will be formed into acid or salt solutions when there exists water in the environment, thus continuing to cause deterioration of sandstone. Due to the complex composition and structure of sandstone, there may exist significant differences in physical parameters between samples taken from the same rock. To ensure that the final samples are minimally affected by the surface weathering layer, the sampling location should be at least 20 cm away from the surface of the rock (weathering layer is about within 5 cm).

In order to ensure the reliability of the experimental results, it is necessary to screen sandstone samples with relatively small differences in physical properties using velocity of compression wave as screening criteria, and also to compare the properties of sandstone before and after simulated weathering. Finally, 156 samples with relatively small differences in physical parameters were selected from the numerous cut sandstones as experimental samples. The rock samples were made into cubic blocks with a size of 5 × 5 × 5 cm^3 using cutting machine. Destructive and non-destructive tests of weathered sandstone were carried out. In this tests, deionized water was used for the freeze–thaw cycle, and the solutions of acid, alkali, and salt are sulfuric acid (H_2SO_4), sodium hydroxide ($NaOH$) and sodium sulfate ($NaSO_4$) respectively. To accelerate the weathering rate of sandstone in each simulated weathering cycles, the pH of H_2SO_4 solution is confirmed as 1 and pH of $NaOH$ solution is confirmed as 14. The saturated $NaSO_4$ solution are also used in this study. It should be noted that, the pH of groundwater in Yungang Grottoes was 6–8; the TDS was lower than 3 g/L and the concentration of SO_4^{2-} was no more than 1.5 g/L. Hence, pH = 1 and pH = 14 is not the actual situation, which indicate that extreme acidic and alkaline solution cannot reflect the actual corrosion of sandstone in Yungang Grottoes very accurately. In addition to the faster corrosion rate compared to weak acids (pH = 3–6), the specific reasons for selecting strong acids in this experiment are: (a) the greater corrosion depth of extreme acid solution allows for the easy formation of a weathering layer extending from the surface to a certain depth within the interior; (b) extreme acid solution can rapidly dissolve minerals, leading to the loosening of the sandstone structure and even its disintegration. The specific information are shown in Table 1.

The weathering of the Yungang Grottoes is the result of long-term and random processes. Simulating weathering under real environmental conditions would require

Table 1 Information of solutions

Solution	Molecular formula	Concentration (mol/L)	pH	Temperature (℃)
Sulfuric acid	H_2SO_4	0.05	1	20
Sodium hydroxide	$NaOH$	1	14	20
Sodium sulfate	$NaSO_4$	3.35 (saturation)	7.5	20

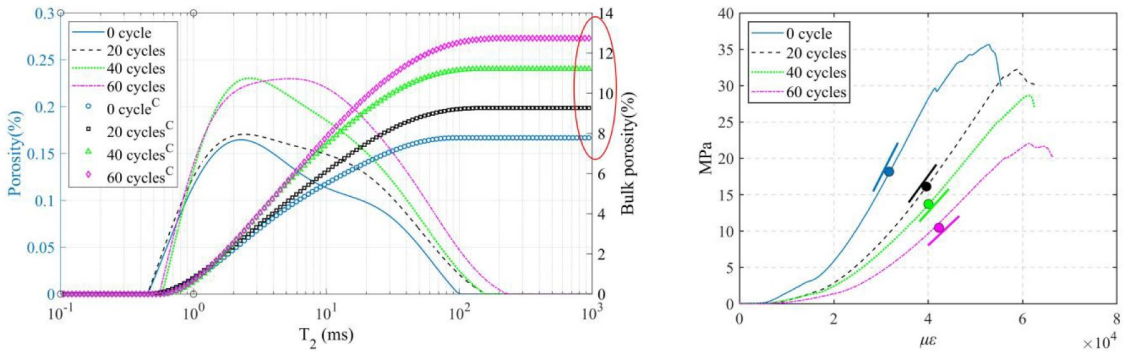
extremely high costs. The purpose of the accelerated weathering experiments under various environmental conditions in this study is to rapidly replicate the current weathering state of the Yungang Grottoes as closely as possible. Although the results may not precisely simulate the actual weathering processes, it can produce sandstone samples with different types and apparent gradients of weathering. Such samples, exhibiting distinct weathering gradients, are crucial for enhancing the generalization capability and performance of subsequent machine learning models.

Meanwhile, The detailed procedure freeze–thaw and solution soaking cycles are as follows respectively:

Firstly, the procedure for freeze–thaw cycle is introduced: (a)All samples are dried to a constant weight in an oven at 105 °C, then placed in a sealed container to cool down to room temperature. (b) With the exception of the unweathered sandstone in the control group (0 cycles), 39 samples undergo vacuum saturation. The specific steps are as follows: place the samples in a desiccator filled with distilled water (ensuring that the samples are fully submerged); seal the container and connect the vacuum pump to evacuate the air from the desiccator; after 4 h of evacuation, allow the samples to remain in distilled water at atmospheric pressure for an additional 4 h. (c) Con·sidering that the minimum temperature in the Yungang area may exceed − 20℃, the minimum temperature for the freeze–thaw cycle in our experiment is − 20℃, and the thawing temperature at 20 °C, with both freezing and thawing phases lasting 4 h each. The freeze–thaw cycle process for the saturated sandstone samples is as follows: place the samples in a refrigerator set to − 20 °C and freeze for 4 h; then, thaw them in distilled water for 4 h. Each 8-h period constitutes one cycle. After every 5 cycles, all samples are dried in an oven at 105 °C for 12 h and then allowed to cool to room temperature.

And the procedure for 3 solutions soaking cycles are also as follows: (a) Before the cyclic tests for each group, all rock samples were placed in an oven and dried at 105 °C for 12 h. After drying, the samples were cooled to room temperature. (b) The 36 sandstone samples from each group were then placed into acid, alkali, and salt solutions corresponding to their designated numbers and soaked for 12 h. During soaking, care was taken to avoid contact between samples to ensure full immersion. c) After the soaking period, the samples were removed and rinsed with deionized water to clean the surface. The samples were then placed in an oven and dried at 105 °C for 12 h. Upon completion of drying, one acid/alkali/salt cycle was considered complete. After cooling, the samples were placed back into the solution to continue the next cycle. It should be noted that the conductivity and pH of deionized water used in laboratory are about 0.06 μS/cm and 6.6 (± 0.2), because it easily absorbs CO_2 from the air due to the high purity of deionized water, and trace amounts of carbonic acid are formed.

Two important indicators for describing sandstone weathering, namely, uniaxial compressive strength test and nuclear magnetic resonance were selected to determine the values of the bulk porosity and static elastic modulus in every weathering cycles. The bulk porosity (total bulk porosity of the sample) is calculated by using accumulated T_2 spectra. The static elastic modulus can be measured through uniaxial compression tests and is



(a)bulk porosity (b)Static elastic modulus

Fig.2 Determination of the bulk porosity and static elastic modulus (freeze–thaw cycles)

calculated from the linear portion of the stress–strain curve according to the recommendations of the International Society for Rock Mechanics (ISRM) [46]. Figure 2 shows the determination of the bulk porosity and static elastic modulus via T_2 spectra curves (milliseconds) and stress–strain curves (megapascals-microstrains) respectively. Where, The superscript "C" represents accumulated values of T_2 spectra and the red elliptical symbol represents the bulk porosities; the colorful dots represent the 50% of maximum load and colorful straight lines represent the slope of a tangent in every cycles in Fig. 2b. The specific patterns of change will be explained in the next Sect. "The variation of input features and outputs". X-Ray Fluorescence (XRF), X-ray diffraction (XRD) were selected to determine the change trends of chemical elements and minerals in weathering processes. Meanwhile, some portable devices such as, electronic balance, ultrasonic emission equipment, magnetic susceptibility meter and leeb hardness meter were selected to measure other indicators. And some macro and micro detection indicators are obtained.

Considering the physical and chemical properties of sandstone will be changed simultaneously subjected to complex weathering processes, the compression wave velocity (V_p), variation of mass before and after weathering (M), magnetic susceptibility (cm), hardness of leeb (HRC), the ratio of silicon to aluminum (S-A) [47], the ratio of the sum of calcium and potassium to aluminum(Ca + K-A) [48], the relative content of kaolinite (Kln), feldspar (Kfs) are all important parameters to consider. The ratio of chemical elements were measured by XRF, and the ratio of mineral was characterized by XRD semi-quantitatively [5]. The equipments' information and key parameters required for testing are as follows: (a) The PB4002-S electronic balance is used to measure M, with a range of 0.5 g to 4100 g and an accuracy of 0.01 g. The mass of the samples is recorded before the experiment and after every 5 cycles. (b) The the GTJ-U200 ultrasonic rebound tester is used to measure V_p, with an emission frequency of 50 kHz and a sound time measurement accuracy of 0.1 μs. Each opposing surface will be measured three times, and the average value is taken as the final result. (c) The SW-6210S Leeb hardness tester from Sndway is used to measure HRC, with a detection range of 170–960 HLD. (d) The magnetic susceptibility meter is used to measure the magnetic susceptibility of rock outcrops, rock specimens, or soil samples. This experiment uses the KM-7 magnetic susceptibility meter is used to measure cm with an operating frequency of 10 Hz and a sensitivity of 10^{-6} SI. (e) The MacroMR12-110H-GS-HTH nuclear magnetic resonance detecter is used to measure the relaxation time spectra of saturated rock samples, and the bulk porosities will be obtained based on the cumulative relaxation time spectra. (f) The uniaxial compressive strength test is conducted using the CTM 9200 electro-hydraulic servo universal testing machine, with a maximum load of 200 kN. The mode of force control is applied in the experiment and the compression rate is 0.5 MPa/s constantly. The compression direction is perpendicular to the bedding plane of the sandstone, and after the test, and the force–displacement curves of the actuator are recorded to calculate static elastic modulus. Where, the area of surfaces under compressive pressure and edge length of all sandstones are equal, namely 2.5×10^3 mm^2 and 0.5×10^2 mm^2,which can be used to estimate the stress–strain curves of sandstones. Actually, the elastic modulus should be final determined via strain gauges because the displacement values recorded during uniaxial compressive strength tests were not measured directly on specimens but were actuator values of displacement. Therefore, the stress–strain curve obtained using the experimental methods in this study is not entirely accurate and the curves may exhibit some lag (strain is underestimated), indicating that the final obtained value of the elastic modulus may be overestimated slightly. Although the localized small strain in sandstone is not captured, the static elastic modulus can be equivalently calculated when considering the linear segment of the rapidly increasing load in the sandstone. (g) The Niton XL3t GOLDD + handheld XRF analyzer is used to effectively characterize the chemical element distribution on the surface of the sandstone and calculate values of S-A and Ca + K-A. (h) The Smartlab X-ray diffractometer is used to measure the relative content of Kln and Kfs, with an angle range of 10°–90° and an angular accuracy of 0.0001. Figure 3 is muti-weathering tests of Yungang sandstone. The changing trends of some indicators detected will be mentioned in the next subsection.

The variation of input features and outputs

Before making quantitative predictions in machine learning, it is necessary to explore the variation patterns of each input features with the weathering process. In order to improve the accuracy of subsequent machine learning, we removed obvious outliers caused by improper human operation (in the box plot). Figures 4, 5, 6, 7 show a box plot of the variation of compression wave velocity, mass change, absolute magnetic susceptibility, and leeb hardness with the number of weathering cycles under four different weathering conditions. It should be noted that Yungang sandstone belongs to medium coarse feldspar-quartz sandstone, and its main minerals are such as quartz, feldspar, calcite, clay minerals and kaolin are all diamagnetic minerals. Therefore, the original test results of the magnetic susceptibility of Yungang sandstone were

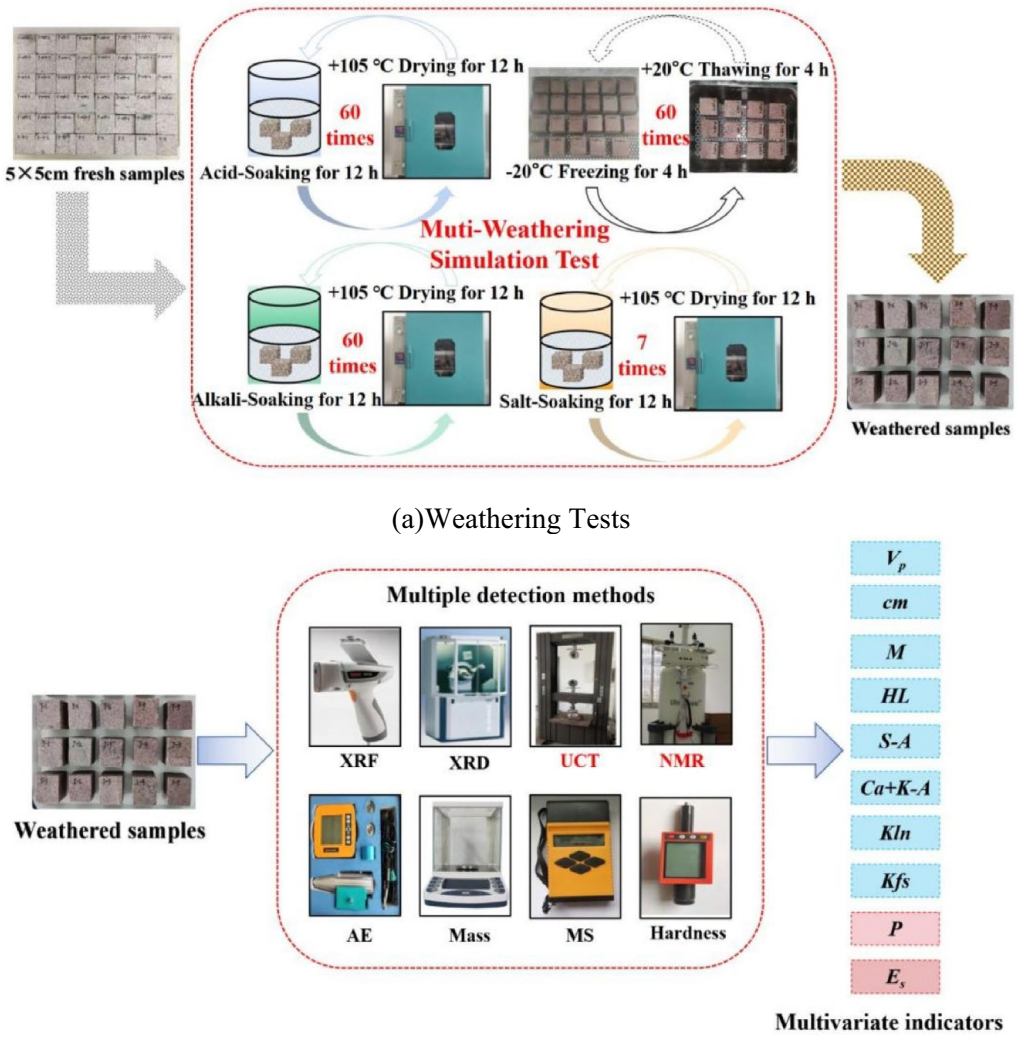


Fig.3 Muti-Weathering Tests of Yungang Sandstone

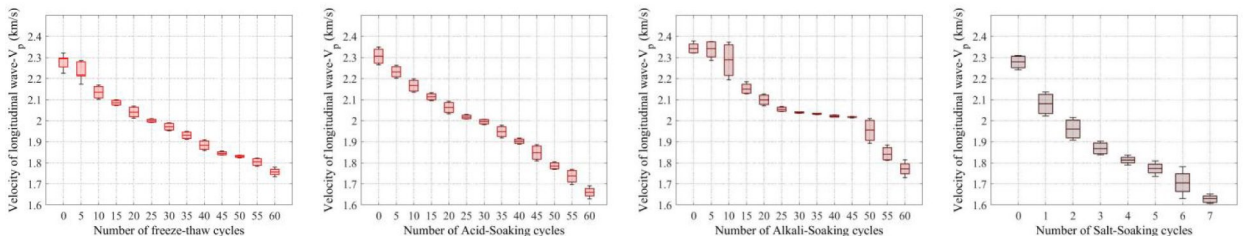


Fig. 4 Box plot for Variation of V_p of different weathered sandstone

negative. Based on Figs. 4, 5, 6, 7, it can be seen that there is a strong correlation between the four macroscopic non-destructive testing indicators, namely decreasing of wave velocity, decreasing of mass, increasing of absolute magnetic susceptibility and decreasing of surface hardness. All of which can reflect the degree of weathering of

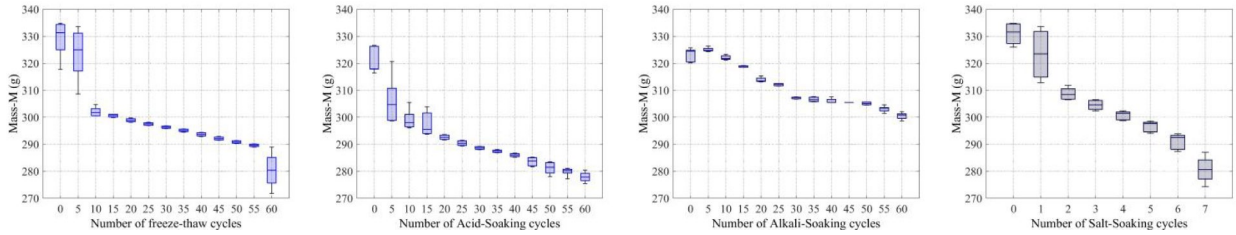


Fig. 5 Box plot for Variation of M of different weathered sandstone

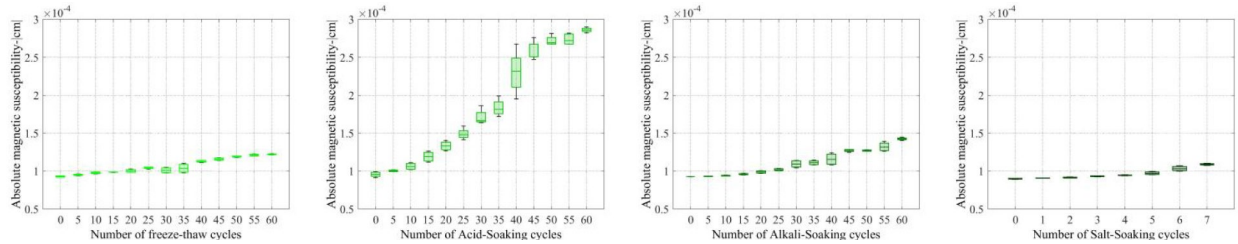


Fig. 6 Box plot for Variation of |cm| of different weathered sandstone

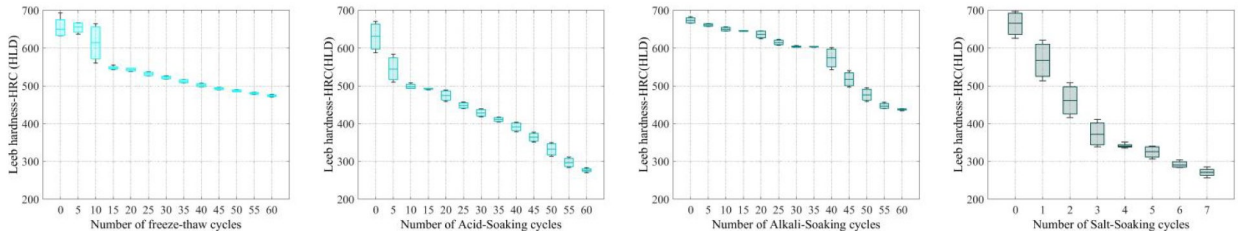


Fig. 7 Box plot for Variation of HRC of different weathered sandstone

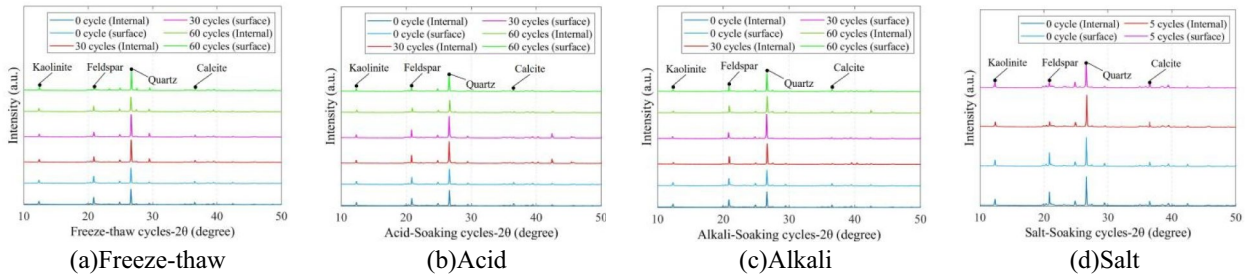


Fig. 8 The XRD spectra of minerals under four different weathering conditions

sandstone. However, experimental measured data may lead to some reverse trend, such as in freeze–thaw cycles and alkaline weathering conditions.

Figure 8 and Table 2 show the variations of chemical elements' content and minerals diffraction spectra under four different weathering conditions respectively.

Table 2 Proportion of chemical elements in different weathered sandstone via XRF (%)

Types	Cycles	Si	Al	Ca	K	Fe	Mg	Ti	S
Freeze–thaw	0	45.76	5.94	1.47	1.31	1.98	0.50	0.29	0.01
	20	44.41	6.71	1.49	1.57	1.80	0.35	0.21	0.01
	40	44.47	6.29	0.76	1.30	2.10	0.40	0.20	0.01
	60	42.23	5.77	0.68	1.40	2.27	0.39	0.19	0.01
Acid	0	48.86	4.60	2.24	1.09	1.21	0.44	0.08	0.01
	20	47.31	5.85	0.45	1.23	1.29	0.41	0.13	1.53
	40	46.89	5.83	0.38	1.11	0.93	0.38	0.12	1.67
	60	43.35	5.74	0.48	1.05	0.78	0.34	0.11	2.18
Alkali	0	45.85	5.55	2.35	1.26	1.73	0.58	0.13	0.02
	20	45.31	3.12	2.40	1.05	1.54	0.47	0.14	0.01
	40	44.77	3.49	2.14	1.03	1.82	0.41	0.15	0.02
	60	44.92	3.27	1.99	1.01	1.67	0.32	0.13	0.01
Salt	0	46.71	5.34	2.40	1.24	2.05	0.50	0.20	0.01
	3	28.53	1.42	1.57	0.62	1.12	0.61	0.12	13.72
	5	25.69	1.64	1.37	0.68	1.31	0.69	0.14	13.55
	7	23.96	1.51	1.18	0.61	1.23	0.55	0.15	15.31

It can be seen that the main elements in sandstone samples are Si, Al, Ca, Mg, K, which constitute minerals such as quartz, feldspar, calcite. While Fe, Ti, S, and several trace elements constitute other metal minerals in sandstone. During the freeze–thaw cycles, the Si-content is decreased by 3.53% after total 60 cycles, and the Al-content is increased by 0.83%, resulting in the decreasing the ratio of silicon to aluminum (S-A). Ca-content is also decreased by 0.79% and Ca is the main element that makes up feldspar, calcite, dolomite, and other metallic minerals in sandstone. The decreasing Ca-content indicates a certain degree of decrease in the content of above mentioned minerals. After acid cycling test, the Ca-content and Fe-content of sandstone are decreased by 1.76% and 0.43%, respectively. This is because feldspar, calcite, dolomite, metallic iron, and their oxides all undergo dissolution reactions when exposed to acid. During alkali cycling test, the changes of Al-content and Ca-content were large relatively, and decreased by 2.28% and 0.36% respectively. The K-content and Mg-content are decreased slightly, while the Si-content are decreased by 1.73%, which is smaller than other weathered sandstone. The changes of other element content were not significant because the dissolution reactions with alkaline solutions in Yungang sandstone are quartz and feldspar mainly. After salt cycling, the content of Si, Al, Ca, Fe, K and Ti are significantly decreased, especially the content of Si, Al, and Ca are decreased by 22.75%, 3.83%, and 1.22% respectively. The S-content is increased by 15.30%. After 5 cycles, the characteristic peaks of feldspar and calcite in sandstone have weakened, and the salt solution migrates along the pore channels of sandstone to various internal parts of sandstone. Similar conclusions have been obtained in numerous previous studies [2, 9, 10]. Furthermore, XRD analysis determined that the main mineral components of the Yungang sandstone is feldspathic quartz sandstone, consisting of quartz, feldspar, calcite, kaolinite along with other clay minerals, like illite. Table 3 shows the average proportion of main minerals in different weathered sandstone. The relative content of quartz, feldspar, calcite and kaolinite are selected for analysis because the content of illite is less than 1.5%, while that of other minerals are even less lower (e.g.,

Table 3 Average proportion of main minerals in different weathered sandstone via XRD (%)

Types	Cycles	Quartz	Feldspar	Calcite	Kaolinite
Unweathered	0	53.5–61.5	12.3–19.5	4.4–6.6	16.3–19.1
Freeze–thaw	30-S	59.0	11.8	4.1	19.9
	30-I	58.2	12.5	3.6	18.1
	60-S	59.2	9.4	2.8	23.3
	60-I	61.1	9.1	4.1	21.2
Acid	30-S	58.3	11.8	2.9	19.0
	30-I	57.9	11.9	4.3	18.5
	60-S	60.5	7.7	0.7	24.1
	60-I	61.4	8.0	1.9	25.2
Alkali	30-S	58.3	11.8	3.9	18.5
	30-I	59.9	11.5	2.6	19.7
	60-S	57.9	9.7	3.3	19.6
	60-I	60.8	10.2	2.4	20.9
Salt	5-S	59.7	7.8	1.7	26.2
	5-I	60.6	8.3	1.9	24.1

hematite, ankerite, etc.). Where, "-S" and "-I" represent surface and internal of sandstone samples respectively. Comparing the proportion changes of mineral of 4 weathered sandstones, the commonalities are the kaolinite-mineralization of feldspar and the loss of calcite (cement). The most severe loss of feldspar and calcite is due to acid and salt cycles, hence a small amount of sandstone samples will also be disintegrated under these two weathering conditions. Although the freeze–thaw cycle is accompanied by the loss of feldspar, the loss of calcite is not severe. Compared to the cyclic conditions of acid and alkali solutions, the weathering state of surface and internal in the freeze–thaw cycle sandstone are almost the same. Due to the sodium sulfate will be generated in salt weathering, weathered sandstone samples are affected by the crystallization pressure of sodium sulfate, leading to the expansion of pore channels (connected pores) and intensified internal weathering.

Figures 9 and 10 show box plots of the variations of bulk porosity and static elastic modulus of weathered sandstone. It can be seen that the increase of bulk porosity and the decrease of static elastic modulus can characterize different types of sandstone weathering effectively, but the degree of dispersion of acid weathering and salt weathering is relatively larger. The initial bulk porosity and elastic modulus of 12 fresh samples were screened, and the usable data range for fresh sandstone was obtained, which is 7.2–8.3% and 1.33GPa–1.36GPa. The initial bulk porosity of Yungang sandstone is small, and the static elastic modulus is large. Compared with other sites of grottoes sandstone, Yungang sandstone is more dense in diagenesis. Compared with other indicators such as wave velocity and hardness, bulk porosity can better reflect the microscopic weathering characteristics. At the same time, the variations of static elastic modulus can accurately reflect the uniaxial compressive strength and the deformation ability of weathered sandstone, which is the macroscopic weathering characteristics. However, it is not possible to test bulk porosity and staticelastic modulus directly on sandstone cultural heritages sites due to non-destructive testing methods can only be used. Hence, the prediction of bulk porosity and static elastic modulus is particularly important. In the Sect. "Experimental database collection and machine learning approach", statistical characteristics analysis will is conducted on the experimental data firstly, and four different types of machine learning algorithms and criteria for measuring their regression performance are selected.

Experimental database collection and machine learning approach

Statistical characteristics and correlation of database

There is currently no appropriate code for evaluating the bulk porosity and static elastic modulus of Grotto sandstone quantitatively, especially subjected to multiple

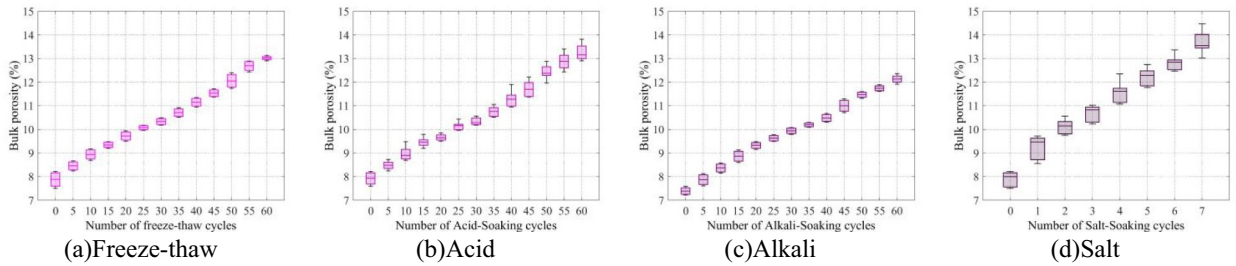


Fig. 9 Box plot for Variation of P of different weathered sandstone

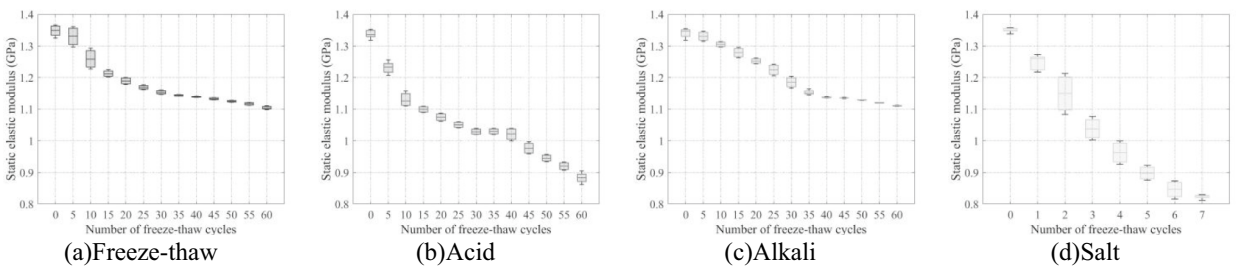


Fig. 10 Box plot for Variation of E_s of different weathered sandstone

weathering conditions. A few traditional weathering conditions are considered in previous studies, such as freeze–thaw cycle or salt cycle tests. As a result, a diversify database of multi-weathered Grotto sandstone should be gathered. The database contains the experimental results of 432 sets data available in Sect. "Dataset description and analysis of weathering experiments of sandstone" under acid, alkali, salt, freeze–thaw cycles, containing typical diseases of sandstone such as surface powdering, discoloration, cracks, etc. Figure 11 shows the statistical characteristics of each input and output features based on relative frequency distribution of test datasets. Nine input variables have been used to construct the prediction models, which includes 4 macro

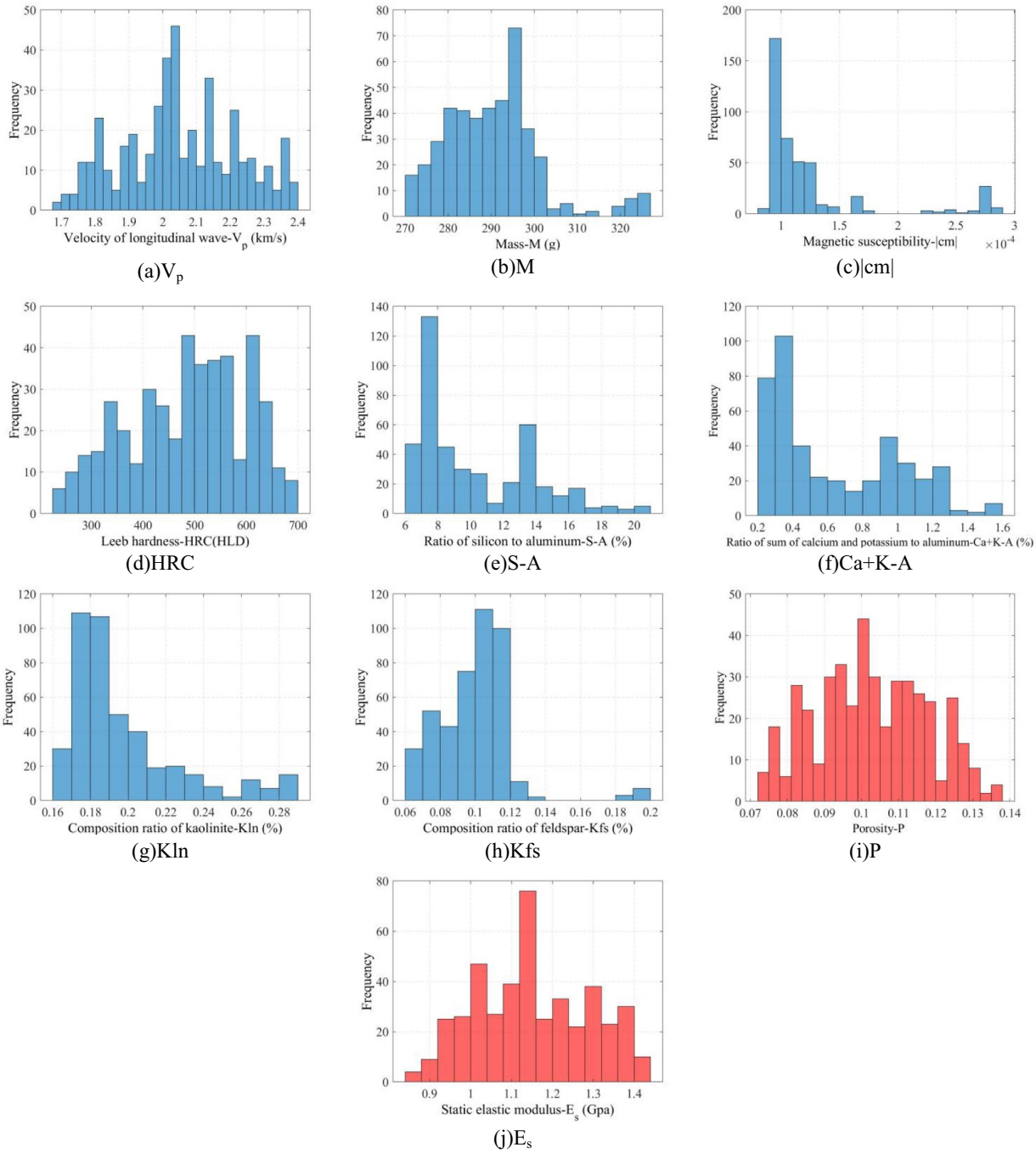
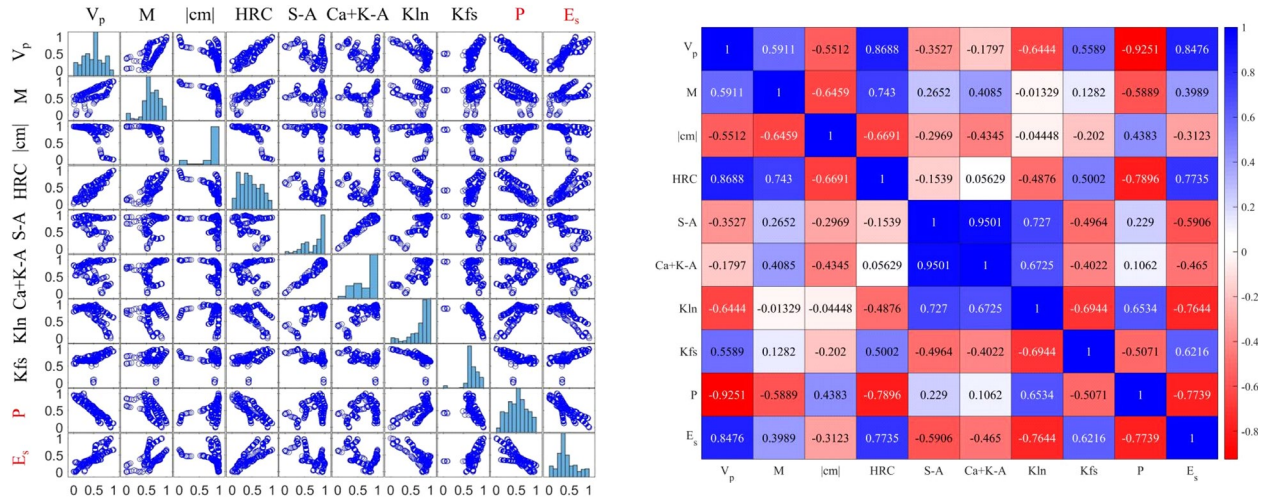


Fig. 11 Relative frequency distribution of experimental datasets



(a)Plotmatrix of normalized dataset (b)Pearson correlation matrix plot

Fig. 12 Correlation analysis of dataset

parameters parameters and 4 microscopic parameters, and the bulk porosity (P) and static elastic modulus (E_s) were specified as an output. The data were collected from Muti-Weathering Simulation Test of Yungang Sandstone in Sect. "Dataset description and analysis of weathering experiments of sandstone".

The plotmtrix of the normalized data and pearson correlation matrix plot are shown in Fig. 12a and b using normalized data. As can be seen that there exist strong correlation between P, E_s and macro indicators, especially the compression wave velocity (V_p) and hardness of leeb (HRC). Previous studies of the weathered sandstone in grottoes have also used these two indicators (V_p, HRC) as quantitative indicators of weathering degree directly. However, due to the large number of chemical reactions accompanying sandstone weathering processes, this is also a vital reason for introducing changes in indicators such as elements and minerals. It can be seen that due to the almost absence of mono-elemental substance in sandstone, the correlation between elements' changes and P, E_s is weak. The change in mineral composition implies a direct change in the internal structure and pore size distribution of sandstone, hence the relative content of minerals can better represent the degrees of weathering, especially the content of kaolinite and feldspar. After that, 70% of the data were chosen as a training set and 30% of the datasets were utilized as a testing set to verify the models' performance and generalization capabilities.

Machine learning models

We selected 4 different types of state-of-the-art and frequently used algorithms for performance comparisons, that is: Multiple Linear Regression (MLR), Support Vector Regression (SVR), Gaussian Process Regression (GPR) and Artificial Neural Network (ANN). These algorithms are common types in the regression prediction field. Some scholars have also summarized its operating mode and mechanism [49]. Due to the parameter settings for each type of algorithm are different, we will select the best settings considering final performance and total execution time of above mentioned 4 algorithms (e. g., types of kernel function, number of layers, etc.).

MLR

MLR is one of the most commonly used type of regression analysis for statistical examination that is often used to scrutinize the connection between a dependent variable and more independent variables [50]. Linear regression models can be expressed by linear equations, and this information may be acquired by employing the regression line. For a total of n number of input features (x_n), the MLR can be expressed as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (1)$$

where: y is the predicted value; β_0 is the intercept; $\beta_1 \sim \beta_n$ are the line slope of x_n.

SVR

SVR examines the extreme limits and draws the edges, and the extreme data points assist in determining the limitations (support vectors) [51]. SVR regression processes are expressed as follows:

$$Z = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\} \text{ and}$$
$$y_i = \sum_{i=1}^{i} (\alpha_i - \beta_i) \times k(x, x_i) + b \tag{2}$$

where α_i and β_i represent Lagrange multipliers and b is the bias, and $k(x, x_i)$ is the kernel function that is a special type of function used for nonlinear mapping in SVR. The kernel function is to map input data from the original feature space to a high-dimensional feature space, making nonlinear problems in the original space linearly separable or approximately linearly separable in the high-dimensional space as illustrated in Fig. 13. And polynomial kernel functions is used in this work.

GPR

A Gaussian processes can be represented by a mean function and a covariance function [52], which can be defined as:

$$f_{GP}(x) \sim GP(\mu(x), k(x, x\prime)) \tag{3}$$

where $f_{GP}(x)$ is Gaussian processes; $\mu(x)$ is mean, and $k(x, x_0)$ is positive-semi definite kernel function which define the covariance between any two realization of $f(x)$ and $f(x_0)$:

$$k(x, x\prime) = \text{cov}\big(f(x), f(x\prime)\big) \tag{4}$$

The mean is often assumed to be zero, and the kernel has parameters θ, i.e., $k(x, x_0|\theta)$. For any infinite collection of inputs $x_1, x_2, \ldots, x_n$, the $f(x_1, x_2, \ldots, x_n) = (f(x_1), f(x_2), \ldots, f(x_n)$ have joint multivariate Gaussian distribution.

$$f(x) \sim N(0, K_{XX}(\theta)) \tag{5}$$

where covariance matrix are defined by the kernel function:

$$[K_{XX}(\theta)]_{i,j} = k(x_i, y_i|\theta) \tag{6}$$

The covariance function aids in the implicit specification of model properties. As shown in Eq. (7), the fundamental and extensively used GPR is made up of a squared exponential covariance function and simple zero mean:

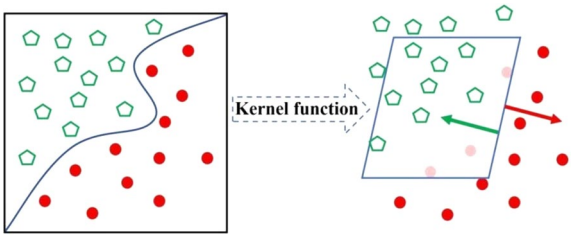


Fig. 13 Nonlinear mapping in SVR

$$k(x, x\prime) \sim \sigma_f^2 \exp\left[\frac{-|x - x\prime|^2}{2l^2}\right] \tag{7}$$

where, l and σ_f are the hyper-parameters and effect the performance of Gaussian Process. σ_f denotes the model noise and l is the scale of length. And rational quadratic kernel function is used in this study.

ANN

ANN abstracts the human brain neural network from the perspective of information processing and forms different networks according to different connection methods [53]. ANN is composed of a large number of nodes (or neurons) connected to each other, with each node representing a specific output function called the activation function. Each connection between two nodes represents a weight for the signal passing through that connection, which is equivalent to the "memory" of an artificial neural network. The output of the network varies depending on the connection method, weights, and activation function of the networks. In addition, back-propagation (BP) are commonly employed to decrease error. The suggested network design is depicted in a simplified form in Fig. 14a.

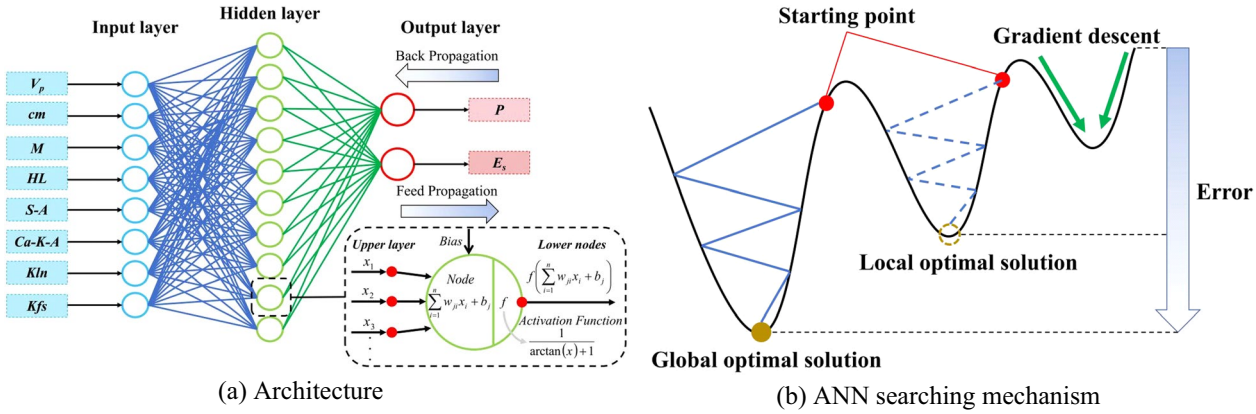
The training process is performed through a gradient descent algorithm on the error space that includes local minima as outlined in Fig. 14b. A network structure with 4 hidden layers is used during training. The sigmod function (or named as S-shaped function) is selected as a activation function $f_{ac}(x)$ of the hidden layer and output layer in this work, which can reflect the nonlinear relationship of the model. The specific expression is shown in Eq. (8):

$$f_{ac}(x) = \frac{1}{\arctan x + 1} \; or \; \frac{1}{e^{-x} + 1} \tag{8}$$

This sigmoid function has smooth continuity properties, which makes it very useful in optimization algorithms such as gradient descent. Its derivative is also easy to calculate, which is helpful for the application of backpropagation algorithm when training.

Performance criteria of ML models

Some statistical assessment criteria can be used to evaluate the performance of the selected ML models. Coefficient of determination (R^2) is a most commonly used criterion, which is evaluated by model prediction performance and an R-value more than 0.85 shows a significant connection. Some error-based criteria are also included in the evaluation methods, such as MAE (Mean



(a) Architecture (b) ANN searching mechanism

Fig. 14 Artificial neural network

Table 4 Performance criteria of ML models

Name of Criteria	Expression	Range
Coefficient of determination (R^2)	$R = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2(y_i-\bar{y})^2}}$	0–1
Coefficient of Efficiency (CE)	$CE = 1 - \frac{\sum_{i=1}^{N}(E_i-P_i)^2}{\sum_{i=1}^{N}(E_i-\bar{E}_i)^2}$	0–1
Root means Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(E_i-P_i)^2}$	0 Represents the best performance
Mean Absolute Error (MAE)	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert E_i-P_i\rvert$	0 Represents the best performance

Absolute Error) and RMSE (Root Mean Square Error) [54]. The related mathematical expressions are expressed in Table 4. Where N is the number of datasets, bar x and y are the mean of measured and predicted values respectively; E_i and P_i are the experimental in Sect. "Dataset description and analysis of weathering experiments of sandstone" and predicted values respectively.

Combined with inputs and outputs data in Sect. "Dataset description and analysis of weathering experiments of sandstone", the methodology adopted in this work to achieve the objective is presented in Fig. 15. In this section, four types of machine learning models will be used in 3.2 to form a surrogate model for predicting bulk porosity and static elastic modulus. In the Sect. "Predicted performances and comparative analyses", we evaluated the performance of the pretrained prediction model and selected the most suitable pretrained model.

Predicted performances and comparative analyses

Prediction of bulk porosity

The values of R^2, RMSE, CE and MAE comparison of best prediction models for bulk porosity are shown in Fig. 16. The variation of experimental data and machine learning models data with respect to data order are presented in Fig. 17. The R^2-value of this model is 0.9351, which is 5.38% and 3.08% lower than the GPR and ANN model respectively. The other performance indices such as RMSE, CE and MAE have values of 0.0039, 0.9408 and 0.0030 respectively. Similarly, the best fitted model was with quadratic support vector in SVR. In this model, the R^2-value is 0.9525, which is 3.45% and 1.20% less than the GPR and ANN model respectively. The other performance indices such as RMSE, CE and MAE have values of 0.0036, 0.9527 and 0.0024 respectively. Hence, GPR and ANN has shown good fitting performance using on training data, which have higher R^2-value and CE-value and lower RMSE and MAE.

When considering performance in testing process for prediction, there will exist significant changes in the predictive performance of different models for a new database, which may reflect negative performance such as overfitting of different pretrained models and generalization ability. Comparing the mentioned criteria such as fitting performances in each pretrained models, the R^2-value in testing process of MLR, SVR, GPR and ANN are 0.8883, 0.9137, 0.9089 and 0.9537, which is 5.00%, 4.07%, 7.76% and 1.06% lower than that in training process; the CE-value in testing process of MLR, SVR, GPR and ANN are 0.8966, 0.9154, 0.9126 and 0.9517, which is 4.70%, 3.92%, 6.50% and 0.84% lower than that in training process. It can be clearly seen that

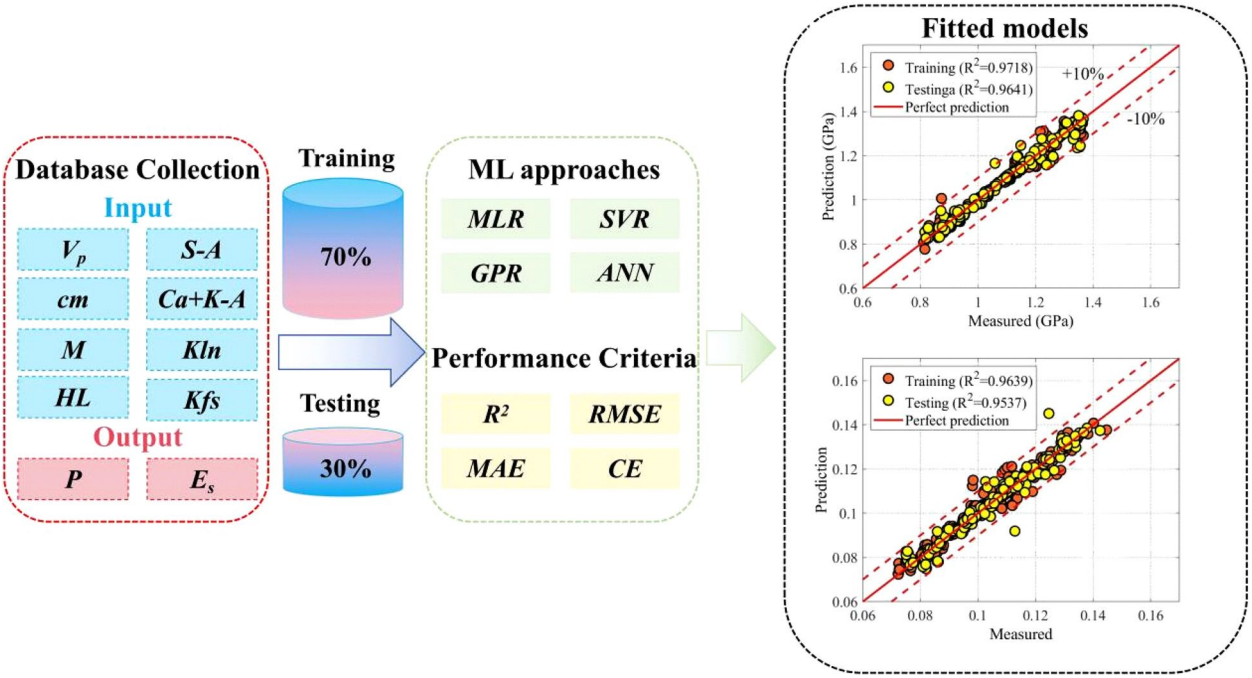


Fig. 15 The flowchart of machine learning prediction

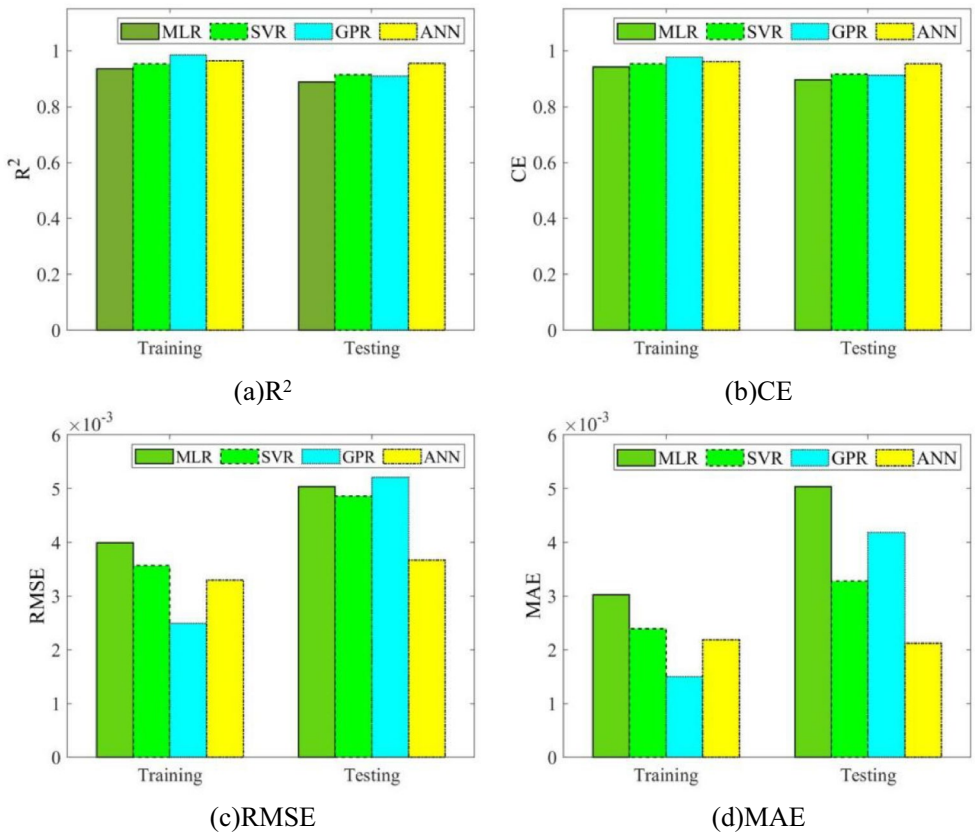
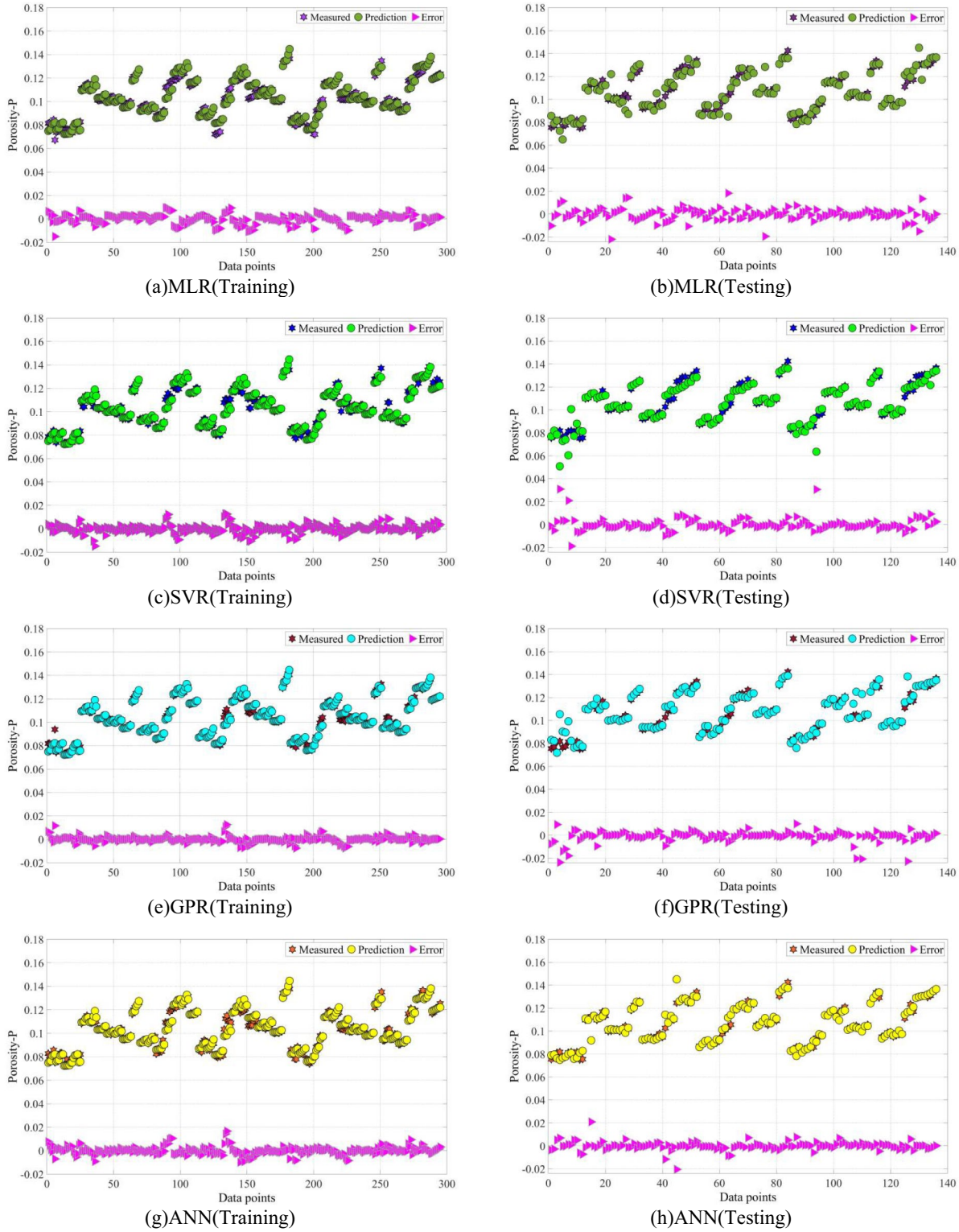


Fig. 16 R, RMSE, CE and MAE of fitted models for P



(a)MLR(Training) (b)MLR(Testing)

(c)SVR(Training) (d)SVR(Testing)

(e)GPR(Training) (f)GPR(Testing)

(g)ANN(Training) (h)ANN(Testing)

Fig. 17 Performances of P predicted by ML approaches

there exist varying degrees of overfitting effects in these three pretrained models based on MLR, SVR and GPR. It was indicated that the best fitted model in training process, that is GPR, may overfit the data when predicting the bulk porosity. And its R^2-value and CE-value in testing process are even lower than the worst model (MLR). But the values of two criteria of ANN in testing process are very closed to that in training process (which is about 1% deviation), indicating higher robustness of pretrained model.

For the error performances, the increasing absolute value can be used to measure the generalization ability of new data. RMSE in testing process is 0.0010, 0.0013, 0.0027 and 0.0004 higher than that in training process; MAE in testing process of MLR, SVR, GPR and ANN are 0.005037, 0.0032781, 0.0041855 and 0.0021249, which is 0.0020, 0.0009, 0.0027 and 0.0001 higher than that in training process. Similar to the trend of changes in fitting performance, the errors of GPR are the largest, indicating lower generalization ability of this pretrained model. The absolute values of two error criteria of ANN in testing process are very closed to that in training process (which are 0.0004 and 0.0001 respectively). indicating the adaptability of ANN to new data is much stronger than other three pre-trained models.

Figure 18 compares the predicted values of all 4 machine learning methods. At the same time, the perfect prediction line and parallel deviation lines (± 10%) are also provided. From the Fig. 18a, it can be seen that the main deviation of MLR occurs when the bulk porosity is greater than 0.1, and the probability of deviations are greater than 10% using the test dataset is 9.8%. This pretrained model may lead to an overestimation of the predicted bulk porosity. From the Fig. 18b, it can be seen that the main deviation of SVR occurs when the bulk porosity is less than 0.1, and the probability of deviations

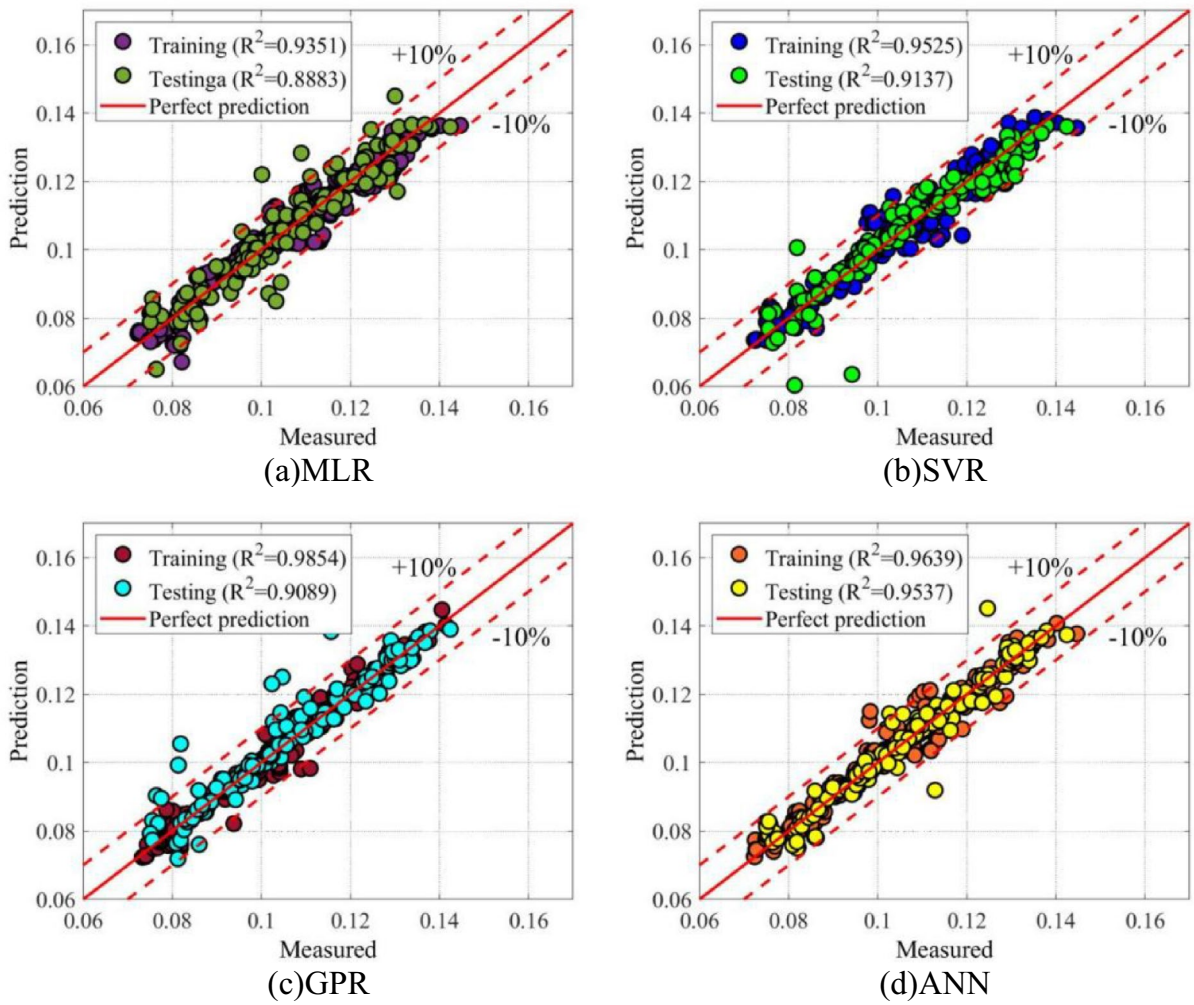


Fig. 18 The scatter plots of observed vs predicted values of the ML models for P in both stages

are greater than 10% using the test dataset is 3.8%. This pretrained model may lead to an underestimation of the predicted bulk porosity, but there exist high accuracy in weathered sandstone. From the Fig. 18c, it can be seen that the main deviation of GPR occurs when the bulk porosity is within 0.08–0.12, and the probability of deviations are greater than 10% using the test dataset is 9.1%. This pretrained model may lead to an overestimation of the predicted bulk porosity even in unweathered sandstone (low bulk porosity). From the Fig. 18d, it can be seen that the main deviation of ANN occurs when the bulk porosity is less than 0.11, and the probability of deviations are greater than 10% using the test dataset is only 2.3%.

Due to the complexity of microscopic structure and mineral composition of sandstone, the multiple detection indicators have significant variability in themselves. Hence, it should be noted that, generalization ability may have a much higher priority than accuracy for predicting the bulk porosity of grotto sandstone in weathering processes. Based on the analysis of the criteria, the most suitable machine learning model for predicting bulk porosity should be ANN, followed by SVR.

Prediction of static elastic *modulus*

The values of R^2, RMSE, CE and MAE comparison of best prediction models for static elastic modulus are shown in Fig. 19. The variation of experimental static elastic modulus data and machine learning models data with respect to data order are presented in Fig. 20. The R^2-value of this model is 0.9578, which is 2.55% and 1.46% lower than the GPR and ANN model respectively. The other performance indices such as RMSE, CE and MAE have values of 0.0261, 0.9635 and 0.0168 respectively. Similarly, the best fitted model was with linear support vector in SVR. In this model, the R^2-value is 0.9561, which is 2.72% and 1.64% less than the GPR and ANN model respectively. The other performance indices such as RMSE, CE and MAE have values of 0.0380, 0.9629 and 0.0297 respectively. Hence, GPR and ANN has shown good fitting performance using on training data, which have higher R^2-value and CE-value and lower RMSE and MAE. This training result is completely consistent with the prediction of bulk porosity. For pretrained models, the R^2-value in testing process of MLR, SVR, GPR and ANN are 0.9196, 0.9453, 0.9357 and 0.9641, which is 3.99%, 1.13%, 4.73% and 0.79% lower than that in training process; the CE-value in testing process of MLR, SVR, GPR and ANN

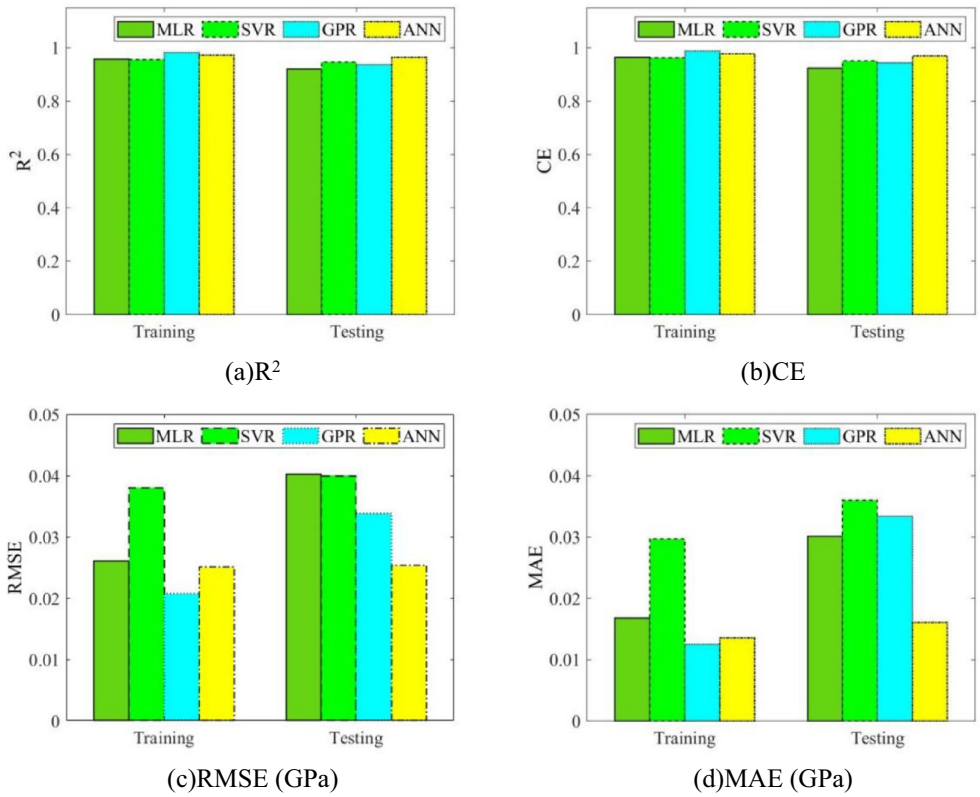
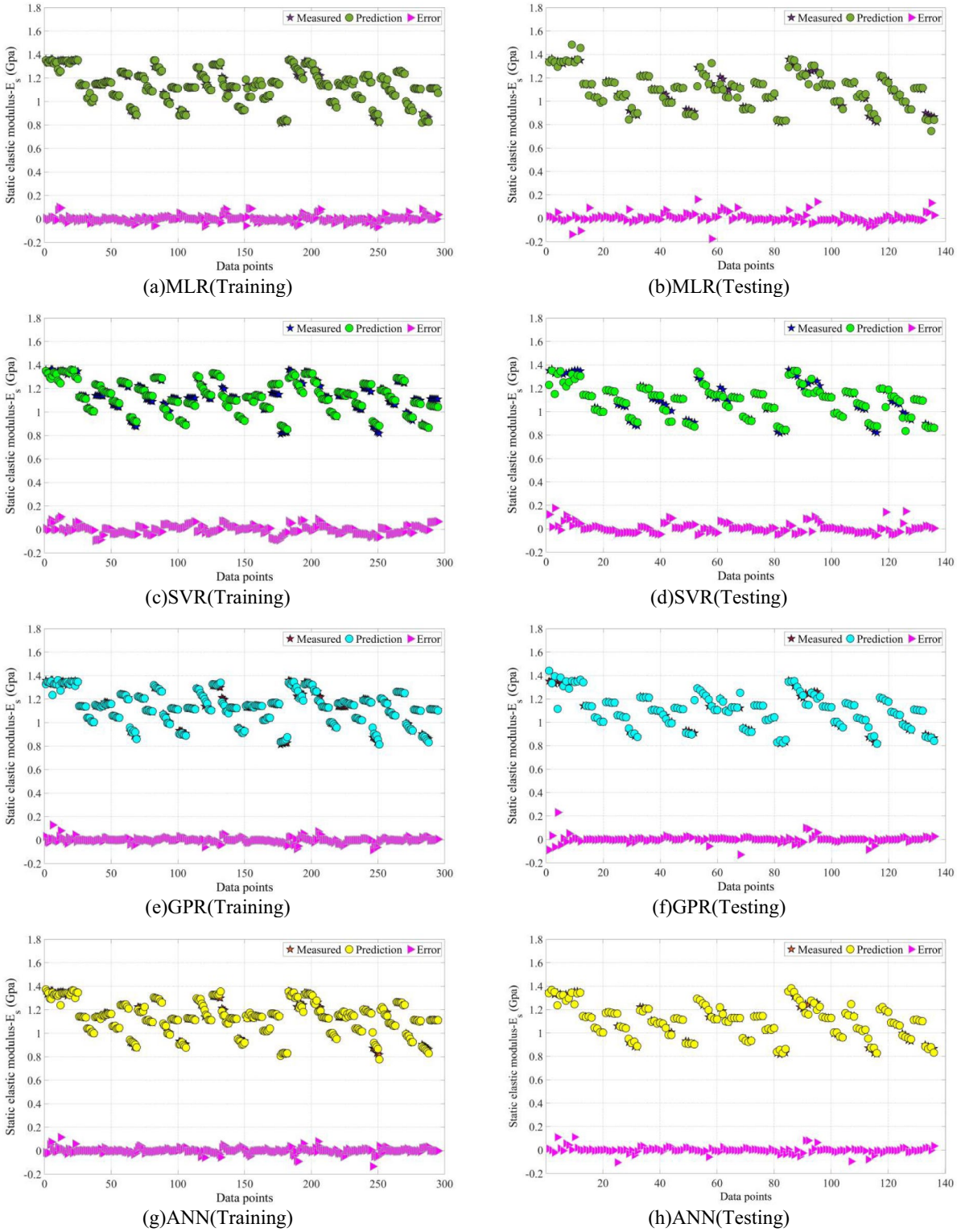


Fig. 19 R, RMSE, CE and MAE of fitted models for E_s



(a)MLR(Training) (b)MLR(Testing)

(c)SVR(Training) (d)SVR(Testing)

(e)GPR(Training) (f)GPR(Testing)

(g)ANN(Training) (h)ANN(Testing)

Fig. 20 Performances of Es predicted by ML approaches

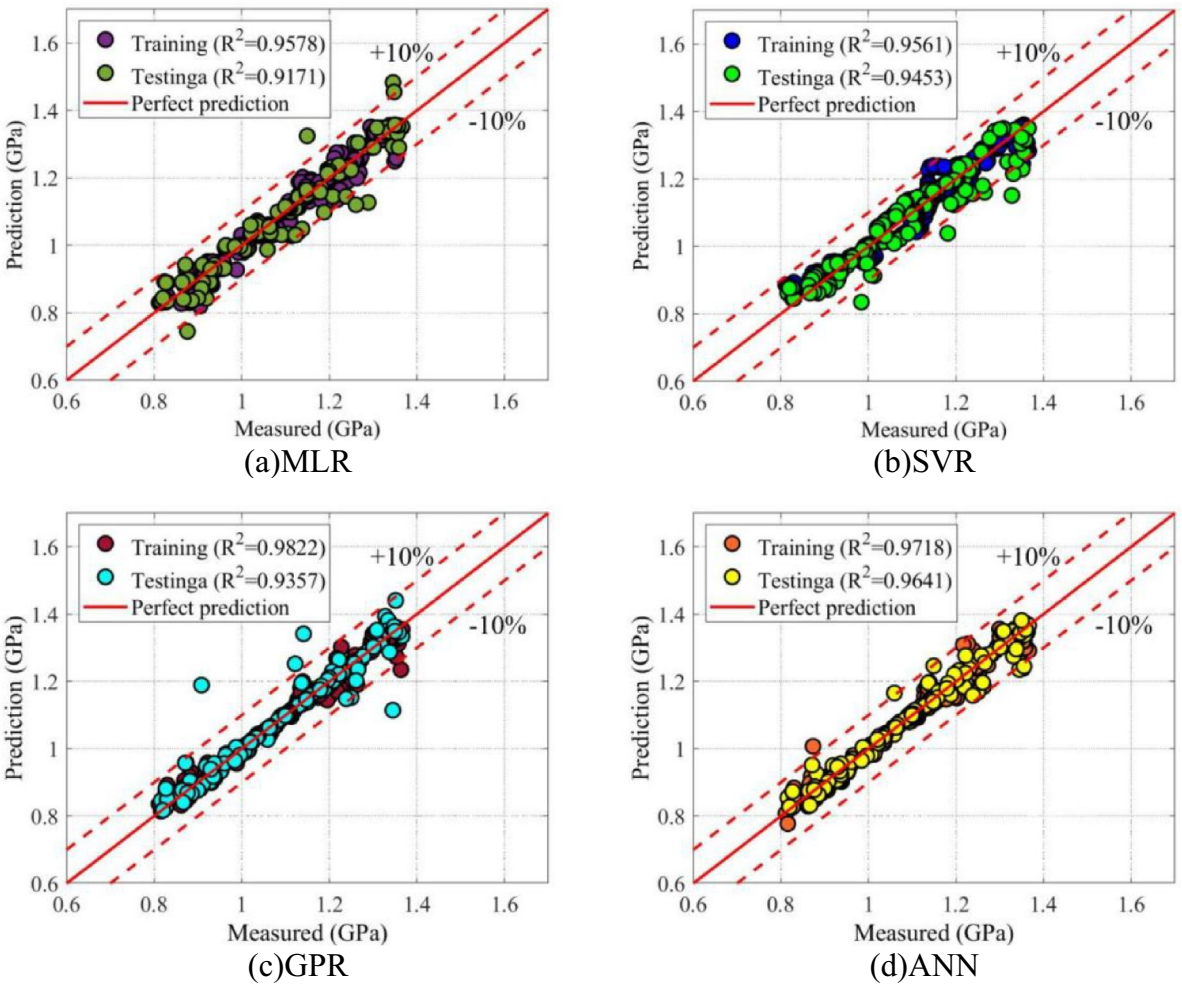


Fig. 21 The scatter plots of observed vs predicted values of the ML models for E_s in both stages

are 0.9226, 0.9505, 0.9431 and 0.9681, which is 4.24%, 1.29%, 4.46% and 0.83% lower than that in training process. There exist overfitting effects in pretrained models based on MLR and GPR. The values of R^2-value and CE-value of ANN and SVR in testing process are very closed to that in training process, indicating higher robustness of pretrained model.

For the error performances, RMSE in testing process of MLR, SVR, GPR and ANN are 0.0402, 0.0400, 0.0338 and 0.0253, which is 0.0141, 0.0020, 0.0131 and 0.0002 higher than that in training process; MAE in testing process of MLR, SVR, GPR and ANN are 0.0301, 0.0360, 0.0334 and 0.0161, which is 0.0133, 0.0063, 0.0209 and 0.0025 higher than that in training process.The errors of MLR and GPR are the large, indicating lower generalization ability of this pretrained model. The absolute values of two error criteria of ANN in testing process are very closed to that in training process (which are 0.0002 and 0.0025 respectively), indicating the adaptability of ANN to new data is much stronger than other three pre-trained models.

Figure 21 compares the predicted values of all 4 machine learning methods for prediction of static elastic modulus. From the Fig. 21a–c, it can be seen that the main deviation of MLR, SVR and GPR occur when the static elastic modulus are greater than 0.85 MPa, 0.99 MPa and 0.91 MPa respectively, and the probability of deviations are greater than 10% using the test dataset are 5.3%, 6.1% and 4.6% respectively. From the Fig. 21d, it can be seen that the main deviation of ANN occurs when the bulk porosity is greater than 1.04 MPa, but its deviations are almost within 10%. Due to the fact that the static elastic modulus of sandstone is the most macroscopic indicator, the decrease of its value can also be served as the final performance of weathered sandstone. Its decreasing patterns are related to changes in

weathering existence environment, mineral composition, and pore structure. At the main time, it can be seen that several regression algorithms have improved the accuracy of predicting elastic modulus to varying degrees compared to predicting bulk porosity. Compared to on-site evaluation of bulk porosity, it's more vital to obtain a more accurate and generalizable pretrained model for static elastic modulus. Based on the analysis of 4 criteria, the most suitable machine learning model for predicting bulk porosity should be ANN.

The comparative analyses and limitations of models for prediction weathering degrees of Yungang Grottoes sandstone

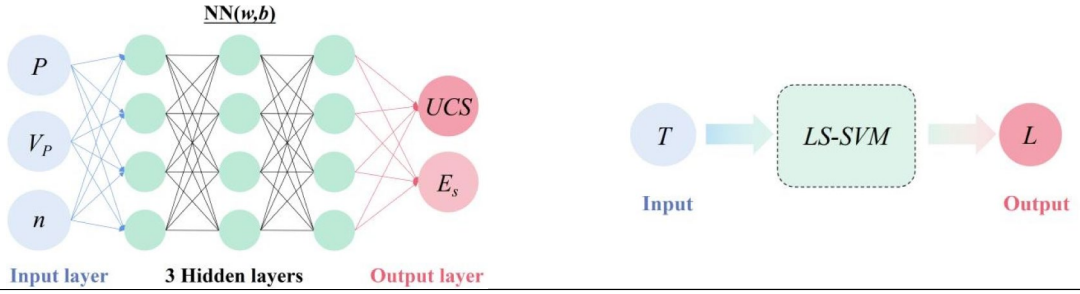
The prediction performance of the two ANN models was evaluated in the previous subsections. To further clarify the practical applicability of these models in the field, we compare The most distinct features of the models from Sects. "Prediction of bulk porosity" and "Prediction of static elastic modulus" with the current surrogate models used for predicting sandstone deterioration in the Yungang Grottoes. Other researchers have also proposed efficient and accurate models for predicting the deterioration of Yungang sandstone. For instance, Liu et al. [55] proposed a three-layers neural network model (named 3-NNs) for predicting the uniaxial compressive strength and elastic modulus of Yungang sandstone, and Meng et al. [34] proposed a least squares-support vector machine (named LS-SVM) for predicting the weathering levels of Yungang sandstone. The dataset for Liu's model is derived from sandstone samples subjected to different freeze–thaw cycles, using compressive wave velocity, bulk porosity, and the number of cycles as inputs to predict mechanical parameters of the sandstone such as uniaxial compressive strength and elastic modulus. The architecture of Liu's model is illustrated below (Fig. 22).

where, n represents the number of freeze–thaw cycles, and UCS is the uniaxial compressive strength. The dataset of Meng's model is derived from different caves in Yungang Grottoes, using the transitivity of terahertz waves T as input to predict the weathering level of the caves. The weathering level L is determined by the compressive wave velocity, as follows:

$$L = 10 - INT\left(10\frac{V_p}{V_{p0}}\right) \tag{9}$$

where, the INT function rounds the object of the function up to the nearest integer, Vp_0 is the compressive ultrasonic velocity of a fresh sandstone core. For the 3-NNs model, bulk porosity is easily detectable due to the use of laboratory data. Therefore, bulk porosity is used as input in this model to predict other destructive detection indicators, namely uniaxial compressive strength and elastic modulus. However, in fact, the bulk porosity of the immovable stone cultural heritages at the Yungang Grottoes in-site are quite difficult to be detected. Meanwhile, the number of cycles n is a known parameter in the simulated weathering experiments, which cannot be measured in-site. Meanwhile, porosity is a direct quantification of the degree of weathering and should be used as an output (predicted value). The predicted value of LS-SVM is actually the compressive wave velocity [weathering level L in Eq. (9)], which is not a direct quantitative indicator representing the degree of weathering in grottoes, but an indirect indicator. Hence, the input of the ANN prediction model for bulk porosity and static elastic modulus proposed in this article are some convenient or collectable indicators in-site, such as wave velocity, leeb hardness, relative content of chemical elements, etc., with the aim of rapid and easy-to-use application; At the same time, this article used various conditions (environments) to simulate weathering, and the models have stronger generalization ability.

However, in reality, there still exist some limitations in proposed ANN models. Firstly, there are usage restrictions: the applicable range of bulk porosity is between 7.8 and 13.3%, and the static elastic modulus is between 0.79 to 1.37 Gpa. The applicable lithology



(a)3-NNs proposed by Liu **[57]** (b)LS-SVM proposed by Meng **[36]**

Fig. 22 Architecture of proposed ML models for predicting the weathering levels of Yungang sandstone

is only for sandstone in Yungang area (Datong, China). Moreover, it is important to note that the training data for this model is derived from equivalent weathering simulation experiments rather than actual weathering data due to the difficulty in obtaining the bulk porosity and static elastic modulus of the Yungang Grottoes. This could result in the model potentially being unsuitable for all actual weathering scenarios of the Yungang Grottoes. Additionally, the ML model (ANN) used is a classic general-purpose model, and it may be necessary to explore advanced AI baseline models that are more suited to weathering data. To obtain a more precise model capable of prediction of bulk porosity and static elastic modulus, we need sandstone samples with more weathering gradients and a denser testing.

In summary, the limitations affecting the model's performance mentioned above could be addressed in the future by expanding the weathered sandstone dataset and the scale of model training, leading to the creation of a more versatile and accurate model, which can be utilized by researchers or cultural heritage conservators. By collecting non-destructive indicators from weathered areas of grotto heritage, researchers can obtain more detailed and accurate weathering information. This contributes to further grotto heritage conservation and informed decision-making.

Conclusion

Four mainstream machine learning techniques are used to evaluate the bulk porosity and static elastic modulus of weathered grottoes sandstone. Datasets are gathered from the experiments, which includes 432 groups effective experimental data including 8 inputs features. bulk porosity and static elastic modulus were considered as outputs to determine the weathering degrees of Yungang Grottoes sandstone. The performance parameters R, RMSE, CE and MAE were used to evaluate the performance of the ANN, GPR, MLR, SVR machine learning models. The following are the quantitative outcomes of this study:

The ANN is the best-fitted models for estimating the bulk porosity and static elastic modulus compared to MLR, SVR and GPR. The accuracy of the trained model for static elastic modulus is slightly higher than that of bulk porosity. The GPR and ANN model can accurately predict the bulk porosity with R-value 0.9854 and 0.9639 for training stages and respectively and predict the static elastic modulus with R-value 0.9822 and 0.9718 for training stages respectively. The modeling results obtained reveal that the ANN with multi-hidden layers developed is competent with high degree of precision and generalization ability for bulk porosity and static elastic modulus compared to other regression-based models (MLR, SVR, and GPR), as indicated by coefficient of determinations in the range of (0.9537–0.9641) during the testing stages of the built models compared to (0.8883–0.9453) coefficient of determination in regression-based models. The ANN model prediction efficiency was well adjusted for the actual and forecast datasets at the training and testing stages, and the error range was no more than 0.7% and 0.15 GPa at both stages of prediction for bulk porosity and static elastic modulus respectively.

Notably, data points in scatter plots of observed vs predicted values simulated by the ANN model were well adjacent to the linear 1:1 line, which was more than 95% and 97% of the data points in prediction for bulk porosity and static elastic modulus respectively. And the ANN based static elastic modulus prediction model's error proportions significantly decreased and were confined to a modest range between + 10% and − 10%. The proposed models are only valid for the bulk porosity ranging from 7.8 to 13.3% and the static elastic modulus ranging from 0.79 to 1.37 Gpa, which can be utilized for the accurate prediction of the weathering degrees of shallow surface layer in Yungang Grottoes sandstone.

Some indicators that can be obtained in site (acoustic emission based wave velocity, hardness, chemical element content, etc.) are used to predict indicators that cannot be directly measured, such as static elastic modulus and bulk porosity mentioned in this work. When the model is trained, static elastic modulus and bulk porosity can be obtained directly by inputing multiple portable indicators of stone cultural heritage in site. In the near future, we will focus on model training for data from other rock types, seeking more detailed weathering related indicators (e.g., moisture content, mechanical properties, mineral content, etc.) and linking them with non-destructive testings (e.g., high-density electrical method, hyperspectral technology, X-ray fluorescence, etc.) quantitatively.

Acknowledgements

We'd like to express our appreciation to Prof. Jizhong Huang for his project support (Science and Technology Major Special Program Project of Shanxi Province), and also express our gratitude to reviewers of Heritage Science and Yue Zhang (Associate professor) for their vital comments to enhance this work. This work is also supported by Shanghai Technical Service Center of Science and Engineering Computing, Shanghai University.

Author contributions

RY.Z worked in methodology, conceptualization, software, investigation, writing—original draft and writing—review and editing; JZ.H worked in supervision, funding acquisition, investigation, conceptualization, validation and writing—review and editing; Y.C worked in funding acquisition, conceptualization, methodology, investigation, validation and writing—review and editing; Y.Z worked in methodology, investigation, validation andwriting—review and editing.

Funding

This research was supported by Science and Technology Major Special Program Project of Shanxi Province (202201150501024), Research Project of

Shanxi Cultural Heritage Bureau (2024KT19) and National Natural Science Foundation of China (42007265).

Availability of data and materials
The data used to support the findings of this study are included within the article.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Mechanics and Engineering Science, Shanghai University, Shanghai 200444, China. [2]Institute for the Conservation of Cultural Heritage, School of Cultural Heritage and Information Management, Shanghai University, Shanghai 200444, China. [3]Key Laboratory of Silicate Cultural Heritages Conservation (Shanghai University) Ministry of Education, Shanghai 200444, China.

Received: 5 July 2024 Accepted: 25 September 2024
Published online: 15 October 2024

References

1. Min Liu JC, Xiao LZ, Xie ZB. Protection of stone cultural heritages in China: analysis of NSFC-funded projects. Sci Conserv Archaeol. 2019;31:112–9.
2. Sun B, Peng N, Fan Y, Zhang H, Wang F. Impact of rock matrix seepage on hollowing and cracking of surface restoration layer in the Leshan Giant Buddha. Int J Archit Herit. 2023. https://doi.org/10.1080/15583058.2023.2284746.
3. Yi Y, Chen Y. An analysis of the statistics on major historical and cultural sites protected at the national level. Southeast Cult. 2021;4:6–15.
4. Wang JH, Chen JQ. Current status and future development of cave temples protection in China. Southeast Cult. 2018;1:6–14+127-128.
5. Hong J, Zhu Y, Zhang Y, Huang J, Peng N. Differentiation study of the damage characteristics of rock cultural heritage sites due to the sulfate weathering process. Appl Sci. 2023;13(23):12831.
6. Steiger M, Asmussen S. Crystallization of sodium sulfate phases in porous materials: the phase diagram Na2SO4·H2O and the generation of stress. Geochim Cosmochim Acta. 2008;72:4291–306.
7. Oguchi CT, Yu S. A review of theoretical salt weathering studies for stone heritage. Prog Earth Planet Sci. 2021;8:32–55.
8. Holynska B, Gilewicz-Wolter J, Ostachowicz B, Bielewski M, Streli C, Wobrauschek P. Study of the deterioration of sandstone due to acid rain and humid SO2 gas. X-Ray Spectrom. 2010;33:342–8.
9. Zhang Y, Zheng Yi, Huang J. Determination of water vapor transmission properties of sandstones in the Yungang Grottoes. Int J Archit Herit. 2024;18(3):357–69.
10. Zhang Y, Cao C, Du H, Huang J, Guo X, Luo Q, Ren J. Investigation into the Gaseous SO2 Attack on Sandstone in the Yungang Grottoes. Minerals. 2023;13(1):123.
11. Zhang Y. Knowledge of earthen heritage deterioration in dry areas of China: salinity effect on the formation of cracked surface crust. Herit Sci. 2023;11:41.
12. Zhang Y, Zhang Y, Huang J. Experimental study on capillary water absorption of sandstones from different grotto heritage sites in China. Herit Sci. 2022;10:25.
13. Huang J, Zheng Yi, Li H. Study of internal moisture condensation for the conservation of stone cultural heritage. J Cult Herit. 2022;56:1–9.
14. Wang F, Huang J, Zhao H. Mechanical sandstone deterioration due to cement binder material materials under dry-wet cycling. Case Stud Constr Mater. 2023;18: e02169.
15. Li L, Wang SJ, Tanimoto C. Study of weathering characteristics of sandstone at longyou grottoes. Chin J Rock Mech Eng. 2008;6:1217–22.
16. Martinez M, Benavente D, Garcia M. Spatial attenuation: the most sensitive ultrasonic parameter for detecting petrographic features and decay processes in carbonate rocks. Eng Geol. 2011;119:84–95.
17. Theoulakis P, Moropoulou A. Microstructural and mechanical parameters determining the susceptibility of porous building stones to salt decay. Constr Build Mater. 1997;11:65–71.
18. Becherini F, Bernardi A, Frassoldati E. Microclimate inside a semi-confined environment: valuation of suitability for the conservation of heritage materials. J Cultur Herit. 2010;11:471–6.
19. Camuffo D. Microclimate for cultural heritage. Amsterdam: Elsevier; 1998.
20. Liu RZ, Zhang BJ, Zhang H, Shi MF. Deterioration of Yungang Grottoes: diagnosis and research. J Cult Herit. 2011;12:494–9.
21. Weng LQ, Yang HF, Wang FR, Song SH. Research on characterization of sandstone weathering of Yungang Grottoes. Mater Rev. 2011;25:425.
22. Vergès-Belmin V. Illustrated glossary on stone deterioration patterns. Paris: ICOMOS-ISCS; 2008.
23. André M-F, Voldoire O, Roussel E, Vautier F, Phalip B, Peou H. Contrasting weathering and climate regimes in forested and cleared sandstone temples of the Angkor region. Earth Surf Process Landf. 2012;37:519–32.
24. André MF, Phalip B, Voldoire O, et al. Weathering of sandstone lotus petals at the Angkor site: a 1000-year stone durability trial. Environ Earth Sci. 2011;63:1723–39.
25. Liu RZ, Zhang BJ, Zhang H, Shi MF. Deterioration of Yungang Grottoes: diagnosis and research. J Cult Herit. 2011;12(4):494–9.
26. Mammoliti E, Cupido M, Teloni R, et al. Implementation of a non-destructive method to assess weathering deterioration of sandstones in cultural heritage. Bull Eng Geol Environ. 2024;83:110.
27. Egartner I, Schnepfleitner H, Sass O. Influence of salt and moisture on weathering of historic stonework in a continental-humid, urban region. Eng Geol Soc Territ. 2015;8:569–75.
28. Meng T, Lu Y, Zhao G, et al. A synthetic approach to weathering degree classification of stone heritages case study of the Yungang Grottoes. Herit Sci. 2018;6:1.
29. Basu A. Applicability of weathering classification to quartzitic materials and relation between mechanical properties and assigned weathering grades: a comparison with investigations on granitic materials. Bull Eng Geol Environ. 2015;6:865–8.
30. Ceryan S, Tudes S, Ceryan N. A new quantitative weathering classification for igneous rocks. Environ Geol. 2008;55(6):1319–36.
31. Fehmi A, Resat U, Nihal A. Characterization of weathered acidic volcanic rocks and a weathering classification based on a rating system. Bull Eng Geol Environ. 2007;66(4):415–30.
32. Hu R, Oyediran IA, Gao W, Zhang XY, Li LH. "Plagioclase solution degree index": a new index to evaluate the weathering degree of granite. Bull Eng Geol Environ. 2014;73(2):589–94.
33. Marszałek M, Alexandrowicz Z, Rzepa G. Erratum to: composition of weathering crusts on sandstones from natural outcrops and architectonic elements in an urban environment. Environ Sci Pollut R. 2015;22(6):4786–7.
34. Meng T, Huang R, Lu Y, et al. Highly sensitive terahertz non-destructive testing technology for stone heritages deterioration prediction using SVM-based machine learning models. Herit Sci. 2021;9:24.
35. Kuchař M, Oswald C, Fišer J, et al. Corrosion prediction for preventive protection of aircraft heritage. Herit Sci. 2024;12:102.
36. Boesgaard C, Hansen BV, Kejser UB, et al. Prediction of the indoor climate in cultural heritage buildings through machine learning: first results from two field tests. Herit Sci. 2022;10:176.
37. Yan L, Chen Y, Zheng L, et al. Application of computer vision technology in surface damage detection and analysis of shedthin tiles in China: a case study of the classical gardens of Suzhou. Herit Sci. 2024;12:72.
38. Bewes J, Low A, Morphett A, Pate FD, Henneberg M. Artificial intelligence for sex determination of skeletal remains: application of a deep learning artificial neural network to human skulls. J Forens Leg Med. 2019;62:40–3.
39. Cintas C, Lucena M, Fuertes JM, Delrieux C, Navarro P, Gonzalez-Jose R, Molinos M. Automatic feature extraction and classification of Iberian ceramics based on deep convolutional networks. J Cult Herit. 2020;41:106–12.

40. Wang N, Zhao X, Zhao P, Zhang Y, Zou Z, Ou J. Automatic damage detection of historic masonry buildings based on mobile deep learning. Autom Construct. 2019;103:53–66.
41. Wang N, Zhao X, Zou Z, Zhao P, Qi F. Autonomous damage segmentation and measurement of glazed tiles in historic buildings via deep learning. Comput Aided Civ Infrastruct. 2020;35(3):277–91.
42. Hatır ME, Barstugan M, İnce İ. Deep learning-based weathering type recognition in historical stone monuments. J Cult Herit. 2020. https://doi.org/10.1016/j.culher.2020.04.008.
43. Hatır ME, İnce İ. Lithology mapping of stone heritage via state-of-the-art computer vision. J Build Eng. 2020. https://doi.org/10.1016/j.jobe.2020.101921.
44. Hatır ME, İnce İ, Korkanç M. Intelligent detection of deterioration in cultural stone heritage. J Build Eng. 2021. https://doi.org/10.1016/j.jobe.2021.102690.
45. Hatır E, Korkanç M, Schachner A, İnce İ. The deep learning method applied to the detection and mapping of stone deterioration in open-air sanctuaries of the Hittite period in Anatolia. J Cult Herit. 2021. https://doi.org/10.1016/j.jobe.2021.102690.
46. Ulusay R, Hudson JA. 2007. The complete ISRM suggested methods for rock characterization, testing and monitoring: 1974–2006. Prepared by the Commission on Testing Methods, International Society for Rock Mechanics: ISRM Turkish National Group, Ankara. p. 628.
47. Udagedara DT, Oguchi CT, Gunatilake AAJK. Combination of chemical indices and physical properties in the assessment of weathering grades of sillimanite-garnet gneiss in tropical environment. Bull Eng Geol Environ. 2017;76:145–57.
48. Mallory Greenough LM, Gorton MP, Greenough JD. The source of basalt vessels in ancient Egyptian archeological sites: a mineralogical approach. Can Mineral. 2002;40(Part 4):1025–46.
49. Kumar A, Arora HC, Kumar K, Mohammed MA, Majumdar A, Khamaksorn A, Thinnukool O. Prediction of FRCM–concrete bond strength with machine learning approach. Sustainability. 2022;14(2):845.
50. Pappalardo G, Mineo S. Static elastic modulus of rocks predicted through regression models and Artificial Neural Network. Eng Geol. 2022;308:106829.
51. Mur R, Díaz I, Rodríguez M. Comparative study of surrogate modelling techniques applied to three different chemical processes. Comput Aided Chem Eng. 2020;48:145–223.
52. Richardson RR, Osborne MA, Howey DA. Gaussian process regression for forecasting battery state of health. J Power Sources. 2017;357:209–19.
53. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys. 1943;5:115–33.
54. Chaabene WB, Flah M, Nehdi ML. Machine learning prediction of mechanical properties of concrete: critical review. Constr Build Mater. 2020;260:119889.
55. Liu C, Liu Y, Ren W, et al. An adaptive prediction method for mechanical properties deterioration of sandstone under freeze–thaw cycles: a case study Yungang Grottoes. Herit Sci. 2021;9:154.

Publisher's Note

Ruoyu Zhang PhD candidate in Institute for the Conservation of Cultural Heritage, Shanghai University.

Jizhong Huang Professor in Institute for the Conservation of Cultural Heritage, Shanghai University.

Yuan Cheng Lecturer in Institute for the Conservation of Cultural Heritage, Shanghai University, 99 Shangda Road, Shanghai, PR China, 200444.

Yue Zhang Associate professor in Institute for the Conservation of Cultural Heritage, Shanghai University.