



City Research Online

City St George's, University of London

Citation: Yu, A., Liu, T., Miao, T., Chen, X., Deng, X. & Fu, F. (2025). Signal Recognition and Prediction of Water-Bearing Concrete Under Axial Compression Using Acoustic Emission and Machine Learning. *Structural Control and Health Monitoring*, 2025(1), 6633988. doi: 10.1155/stc/6633988

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/35608/>

Link to published version: <https://doi.org/10.1155/stc/6633988>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Research Article

Signal Recognition and Prediction of Water-Bearing Concrete Under Axial Compression Using Acoustic Emission and Machine Learning

Aiping Yu ^{1,2} Tao Liu ^{1,2} Tianjiao Miao ^{1,2} Xuandong Chen ^{1,2} Xuelian Deng ^{1,2} and Feng Fu ³

¹School of Civil Engineering, Guilin University of Technology, Guilin 541004, China

²Guangxi Key Laboratory of Green Building Materials and Construction Industrialization, Guilin 541004, China

³Department of Engineering, School of Science & Technology, City, University of London, Northampton Square, London EC1V 0HB, UK

Correspondence should be addressed to Feng Fu; feng.fu.1@city.ac.uk

Received 25 April 2025; Revised 27 May 2025; Accepted 15 July 2025

Academic Editor: Young-Jin Cha

Copyright © 2025 Aiping Yu et al. Structural Control and Health Monitoring published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The presence of free water in the concrete slurry significantly influences the crack patterns of concrete. In this study, uniaxial compression tests were conducted on concrete specimens with varying moisture contents under acoustic emission (AE) monitoring. Through parametric analysis and machine learning, the cracking process of water-containing concrete was studied, signal patterns during the cracking process were identified, and the impact of moisture content on the damage evolution and fracture mechanism of concrete was understood. The results indicate that free water is capable of absorbing high-frequency signals. With the increase of moisture content, the AE signals decrease. The failure of concrete is mainly of the tensile type, while the shear-type accounts for a relatively small proportion. The presence of free water decreases the likelihood of diagonal shear failure in concrete structures. The unsupervised learning was used for various moisture content analyses. Three distinct AE signal patterns were identified during the concrete compression tests: frictional motion signals of the compression surface, fracture surface activity signals, and aggregate cracking signals. Based on the moisture content, this study analyzes the variations in signal responses across different modes. A predictive model was established utilizing the BP neural network to differentiate signals of various modes, achieving an accuracy rate of 99%.

Keywords: acoustic emission; BP neural network; clustering algorithm; concrete; moisture content; parameter characteristics; RA-AF

1. Introduction

Various large-scale engineering structures, including but not limited to piers, sluices, reservoirs, and tunnels, are frequently exposed to high humidity environments during their service life [1–6]. Prolonged exposure to high humidity environment significantly impacts the strength, stiffness, and durability of concrete [7–9]. With the increase in moisture content, the compressive strength of concrete will decrease, which can easily lead to failures, such as pier

failure, reservoir collapse, and tunnel collapse. Therefore, it is imperative to assess and characterize the damage mechanisms of concrete under varying moisture conditions. This evaluation holds significant importance for maintaining structural integrity and accurately assessing the service life of the structure.

The moisture content exerts a significant influence on the mechanical properties of concrete materials. The static compressive strength of saturated concrete is lower compared to that of dry concrete [8]. As the moisture content

increases, the cracking stress and compressive strength initially rise before decreasing. In addition, the presence of water also reduces the shear strength of concrete [9, 10]. Studies have demonstrated that the compressive strength of concrete is influenced by the displacement rate when fully hydrated under axial loading [11]. Compared to dry specimens subjected to the same displacement rate, the strength of concrete with high moisture exhibits a significant reduction. The moisture content significantly influences the durability of concrete. Specifically, higher moisture content is directly correlated with increased shrinkage and inversely correlated with workability [7]. When concrete is subjected to compressive stress, its elastic modulus increases with higher moisture content; however, the presence of internal voids adversely affects the mechanical properties of concrete [12]. Under fatigue loading, the fatigue life of concrete in a saturated state is significantly lower compared to that in a dry state. In addition, both the strain rate and dissipated energy of concrete during the fatigue stage increase as moisture content and stress levels rise [13]. From the perspective of damage rheological theory, it has been observed that moisture content exerts a significant influence on concrete creep [14]. Zhou et al. [15] investigated the coupling effects of water saturation and dip angle on the mechanical properties and failure characteristics of the rock–concrete interface. They discovered that the extent of strength degradation in rock–concrete disc specimens was significantly correlated with the ratio of shear stress to tensile stress at the interface. Fahim et al. [16] developed a finite element model to predict the drying behavior and internal relative humidity of concrete. These studies have elucidated the impact of moisture content on various mechanical properties of concrete. However, the impact of moisture content on various failure mechanisms of concrete remains poorly understood, primarily due to the challenges in accurately identifying these mechanisms.

Acoustic emission (AE) is a real-time and efficient nondestructive testing technology that has gained widespread application in the damage assessment and health monitoring of concrete structures. Under the action of external force or internal force, the internal deformation or fracture of materials releases energy and generates instantaneous elastic waves, which are called AE [17]. AE monitoring does not induce any damage to the materials during the assessment process and possesses the capability for quantitative evaluation of damage [18]. Evaluating the damage of engineering structures through AE technology and promptly proposing corresponding measures for any detection issues are effective methods to mitigate engineering accidents. Ohno et al. [19], Gao and Sun [20], and Prem et al. [21] conducted AE monitoring during the fracture process of concrete and reinforced concrete structures, revealing a strong correlation between AE parameters and the fracture process. Yu et al. [22] and Chen et al. [23] conducted uniaxial compression and AE monitoring tests on concrete specimens with varying water–cement ratios. They developed a concrete damage evaluation model based on cumulative AE energy and cumulative AE ringing counts and introduced a combined active–passive

AE testing method that accounts for the heterogeneity of concrete. The reliability of the proposed method has been validated through a step loading test, utilizing AE wave velocity and amplitude parameters. Chen and Liu [24] investigated the influence of maximum aggregate particle size on the fracture properties and fracture process zone at the crack tip in high-performance concrete, characterizing these effects through AE hit counts. To investigate the characteristics of AE parameters during the failure process of underground concrete infrastructure, Lee et al. [25] conducted uniaxial compression tests on concrete specimens prepared with varying sand ratios. By analyzing both time-domain parameters (such as ringing counts, energy, and amplitude) and frequency-domain parameters (including initial frequency and peak frequency), it was determined that cumulative energy parameters provide the most effective characterization of the failure process in the specimens. Furthermore, AE technology has been utilized for the classification of shear and tension cracks during the fracture process of concrete structures [26], as well as for monitoring various materials and structures including rocks [27], coal samples [28], and railway substructures [29].

Machine learning (ML) is a subfield of artificial intelligence that enables systems to improve and optimize their performance autonomously by learning from data and recognizing patterns. ML algorithms leverage extensive datasets to identify underlying patterns through training and optimization processes, thereby facilitating accurate predictions or informed decisions. The training of ML models necessitates a substantial volume of data. AE technology, when employed to monitor the material failure process, generates a dataset comprised of parameters that can be used in ML. Pattern recognition of AE signals from various building materials has been made in the application of ML in structural health monitoring. Liu et al. [30] investigated AE signals generated during alloy deformation through an integrated approach that combined supervised and unsupervised ML techniques. Heidary et al. [31] employed unsupervised pattern recognition analysis, specifically fuzzy C-means clustering related to principal component analysis (PCA), to identify the damage modes of composite materials at various stages of drilling. Li et al. [32] successfully classified AE signals during the stress corrosion of 304 stainless steel using the k-means clustering algorithm, achieving promising results. Soltangharai et al. [33] employed a hierarchical clustering algorithm to classify AE signals from cement mortar column specimens under monotonic loading conditions and identified the failure modes based on the characteristics of these signals. Radhika et al. [34] performed a parametric analysis of AE data obtained from 12 beam specimens subjected to three-point bending tests. Through the application of PCA and k-means clustering algorithms, they identified four distinct damage mechanisms in concrete under fatigue loading conditions. In conclusion, ML has been extensively applied in the recognition of material damage, offering a novel and effective approach for identifying damage patterns in concrete structures.

It is important to highlight that previous studies on the identification of concrete damage modes have seldom taken

into account the influence of water content within internal voids on the damage mode. As a constituent within the concrete medium, free water possesses specific functionalities and characteristics that play a significant role during the process of concrete failure. Water exerts a significant influence on the various damage modes of concrete. As the moisture content varies, distinct damage modes are expected to exhibit different patterns. It is viable to determine the damage modes of concrete structures utilizing ML models. In this study, the fracture process of concrete specimens with varying moisture content was monitored using AE techniques, and the influence of free water within voids on the damage and fracture evolution of concrete was investigated. The AE signal patterns during the failure process of concrete are identified using a combination of moisture content analysis and an unsupervised learning algorithm. Furthermore, the impact of varying moisture content levels on distinct AE signal patterns is examined. Utilizing the BP neural network, we have developed prediction models for various signal patterns. This study offers robust theoretical support and comprehensive safety guidelines for the health monitoring of concrete structures operating in aquatic or heavy rainfall environments.

2. Experiment

2.1. Sample Design. In the tests, 15 concrete specimens with a water-to-cement ratio of 0.58 were fabricated, and each specimen had dimensions of $150 \times 150 \times 150$ mm. The specific mix proportions are detailed in Table 1. The cement utilized is standard Portland cement, while the fine aggregate consists of medium sand. The coarse aggregate comprises macadam with a maximum particle size of 20 mm. Cement, sand, and gravel are all produced in Guilin Xiangjiu Sand Field. All samples were fabricated and cured at the same time period under standard conditions for 28 days following demolding. Upon completion of the curing period, the specimens were air-dried in a well-ventilated environment for 1 month, followed by artificial drying.

The test samples were categorized into five groups, with each group undergoing three replicate trials. Following natural drying, the samples were placed in a drying oven set at 105°C for 24 h. Subsequently, the dried samples were labeled and weighed, allowed to equilibrate to room temperature, and sealed with plastic film. Thereafter, the samples were subjected to soaking.

To control the moisture content of the samples, they were categorized into five groups, each comprising three samples. The grouped samples were submerged in water. The control samples were subjected to soaking durations of 6, 12, 24, and 48 h. The samples following each soaking period were designated as C-6, C-12, C-24, and C-48, respectively. To compare their damage characteristics, a set of dry samples, designated as C-0, was reserved. After soaking the samples, we gently blotted the surface moisture using paper towels and then weighed them. Subsequently, the samples were sealed with plastic wrap and allowed to remain in the room for 3 days to ensure complete water absorption. After sealing the test piece with cling film, since there is no air

convection and the temperature change is relatively small, there is basically no loss of moisture. It is worth noting that the water immersion and sealing maintenance of the concrete specimens were all carried out in the laboratory. This approach basically compensates for the impact of environmental changes on the water loss of the concrete specimens. Following this period, a uniaxial compression test was performed. Assuming that the weight of the specimen after drying is G_0 , and the weight of the specimen after soaking is G , then the moisture content M can be calculated as

$$M = \frac{G - G_0}{G_0}. \quad (1)$$

Table 2 presents the pretreatment outcomes for each sample group.

2.2. Testing System. The testing apparatus primarily consists of a mechanical loading system and an AE monitoring system, as illustrated in Figure 1. The mechanical loading system utilized is the Shanghai Sys-Vertical WAW-3000 series electrohydraulic servo universal testing machine, with a capacity of 3000 kN. The testing procedure employs continuous loading at a rate of 0.1 MPa/s [22, 23]. The AE monitoring system utilized in the experiment is the Sensor Highway 3, a third-generation, 16-channel fully digital system, manufactured by Physical Acoustics Corporation, USA. This system comprises preamplifiers, filters, signal acquisition, and processing units, as well as sensors. We set the preamplifier gain as 40 dB, in order to reduce the impact of environmental noise. We tested the environmental noise in the test site, and the test found that the amplitude of environmental noise is below 35 dB. So, the AE instrument threshold was set at 40 dB, which can filter almost all environmental noise.

2.3. Test Program

- (1) Prior to the testing procedure, the concrete specimens were labeled and weighed with the locations and sequence of the eight sensors clearly marked. As illustrated in Figure 2, this figure demonstrates the positioning of the AE sensors on concrete specimens.
- (2) The test sample is positioned on the base of the testing apparatus, after which the sensor is affixed to the sample surface using a hot melt adhesive. Following the installation of the sensor, the AST function should be utilized to verify the adequacy of the adhesive, ensuring optimal contact between the sensor and the sample surface.
- (3) Set the AE system threshold and preamplifier gain to 40 dB. Typically, the hit definition time is twice the peak definition time, while the hit lock time should be marginally longer than the hit definition time. Consequently, the peak definition time, hit definition time, and hit lock time are configured at 250, 500, and 600 s, respectively.

TABLE 1: Proportions of concrete mix.

W/C	Cement (kg/m ³)	Water (kg/m ³)	Sand (kg/m ³)	Macadam (kg/m ³)
0.58	458.72	266.06	534.05	991.80

TABLE 2: Sample pretreatment result.

Groups	Soaking time (h)	The average weight of the sample after drying (kg)	The average weight after soaking (kg)	Water content (kg)	Proportion of moisture content (%)
C-0	0	7.15	—	—	—
C-6	6	7.12	7.42	0.3	4.21
C-12	12	7.1	7.49	0.39	5.49
C-24	24	7.01	7.44	0.43	6.13
C-48	48	7.11	7.56	0.45	6.33

- (4) Set the loading rate of the testing system to 0.1 MPa/s, activate the AE acquisition device at the onset of loading, and record the AE activities of the sample throughout the loading process.

3. Results and Analysis

3.1. Damage AE Characteristics of Water-Bearing Concrete. The damage AE characteristics of water-bearing concrete were investigated on AE time-domain parameters, frequency-domain parameters, and the number of different types of cracks. The commonly utilized AE parameters to quantify the damage extent of concrete include AE energy [35], AE ringing counts [22], AE hit counts [23], and peak frequency [36]. By integrating time-domain and frequency-domain signal characteristic analyses, the failure process of water-bearing concrete can be accurately quantified.

3.1.1. AE Energy and Ringing Counts Analysis. The AE ringing counts refer to the number of times the AE signal waveform exceeds the preset threshold voltage, which can characterize the activity level of the AE signal. The higher the AE ringing counts, the higher the frequency of microscopic damage events occurring within the materials. The AE energy refers to the area under the envelope line of the AE time-domain signal, which can reflect the intensity of the AE events. The higher the energy, the greater the elastic wave energy released by the internal damage within the materials. Figure 3 illustrates the variations in AE ringing counts, cumulative AE ringing counts, AE energy, and cumulative AE energy for specimens with varying moisture content under stress. In general, apart from the C-0 specimen, the remaining specimens exhibited similar AE characteristics. Dry concrete exhibits a significantly higher ringing count and energy level. However, for wet concrete specimens, both their ringing counts and energy level rapidly decrease to a normal state. On one hand, the compressive force exacerbates the frictional interactions between the internal aggregates of dry concrete, leading to more intense internal activities. Owing to the filling effect

of water in the voids, the water adheres to the surface of the aggregates, forming a water film that reduces the friction and collisions between the aggregates. Furthermore, the addition of water enhances the continuity of concrete, leading to increased medium absorption attenuation of AE waves during propagation. Therefore, it is observed that the AE energy and ringing counts of water-containing concrete are much smaller than those of dry concrete. On the other hand, a significant number of voids exist between the aggregates within dry concrete. Despite these gaps leading to more pronounced scattering, refraction, and diffraction attenuation of the AE wave before it reaches the sensor (owing to the significantly lower acoustic impedance of the gas phase compared to the solid and liquid phases, which hinders the propagation of the AE wave in the gas phase), the energy released during the destruction process does not decrease; rather, it increases. This indicates that dielectric absorption-induced attenuation of AE waves is more substantial than that caused by scattering, refraction, and diffraction. The presence of water not only reduces the collision friction between aggregates but also attenuates the energy of AE waves, thereby diminishing the intensity of AE activity.

Based on the AE energy and ringing count characteristics associated with concrete compression failure, the failure process of concrete can be systematically categorized into three distinct phases: the localized compaction and elastic deformation stage, the stable crack propagation stage, and the unstable crack development stage.

Stage I: the localized compaction and elastic deformation stage (about 0%–10% of peak stress). The concrete specimen begins to load, the original cracks and voids inside are compacted, and elastic deformation occurs. The interconnections between the aggregates are robust, and no new cracks are initiated. The AE signals at this stage primarily originate from the collision friction between aggregates as the internal voids within the concrete close. The duration for dry concrete at this stage is longer compared to water-containing concrete, with greater energy release and more intense aggregate collision and friction. Conversely, water-containing concrete compacts more easily due to the

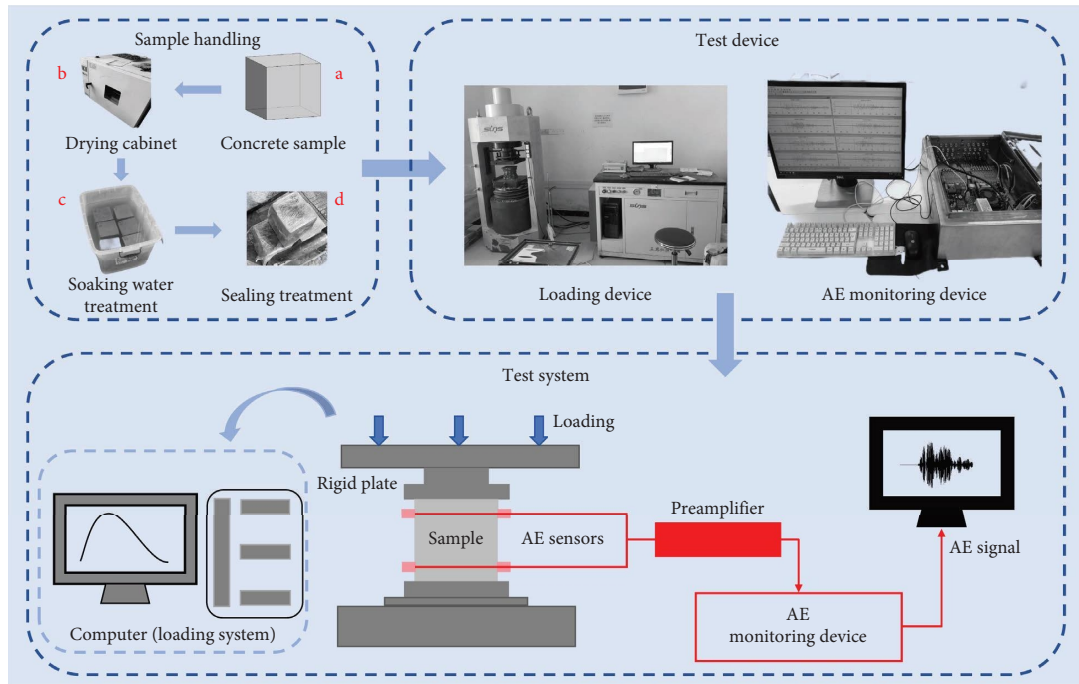


FIGURE 1: Test system diagram.

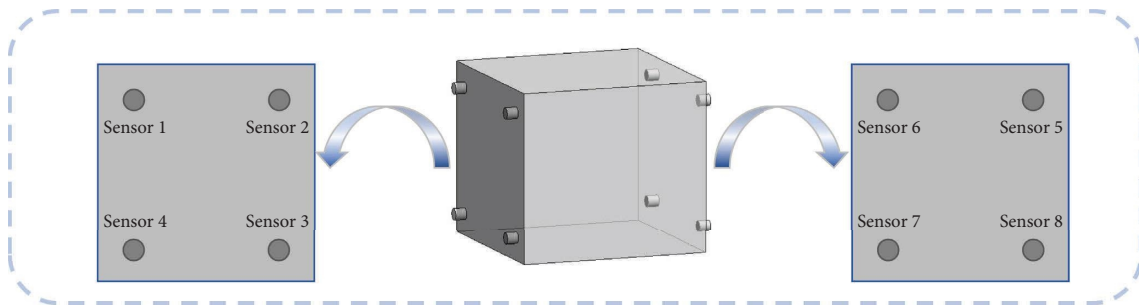


FIGURE 2: Sensor layout.

filling effect of water, leading to a shorter duration, reduced aggregate collision friction, and partial absorption of collision-generated energy by water, thereby weakening AE activity.

Stage II: the steady crack propagation stage (about 10%–70% of the peak stress). With the increase of stress, the initial crack re-initiates propagation, and the mortar and aggregate start to exhibit slip toward the crack plane. At this time, the initial crack also started from the interface between cement mortar and aggregate with the weakest connection strength, and the crack continued to extend into the cement mortar matrix. The cement mortar matrix experiences irreversible plastic strain. At this stage, the AE activity is weak, and the AE energy and AE ringing counts develop steadily. With the increase of moisture content, the duration of this stage gradually lengthens. Dry concrete exhibits a higher number of initial cracks, faster development rates, and greater energy release, leading to more intense internal activities.

Stage III: the rapid crack development stage (about 70%–100% of the peak stress). As the stress on the specimen

progressively approaches the peak stress, the rate of microcrack propagation accelerates. This leads to increased deformation of the specimen. The microcracks begin to interact with and eventually coalesce with the main cracks, resulting in instability and ultimate failure of the specimen. At this stage, there is a sharp increase in the number of AE signals, and macroscopic cracks become evident on the specimen's surface, accompanied by audible destruction sounds. Notably, the increase in AE signals between Stage II and Stage III for specimens in the dry state is relatively modest, suggesting potential interactions among various failure modes. This interaction will be discussed in detail in Section 3.1.3.

By comparison, we observe that the AE activity in dry concrete is relatively intense across the three stages. However, after water absorption, the AE activity is significantly attenuated as a result of the lubricating effect of the water film and the absorption of AE energy by the medium. With the increase of moisture content, Stage II gradually lengthened, Stage III gradually shortened, and tended to be stable.

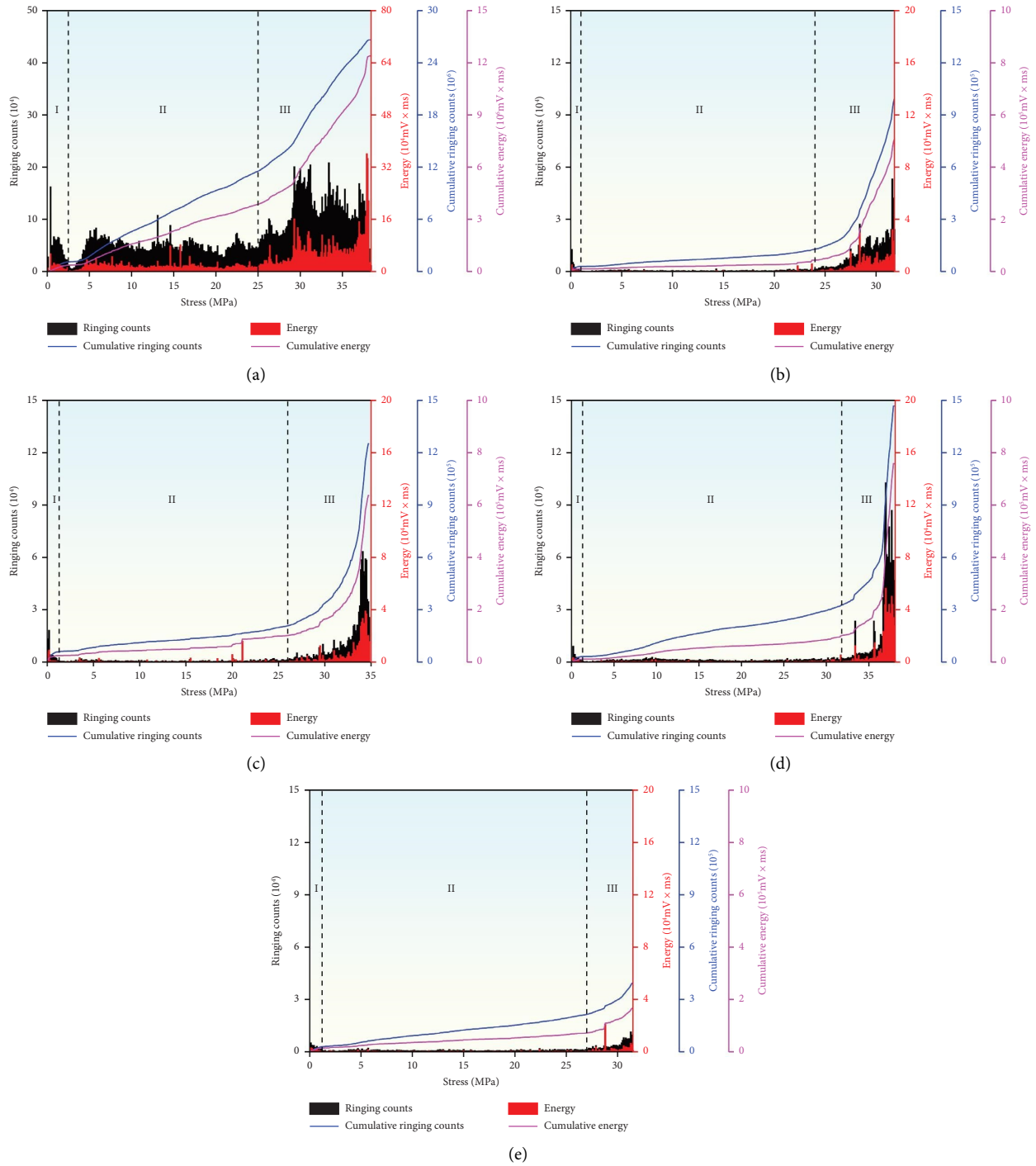


FIGURE 3: Correlation analysis of ringing counts, energy, and pressure: (a) C-0, (b) C-6, (c) C-12, (d) C-24, and (e) C-48.

The aforementioned rules not only demonstrate a significant correlation between AE parameters and the failure process of concrete but also indicate that the presence of free water in concrete has an influence on the failure mode of the specimens. To comprehensively understand the failure modes of concrete specimens with various moisture contents, it is insufficient to solely analyze time-domain parameters. Therefore, we conducted an analysis of the peak frequency extracted from the AE signals after performing fast Fourier

transform (FFT) [37]. The peak frequency can effectively reflect the number of signals in each frequency band and the changes in high-, medium-, and low-frequency signals, which is essential for the characterization of concrete damage.

3.1.2. AE Peak Frequency and Hit Counts Analysis. The characteristics of AE signals cannot be fully obtained only through the time-domain parameter analysis of AE signals.

Therefore, this section analyzes the peak frequency of AE signals by using FFT. The AE peak frequency refers to the frequency component in the frequency domain where the energy of the AE signal is most concentrated, that is, the frequency point with the highest amplitude in the spectrogram. When the AE signal exceeds the threshold and is received by any sensor, a hit occurs. The higher the hit counts, the more frequent the damage activity. For the purpose of facilitating analysis, we categorize the peak frequency into three bands: low, medium, and high. The low-frequency band encompasses signals ranging from 0 to 100 kHz, the medium-frequency band includes signals from 100 to 200 kHz, and the high-frequency band comprises signals exceeding 200 kHz. Figure 4 shows the change process of peak frequency of five groups of samples with different moisture content over time. Overall, the AE cumulative hit counts show a turning point at the junctions of each stage. In different stages of the specimen under dry conditions, the number of signal distributions in the three frequency bands is relatively large. Relatively speaking, partial absorption of high-frequency signals leads to a sharp reduction in their quantity within the specimen after water absorption. Figure 5 shows the percentage of low-, medium-, and high-frequency signals in the total signals of the three stages in each of the five groups of specimens.

In Stage I, the cumulative AE hit counts have rapidly increased, indicating that the damage caused by friction and collision during this stage was significant, and the AE activity has intensified. As the moisture content increased, the proportion of low-frequency signals progressively rose. Upon reaching saturation, the low-frequency signals had increased by 61.03%. The proportion of intermediate frequency signals exhibited a gradual downward trend, decreasing from C-0 to C-48 with a reduction rate of 60.45%. No substantial variation in the frequency of high-frequency signals was detected. In other words, at this stage, the signal is mainly distributed in the middle- and low-frequency bands. With the increase of moisture content, free water not only weakens the collision friction between aggregates but also reduces the energy generated by the collision friction and absorbs part of the signal in the middle band [10], so that the dominant frequency band of the signal is transferred to the low band.

In Stage II, overall, at the junction with Stage I, the increase in cumulative hit counts suddenly slowed down. Except for the dry concrete specimens, the cumulative hit counts of AE in other specimens rose slowly throughout Stage II. This indicates that after entering Stage II, the damage activities within the concrete significantly weakened, and the cumulative hit counts of AE can clearly distinguish between Stage I and Stage II. The frequency band composition of the signals in Stage II is still dominated by the medium- and low-frequency bands. In contrast to Stage I, the frequency band signal of water-containing concrete remains stable, with no substantial variations observed. As the concrete transitions from a dry state to one that contains water, the predominant frequency band shifts from the mid-frequency range to the low-frequency range. This suggests that during the steady crack development stage, the free

water occupying the voids within the specimen progressively accumulates as microcracks expand. The randomness in the number of these microcracks results in varying degrees of absorption of the medium-frequency signal by the free water. Due to the absence of free water in dry concrete, the energy released during crack propagation is significant, leading to the direct reception of mid-frequency signals by the sensor. Therefore, the response of dry concrete specimens primarily relies on mid-frequency signals.

In Stage III, the cumulative AE hit counts rapidly increased, and the internal damage of the concrete intensified significantly. The activity of AE was significantly enhanced. The AE signals were distributed in the low-, medium-, and high-frequency bands, but the low-frequency and medium-frequency bands were predominant. The difference was that compared with Stage I and Stage II, the proportion of high-frequency band signals was larger in Stage III. The proportion of low-frequency and medium-frequency band signals in dry concrete was comparable, while the high-frequency band signals were about 4–15 times that of water-containing concrete. In this stage, as the water content increased, the AE activity decreased slightly. The high-frequency signals and some of the medium-frequency signals were absorbed by water, and the signals shifted from the high- and medium-frequency bands to the medium- and low-frequency bands, with a relatively small shift amplitude.

In summary, the cumulative hit counts of AE can effectively represent the degree of damage activity within the concrete. Analysis of the peak frequency reveals that the water's effect weakens the fracture mechanism of the concrete and absorbs the AE signals in the medium- and high-frequency bands. However, when the specimen reaches the saturated water absorption rate, this change may tend to stabilize. Based on the variations in signal proportions observed during Stages I, II, and III, it is evident that the free water within the voids of the specimen exhibits a higher absorption capacity for mid-to-high frequency signals, whereas low-frequency signals remain relatively stable.

3.1.3. AE RA-AF Analysis. The cracks in concrete under uniaxial compression are mainly shear cracks and tension cracks. Relevant studies have shown that the values of RA and AF are closely related to the cracking type of concrete [38–42]. The RA value refers to the ratio of the AE rise time to the amplitude, which is called the rise angle. AF is the ratio of the AE ringing counts to its duration. The shear damage signal exhibits a higher RA value and a lower AF value, whereas the tension damage signal demonstrates a lower RA value and a higher AF value. This suggests that these two types of signals can be effectively differentiated by establishing an appropriate demarcation line. To facilitate the comparison of cracking signal characteristics across varying moisture content states, this study includes all signals from the five specimen groups. Therefore, using dry-state concrete specimens as a reference, the slope of the dividing line is set to 1. Figure 6 shows the classified signals and the proportion of different types of signals at each stage.

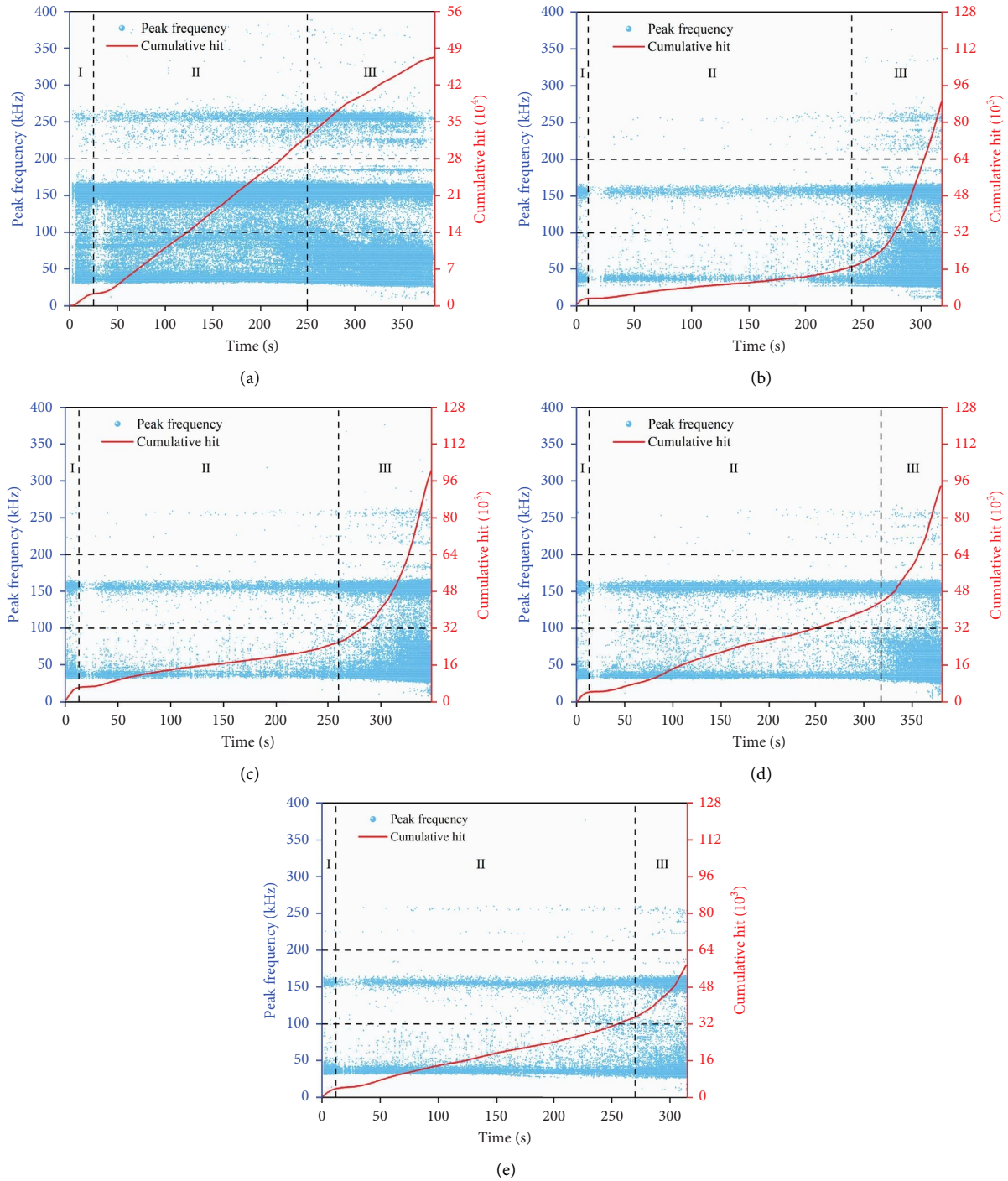


FIGURE 4: Correlation analysis of AE peak frequency and hit counts: (a) C-0, (b) C-6, (c) C-12, (d) C-24, and (e) C-48.

On the whole, the predominant cracking signal observed in the specimen is tensile in nature, with the shear signal comprising only 1.44%–14.09% of the total signals. As the moisture content increases, the proportion of shear signal exhibits a gradual decrease. It is noteworthy that the shear crack signal intensity in dry concrete specimens is significantly higher than that in water-containing concrete, suggesting that the presence of water alters the damage mechanism of concrete. When dry concrete is subjected to

compressive stress, the lack of infill in the voids and the irregular shape and size of internal aggregate particles result in interparticle slippage. This slippage can cause a deflection in the direction of the principal compressive stress acting on the aggregate particles, thereby increasing the angle between the aggregate particles and the principal compressive stress. Consequently, this phenomenon enhances the likelihood of diagonal shear failure in concrete. After concrete absorbs water, the free water fills the voids and bears a minor portion

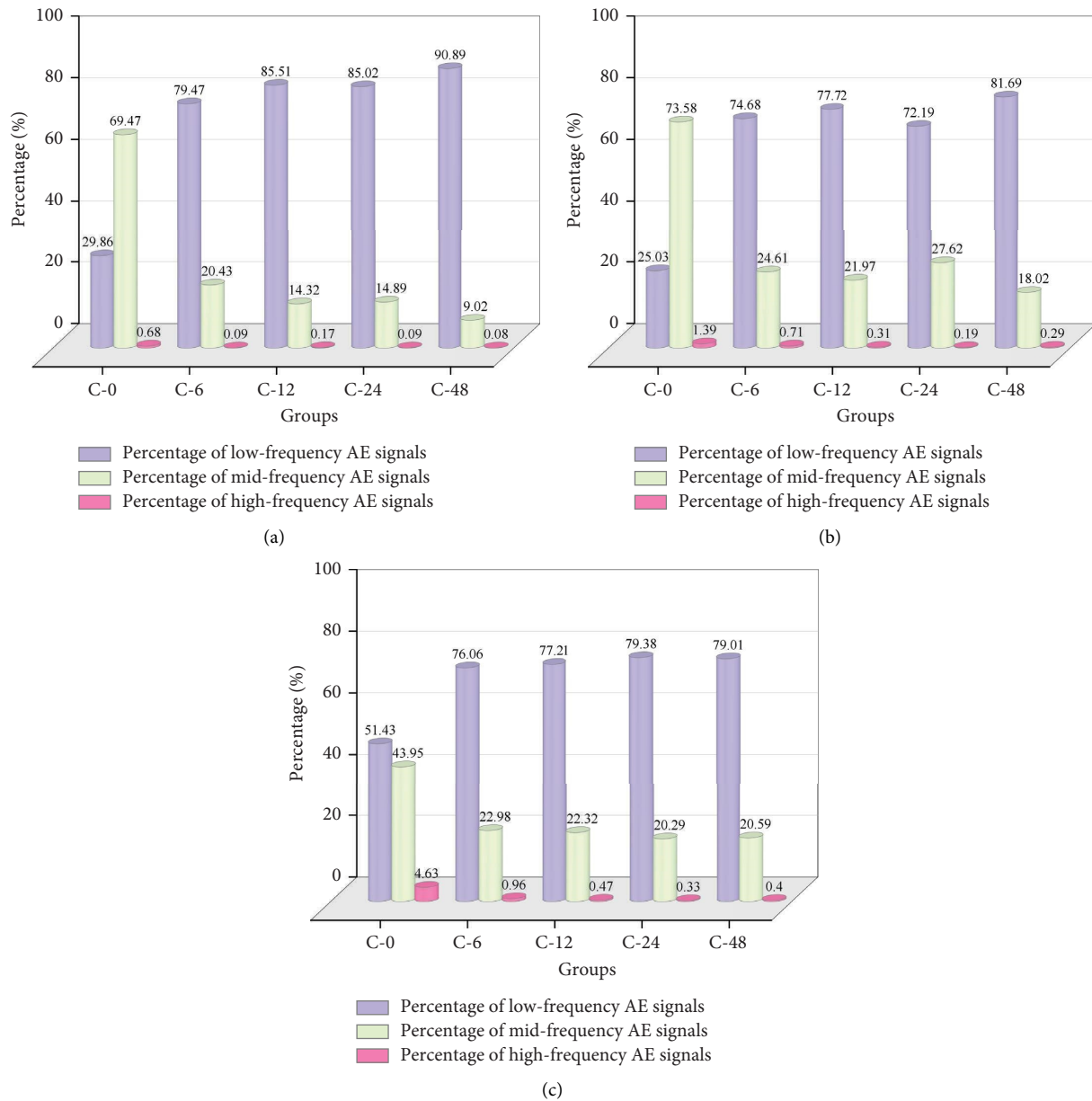


FIGURE 5: Proportions of signals in different frequency bands at different stages: (a) Stage I, (b) Stage II, and (c) Stage III.

of the pressure. Consequently, the angle between the aggregate particles and the principal compressive stress undergoes a slight alteration. When the internal stress within the concrete surpasses its bearing capacity, microscopic damage initiates, ultimately leading to the formation of macroscopic tensile cracks.

From a stage-by-stage perspective, during Stage I, the proportion of shear signals progressively diminishes. This trend suggests that as moisture content increases, the internal voids within the specimen become increasingly filled, thereby mitigating the likelihood of inclined shear failure to some extent. In Stage II, except for the dry concrete specimens, the other specimens hardly produce the shear-type signal. This is because the development of microcracks at this stage is dominated by the main compressive stress, its direction is almost not deflected,

and due to the pressure and lubrication of free water, the failure trend of pressure cracks in Stage I is not overextended. In Stage III, the microscopic crack penetration and macroscopic failure occurred, and the shear signal of 2.76%~22.97% was generated. At this juncture, the shear failure trend established during Stage I commenced its influence. The orientation of the principal compressive stress underwent a shift, leading to an increased angle between the aggregate particles and the principal compressive stress. Consequently, partial shear failure was observed. With the increase in moisture content, the proportion of shear signals exhibits an irregular pattern. This is attributed to the interaction between shear cracks and tension cracks during their development, leading to the formation of composite cracks. Such interactions result in instability and unpredictability in crack behavior.

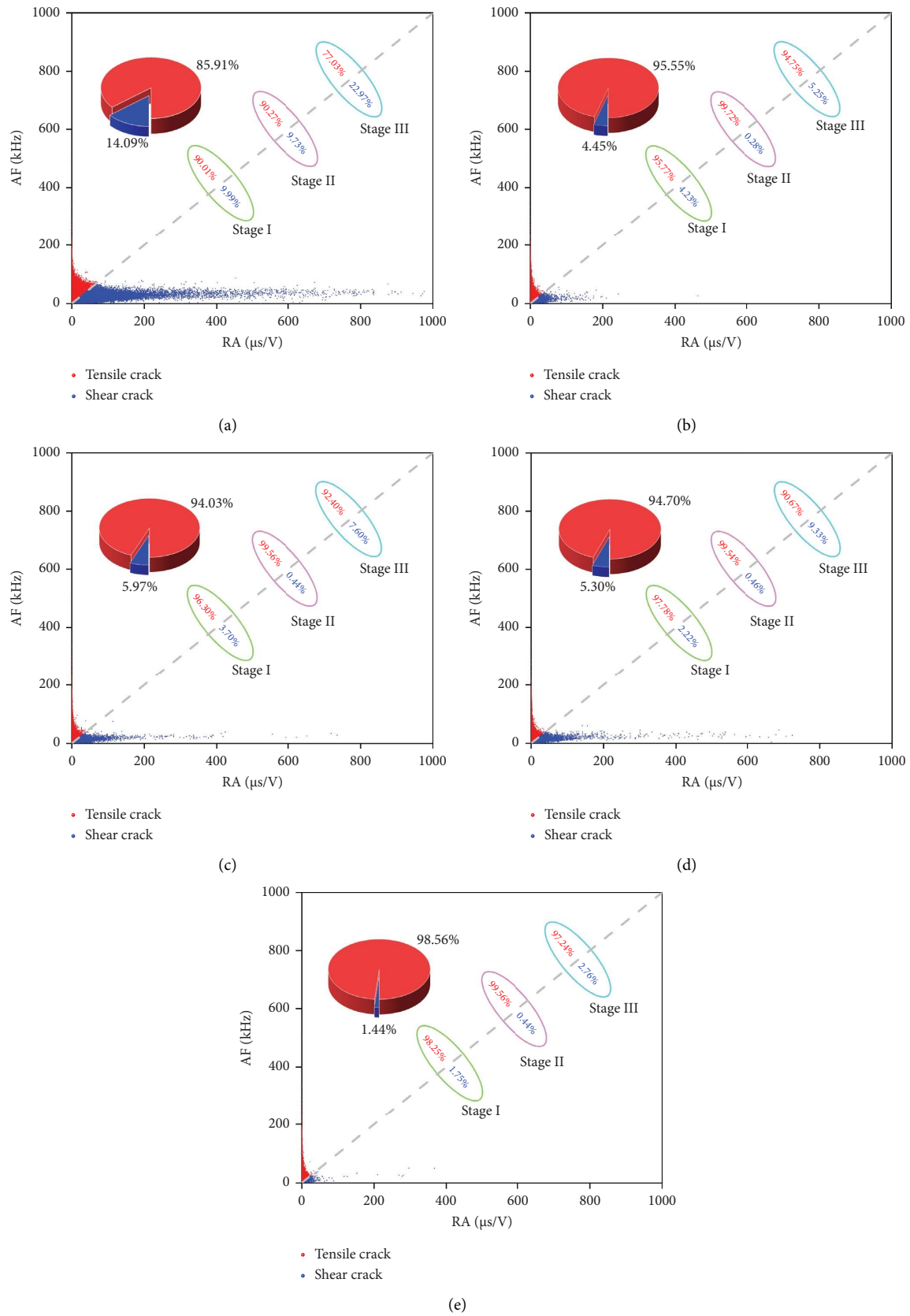


FIGURE 6: The classification of shear and tension cracks and the proportion of cracks in each stage: (a) C-0, (b) C-6, (c) C-12, (d) C-24, and (e) C-48.

3.2. Recognition of Cement Slurry Crack Pattern Based on Unsupervised ML. The aforementioned analysis examines the characteristics of AE signal time-domain and frequency-domain parameters, as well as the development mechanisms of internal cracks in concrete with varying moisture content, based on changes in time-domain parameters, frequency-domain parameters, and RA-AF. Accurately identifying various types of damage signals remains the primary challenge that needs to be addressed, and it is of critical importance for evaluating the failure modes of concrete. It is challenging to distinguish between the various types of damage signals in concrete. The emergence of ML offers a novel approach for the pattern recognition of damaged signals. In this section, we incorporate moisture content analysis, employ ML algorithms, and integrate the method proposed in literature [34] to identify and predict the signal pattern of the concrete cracking process. Figure 7 illustrates the comprehensive methodology employed in this section. The primary contents encompass the correlation analysis of AE parameters, feature engineering of high-dimensional data for ML, concrete cracking signal pattern classification based on an unsupervised learning algorithm, cracking signal pattern identification based on moisture content and AE parameters, and cracking signal pattern prediction based on BP neural network.

3.2.1. Correlation Analysis of AE Parameters. The parameters utilized in ML should exhibit low similarity or substantial dissimilarity to enable the model to more effectively identify patterns within the data. So, it is necessary to analyze the correlation of AE parameters. The main parameters used for AE signal analysis include time-domain parameters and frequency-domain parameters; the time-domain parameters mainly include rise time [43], counts, energy, duration [44], amplitude [45], signal strength [46], and absolute energy [47], and the frequency-domain parameters mainly include frequency centroid [48] and peak frequency.

Correlation analysis is a statistical method employed to quantify the degree of association between two or more variables. Essentially, correlation analysis is used for spotting patterns within datasets. A positive correlation result means that both variables increase in relation to each other, while a negative correlation means that as one variable decreases, the other increases. In general, the relationship between data can be categorized into two types: linear correlation and nonlinear correlation. The relationship between linear data can frequently be described through the formulation of a linear equation, wherein an increase or decrease in one variable corresponds to a proportional increase or decrease in the other variable. The relationship between nonlinear data is inherently more intricate and cannot be adequately represented by a linear model. As one variable undergoes changes, the corresponding changes in the other variable occur at varying rates. Pearson [49] correlation analysis is utilized to quantify the linear relationship between variables, assuming that the data are normally distributed. Pearson correlation analysis quantifies the linear relationship between variables by defining the Pearson correlation

coefficient “ r .” In cases where the data contain outliers or do not follow a normal distribution, the interpretative power of the Pearson correlation coefficient may be compromised. The Pearson correlation coefficient “ r ” is defined as follows:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2)$$

where x_i and y_i represent 2 groups of variables, and $\bar{x}$ and $\bar{y}$ represent the arithmetic mean of the two sets of variables, respectively.

Spearman correlation analysis [49] is employed to assess the nonlinear relationship between variables. This method evaluates the monotonic association based on the ranks of the data points, rendering it relatively robust to the distribution shape and less sensitive to deviations from normality. Spearman correlation analysis measures the degree of correlation between data by defining the Spearman correlation coefficient “ S ,” which is not affected by data distribution and is not sensitive to outliers. Spearman’s correlation coefficient “ S ” is defined as follows:

$$S(X, Y) = 1 - \frac{6 \sum_{i=1}^n D_i^2}{N(N^2 - 1)}, \quad (3)$$

where X and Y represent 2 sets of data, D_i represents the difference between the grades of two variables being observed, and N indicates the sample size. The range of Pearson and Spearman correlation coefficients is $[-1, 1]$ (correlation approaching 0 indicates no correlation; approaching 1 represents a positive correlation; approaching -1 means negative correlation).

Figure 8 presents the Pearson and Spearman correlation coefficient matrices for 9 AE parameters, specifically rise time, ringing counts, energy, duration, amplitude, signal strength, absolute energy, frequency centroid, and peak frequency. In general, the parameters showed a strong nonlinear correlation, so the parameters used in the subsequent process were selected according to Spearman correlation coefficient analysis. Spearman correlation analysis revealed a strong association between energy, absolute energy, and signal intensity, with correlation coefficients ranging from 0.91 to 1.0, and concentrated distribution around 0.97. Consequently, the selection of energy parameters can effectively characterize both absolute energy and signal intensity. The absolute values of the correlation coefficients between rise time, ringing counts, duration, amplitude, frequency centroid, and peak frequency are all below 0.95, indicating that these parameters can be utilized for subsequent data processing. However, considering that both the rise time and the duration are defined based on the time axis, both reflect the time evolution process of the signal and are related to the energy release characteristics of the signal. That is, the shorter the rise time, the more concentrated the signal energy release is, and the longer the duration, the more persistent the signal energy release process is. Moreover, the rise time can only represent part of the characteristics of the AE signal, while the duration can directly reflect the length of the AE signal in the time

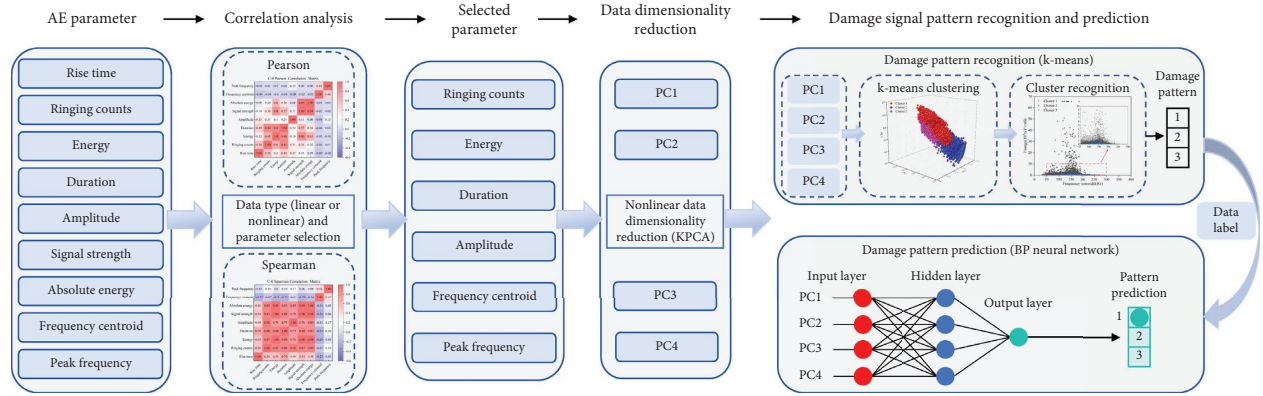


FIGURE 7: Methodological introduction.

domain. Therefore, the rise time is excluded. The final AE parameters determined for subsequent processes include ringing counts, energy, duration, amplitude, frequency centroid, and peak frequency.

3.2.2. Feature Engineering of High-Dimensional Data for ML. There are 6 parameters selected after Spearman correlation analysis, which is still difficult for ML models to process. Feature engineering of high-dimensional data is the focus of ML. Commonly employed data dimensionality reduction and feature extraction techniques encompass PCA [50] and kernel PCA (KPCA) [51]. The traditional PCA algorithm is primarily designed for linearly separable data, and its effectiveness is notably constrained when applied to non-linearly separable data. Specifically, a key limitation of the PCA algorithm is its inability to effectively separate non-linearly separable data using a linear classifier in the original feature space. The KPCA algorithm employs a kernel function to project the data into a high-dimensional feature space, thereby enhancing the linear separability of the data within this space. Subsequently, it applies the PCA algorithm to the transformed data for dimensionality reduction and feature extraction. Common kernel functions utilized in KPCA encompass the linear kernel, polynomial kernel, and radial basis function (RBF) kernel, among others. Notably, both the polynomial and RBF kernels fall into the category of nonlinear kernels. The RBF kernel is one of the most common and efficient kernel functions, and its expression is as follows:

$$V(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \quad (4)$$

where x_i and x_j represent the input sample, and σ is the width parameter of the function, which controls the radial range of the function.

As evidenced in the preceding section, AE parameters exhibit characteristics of nonlinear data. Consequently, KPCA is applied to these data, utilizing a RBF as the kernel. In the computation of KPCA, the proportion of variance explained by each principal component relative to the total variance can be derived. This proportion serves as an

indicator of the information encapsulated within each principal component. The formula for calculating covariance is as follows:

$$\text{Cov} = \frac{1}{n} \sum_{i=1}^n \phi(x_i) \phi(x_i)^T, \quad (5)$$

where n is the dimension of the original data, x_i is the i -dimensional raw data, and ϕ is the mapping function. The characteristic equation is

$$\lambda \cdot v = \text{Cov} \cdot v, \quad (6)$$

where λ is the eigenvalue, and v is the corresponding eigenvector. The variance contribution rate is calculated as follows:

$$\lambda \cdot v = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i}. \quad (7)$$

Figure 9 illustrates the variance interpretation rates of each principal component for AE parameters. The cumulative interpretation rates for PC1, PC2, PC3, and PC4 reached 95%, effectively capturing the majority of parameter characteristics. Furthermore, the data transformed by KPCA are more amenable to visualization and interpretation. The outcomes of KPCA will subsequently be employed in the ensuing ML procedures, which will be elaborated upon in the following section.

3.2.3. Concrete Cracking Signal Pattern Classification Based on the Unsupervised Learning Algorithm. To eliminate the dimensional discrepancies among various principal components and facilitate subsequent analysis, PC1, PC2, PC3, and PC4 were normalized. Unsupervised learning refers to a type of ML algorithm where the data are not pre-labeled, and the algorithm autonomously identifies features and patterns within the data. Commonly employed unsupervised learning algorithms include k-means clustering [52], spectral clustering [53], and density-based clustering [54]. The k-means clustering algorithm is currently the most widely utilized unsupervised learning method. It boasts high computational efficiency, making it particularly suitable for large-scale datasets. In addition, it can achieve convergence

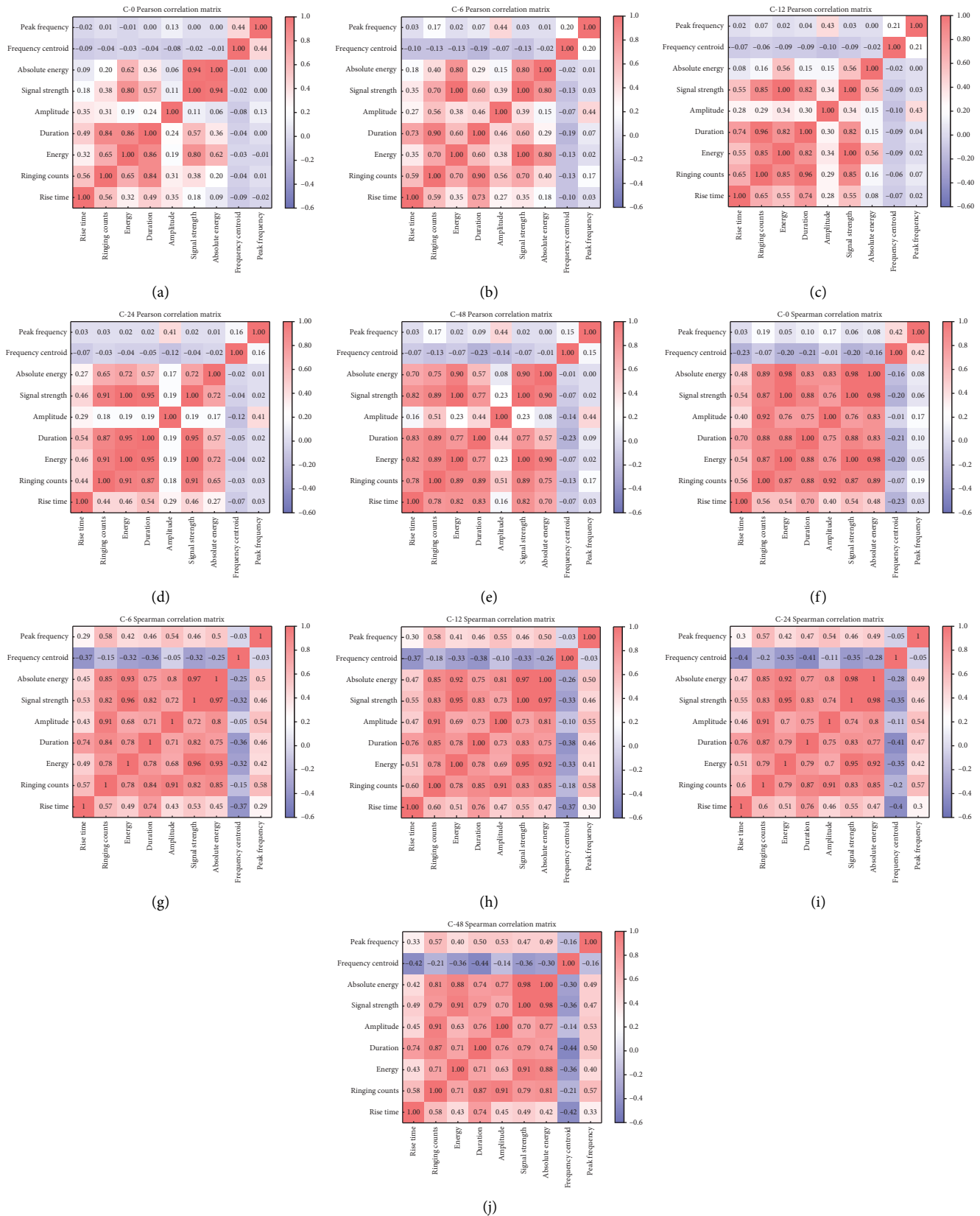


FIGURE 8: Pearson and Spearman correlation analysis: (a) C-0 Pearson, (b) C-6 Pearson, (c) C-12 Pearson, (d) C-24 Pearson, (e) C-48 Pearson, (f) C-0 Spearman, (g) C-6 Spearman, (h) C-12 Spearman, (i) C-24 Spearman, and (j) C-48 Spearman.

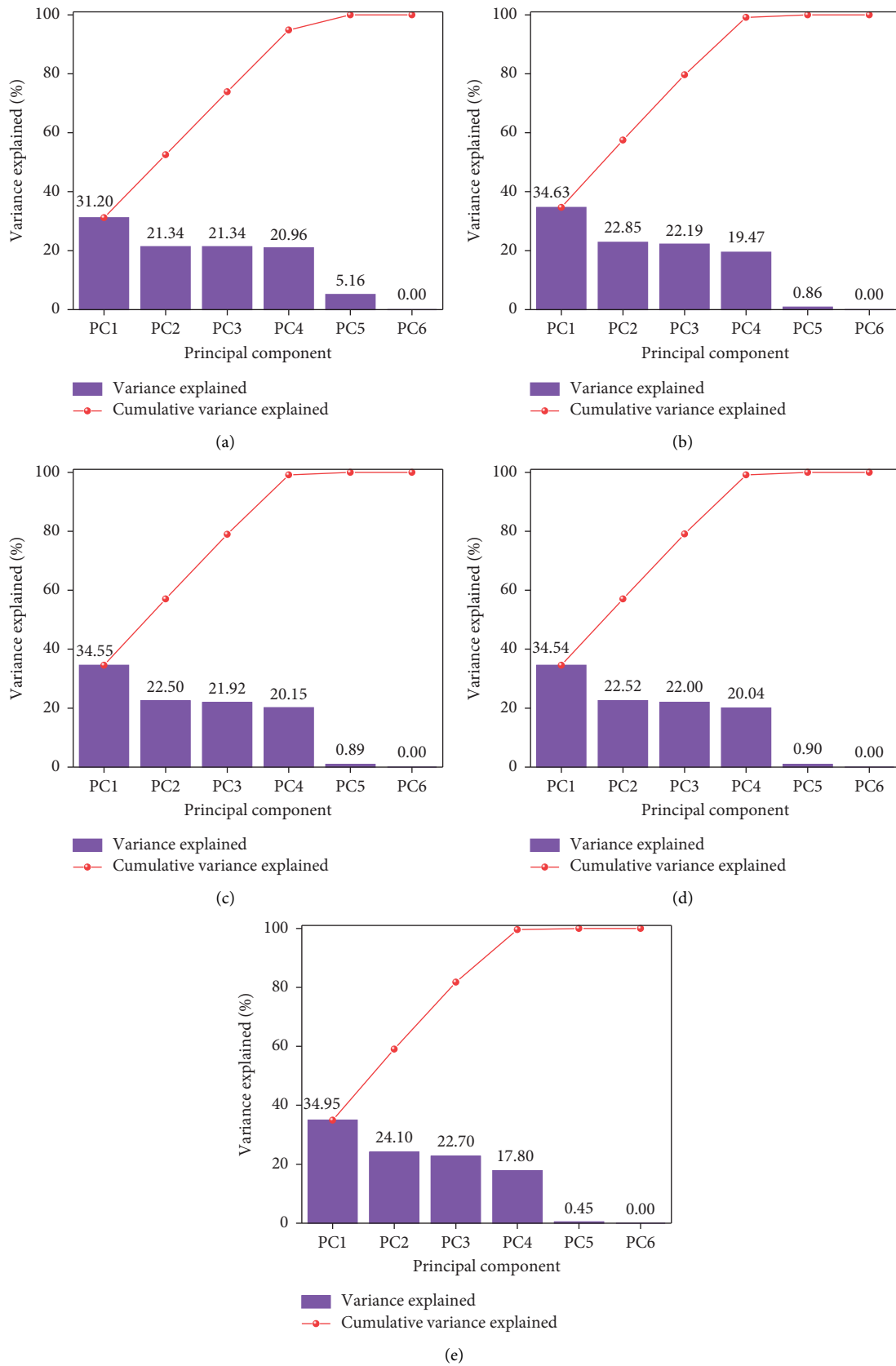


FIGURE 9: Variance explained rate of each principal component: (a) C-0, (b) C-6, (c) C-12, (d) C-24, and (e) C-48.

within a finite number of iterations. It has a simple structure and is highly interpretable. The fundamental principle of the k-means algorithm is to minimize the within-cluster sum of squares, thereby ensuring that data points are grouped into clusters, where the intracluster similarity is maximized. Through multiple iterations, the algorithm will refine the distribution of clusters to minimize the distance between data points and their respective centroids, ultimately achieving convergence. The sum of squares of the distance between each point and the center of its cluster is calculated as follows:

$$W = \sum_{i=1}^n \sum_{x \in N_i} \|x - \mu_i\|^2, \quad (8)$$

where n represents the number of clusters, N_i is the point set of cluster i , x is the data point in N_i , μ_i represents the cluster center of cluster i , and $\|x - \mu_i\|^2$ is the square of the Euclidean distance between x and μ_i . k-means usually uses Euclidean distance to measure the distance between a point and the center of a cluster, and its formula is

$$d(x, \mu) = \sqrt{\sum_{j=1}^n (x_j - \mu_j)^2}. \quad (9)$$

Prior to employing the k-means algorithm, it is imperative to predetermine the number of clusters, a factor that significantly influences the clustering outcome. To determine the optimal number of clusters, we concurrently calculate the silhouette coefficient, Calinski–Harabasz (CH) index, and Davies–Bouldin (DB) index, with the number of clusters ranging from 2 to 5. It is worth noting that the silhouette coefficient, DB value, and CH value mentioned here are all evaluation indicators calculated based on the k-means clustering results. The silhouette coefficient focuses on the rationality of the assignment of individual samples. The DB index is used to measure the balance between the compactness within clusters and the separation between clusters. The CH index assesses the separation degree of clusters by comparing the ratio of intercluster dispersion to intracluster dispersion. A lower DB value, coupled with a higher contour coefficient and CH value, indicates a more effective clustering outcome. Figure 10 presents the calculated results of the contour coefficient, CH value, and DB value for the C-0, C-12, and C-48 specimens. Two clusters are appropriate for dry concrete, whereas three clusters are recommended for water-containing concrete. This indicates that a certain damage mechanism is expanding as the specimen absorbs water, so we determined the cluster as 3 in order to explore this damage mechanism.

The k-means clustering algorithm was applied to five distinct datasets, effectively partitioning the original data into three distinct clusters. Figure 11 shows the spatial distribution of the three clusters. Although PC1, PC2, PC3, and PC4 are utilized for clustering, for the sake of descriptive simplicity, we have chosen to visualize the data in three-dimensional space using only PC1, PC2, and PC3. As can be seen from the figure, although the clustering shapes of other

samples except C-0 are different, the overall clustering effect is good. The C-0 specimen generated 2–3 times more signals compared to other specimens, as the damage occurred in an environment devoid of water, thereby eliminating the absorption of sound signal energy by water. Furthermore, the substantial energy release has resulted in elevated values of the measured acoustic parameters, which are identified as the primary factor contributing to the suboptimal clustering performance. A viable approach is to evaluate the signal pattern following clustering based on the variation of moisture content and AE parameters. In the next section, we will identify the signal pattern of the whole process of concrete compression through moisture content and AE parameter characteristics.

3.2.4. Cracking Signal Pattern Identification Based on Moisture Content and AE Parameters. The proportion of signals from each cluster relative to the total number of signals is presented in Table 3. We will consider identifying clustering features with as few AE parameters as possible. Energy and amplitude can be used to quantify the strength of the signal, while duration reflects the temporal extent of the signal. In addition, frequency centroid provides insight into the signal's frequency-domain characteristics. Consequently, these four parameters, energy, amplitude, duration, and frequency centroid, are selected as the key indicators for identifying damage signal patterns associated with clustering analysis. Figure 12 illustrates the correlation analysis diagram of AE signal parameters for different clustering segments. Three distinct types of characteristic signals can be identified from the figure, detailed as follows:

- (1) Cluster 1 exhibits the lowest signal energy, with its amplitude and duration being marginally higher than those of Cluster 2, while the frequency centroid remains comparable. In addition, the proportion of Cluster 1 among all specimens did not exhibit significant variation and was consistently the smallest. In other words, the damage mode associated with Cluster 1 is commonly observed in both dry and water-containing concrete specimens, exhibiting minimal correlation with their moisture content. The moisture content will affect the strength of concrete, so the damage mechanism inside the concrete is basically related to the moisture content. Owing to the development and propagation of microcracks within the concrete, the compressive surface may experience minor slippage, leading to friction. The friction is relatively mild, resulting in a low intensity of the measured AE signal. The low-frequency centroid suggests that the energy of this signal is predominantly concentrated in the low-frequency components. Cluster 1 may be related to the frictional motion signals of the concrete compression surface.
- (2) The signal of Cluster 2 exhibits moderate energy levels (higher than those of Cluster 1), a moderate frequency centroid, relatively lower amplitude, and

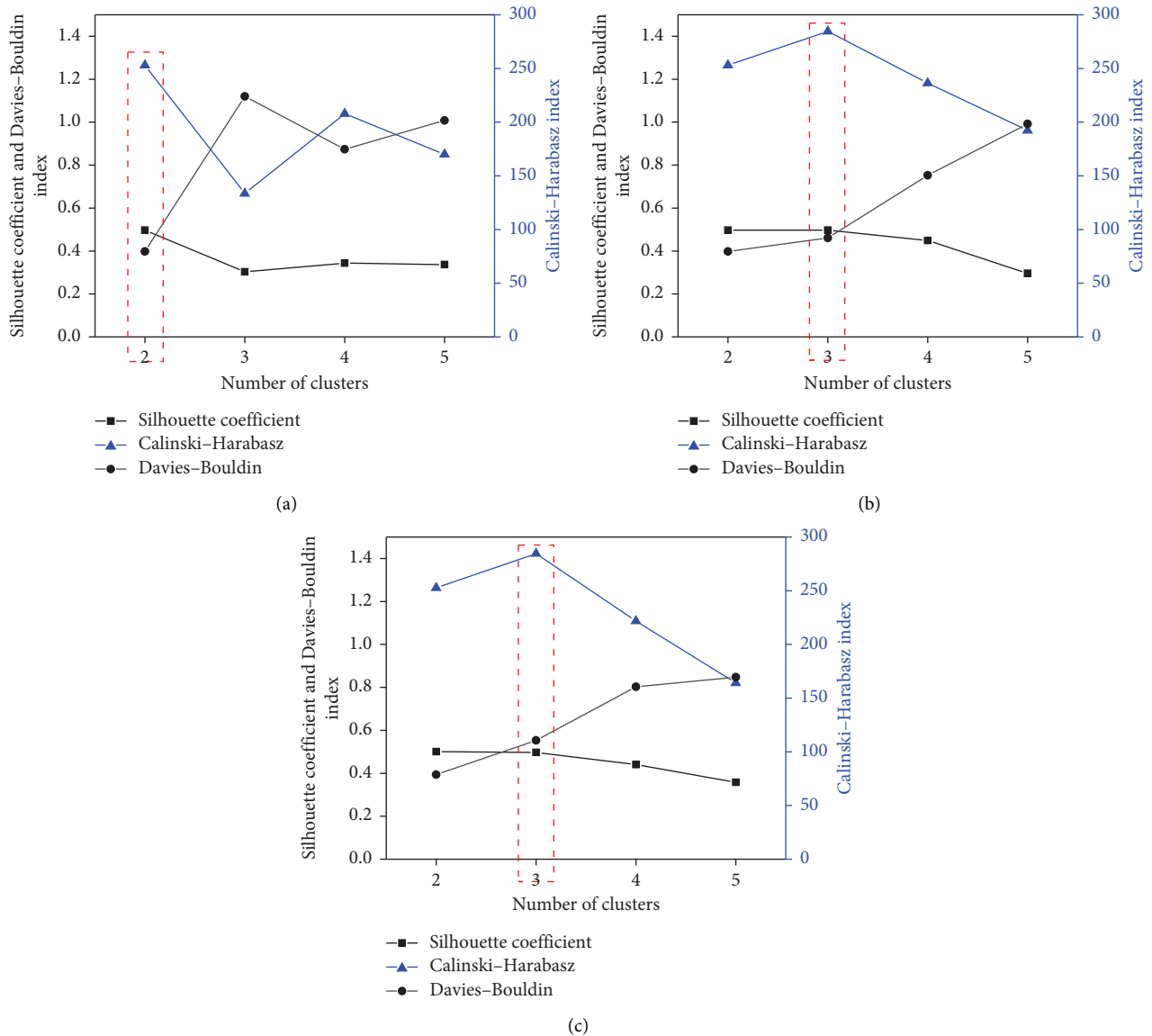


FIGURE 10: Selection of cluster number: (a) C-0, (b) C-12, and (c) C-48.

shorter duration. The medium energy and low amplitude indicate that the signal is not violent. The proportion of Cluster 2 exhibited a gradual increase as the moisture content rose. In the process of concrete fracture, the fracture surface will be active, including frictional slip, collision, intersection, and so on. In dry concrete, the movement of the fracture surface is weak. After the specimen absorbs water, the free water in the void will concentrate to collect and fill the microcracks in the cement mortar. The accumulated free water will produce a “wedging [10]” effect on the tip of the microcrack, accelerate the development of microcrack, and produce more AE signals. Due to the lubrication and wedging action of water, the movement of the fracture surface is intensified. Combined with AE parameter features, it can be determined that these are the fracture surface activity signals.

(3) The signal distribution range of Cluster 3 is wide, and the values of signal characteristic parameters are distributed at all levels. With the increase in moisture content, the proportion of Cluster 3 progressively diminishes. The smaller the moisture content, the more intense this signal pattern is, which may be related to aggregate cracking. Due to the filling and lubricating effect of free water on the void inside the concrete, the friction between the cement mortar and the particles inside the interfacial transition zone (ITZ) is reduced, making the aggregate more prone to cracking. Coarse aggregate is the hardest part of concrete and releases a lot of energy when broken, corresponding to a signal with the highest energy and maximum amplitude. The ITZ is the interface area between aggregate and cement mortar in concrete, which has high porosity, is the weak part in concrete, and is also one of the areas most susceptible to external

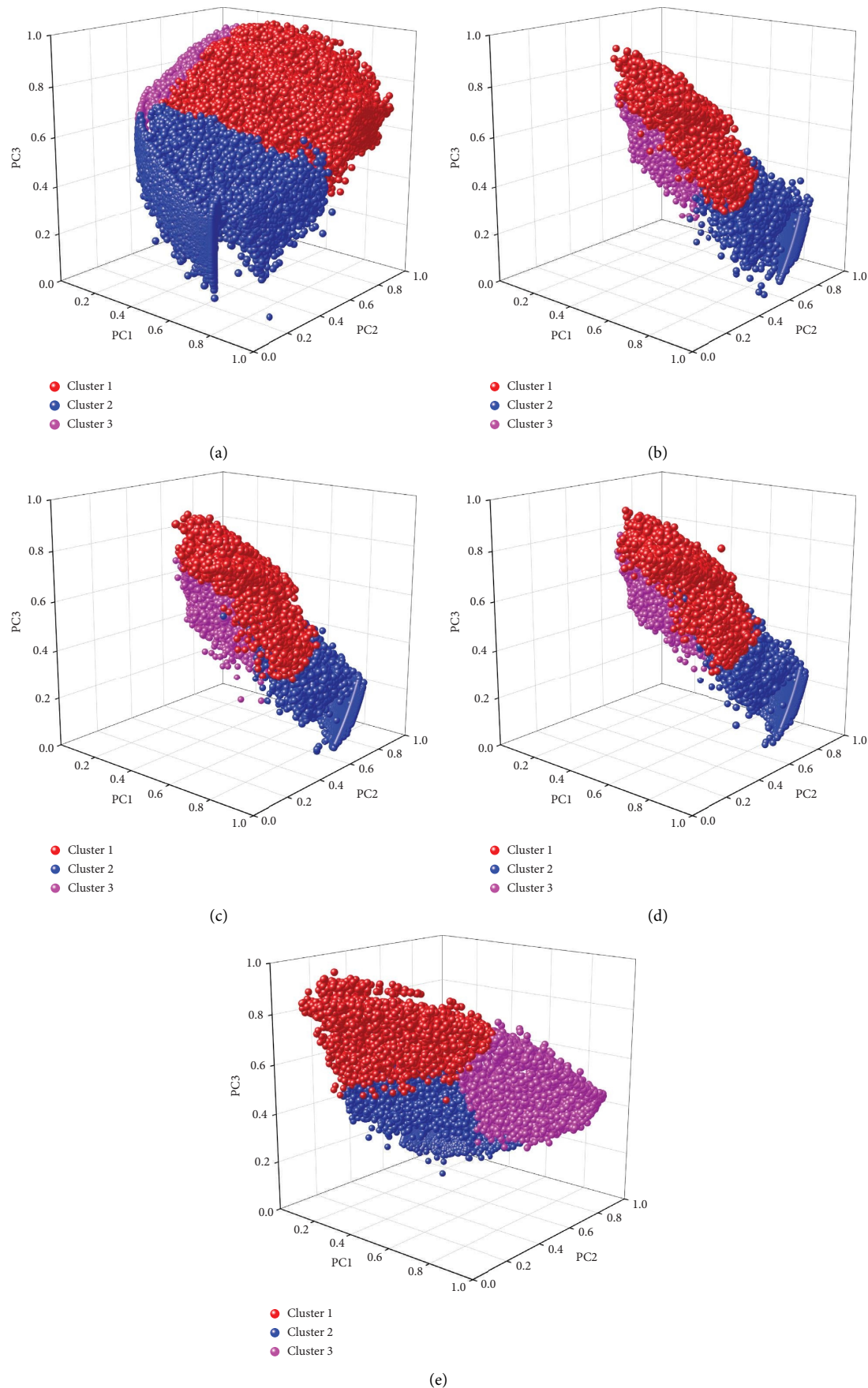


FIGURE 11: Three-dimensional visualization of clustering data: (a) C-0, (b) C-6, (c) C-12, (d) C-24, and (e) C-48.

TABLE 3: Percentage of signals from each cluster of the total signal.

Groups	Cluster 1(%)	Cluster 2(%)	Cluster 3(%)
C-0	16.4	9.6	74.0
C-6	16.8	39.5	43.6
C-12	16.3	41.4	42.4
C-24	18.0	41.5	40.6
C-48	15.9	52.2	31.9

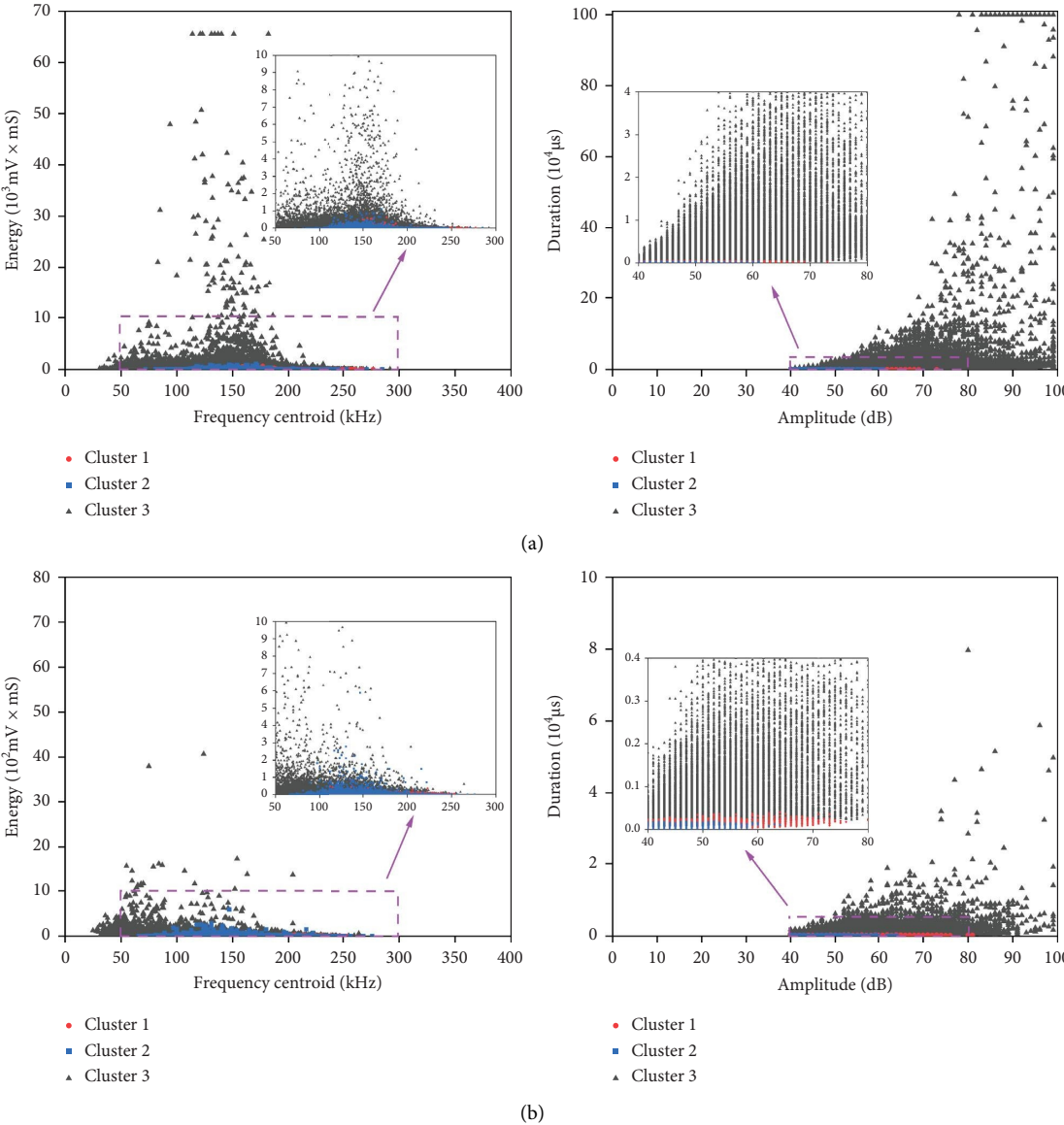


FIGURE 12: Continued.

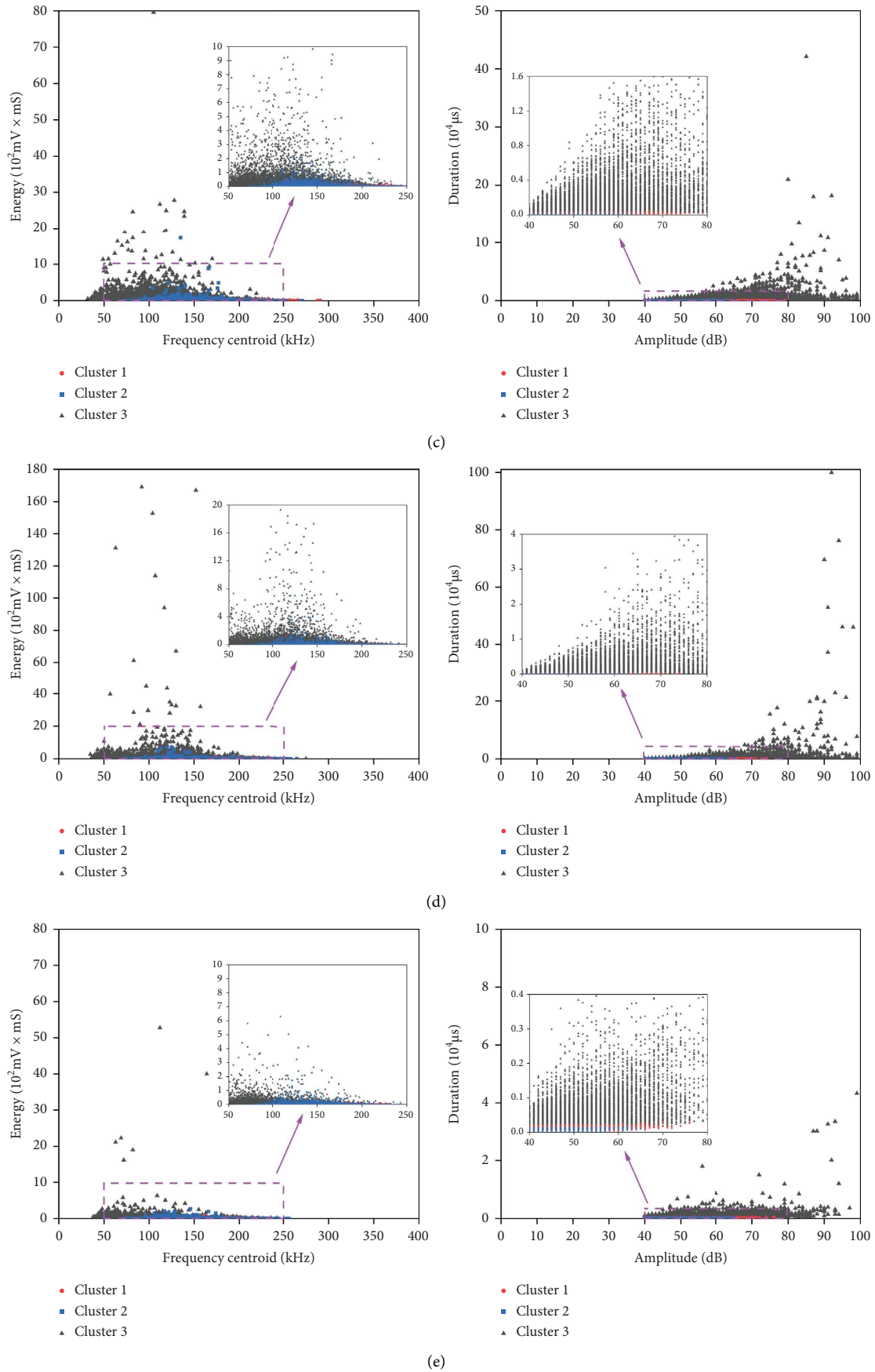


FIGURE 12: Correlation analysis of AE parameters: (a) C-0, (b) C-6, (c) C-12, (d) C-24, and (e) C-48.

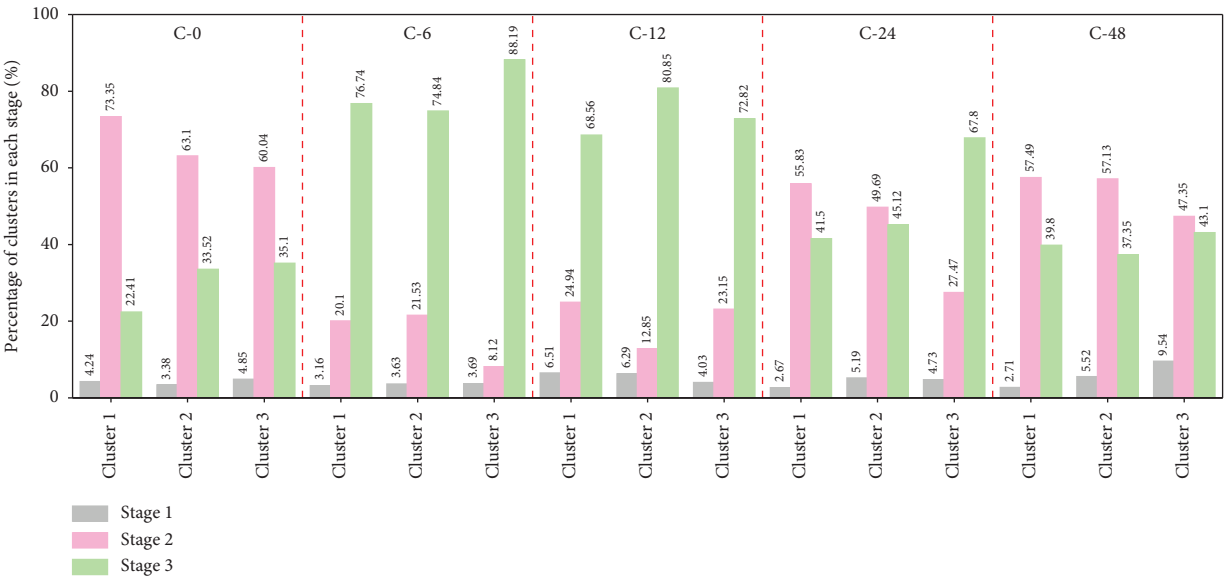


FIGURE 13: The distribution of clusters in each stage.

TABLE 4: Accuracy of test sets with different numbers of hidden layers.

The number of hidden layers	2	3	4	5	6	7	8
C-0(%)	94.4	95.1	97.2	96.3	98.1	99.1	99.2
C-6(%)	95.2	96.6	98.8	97.6	99.1	99.8	99.7

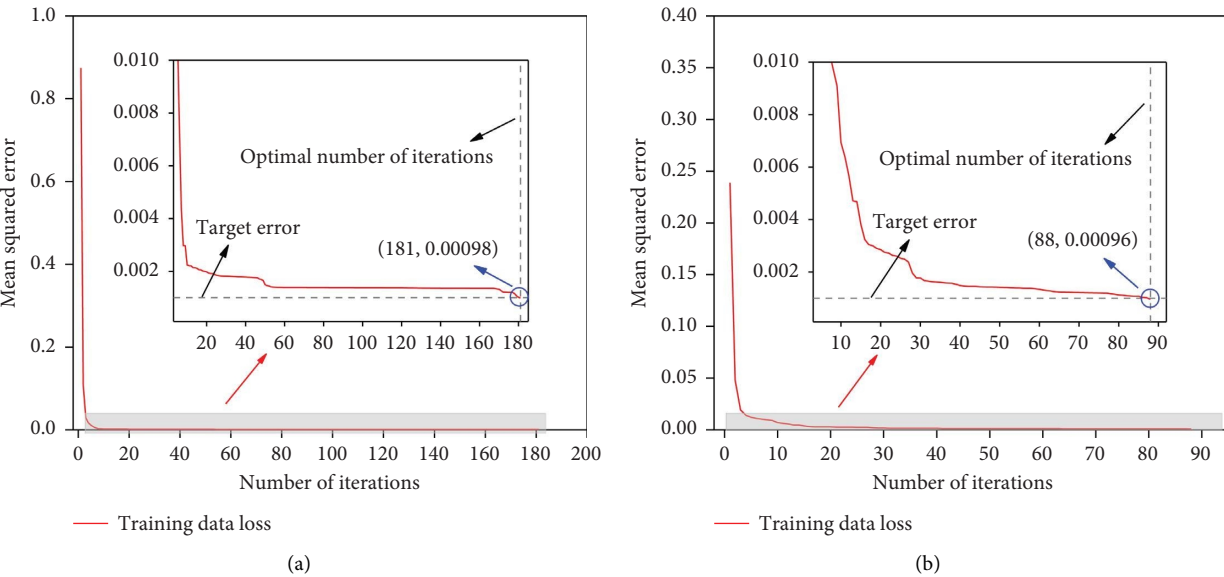
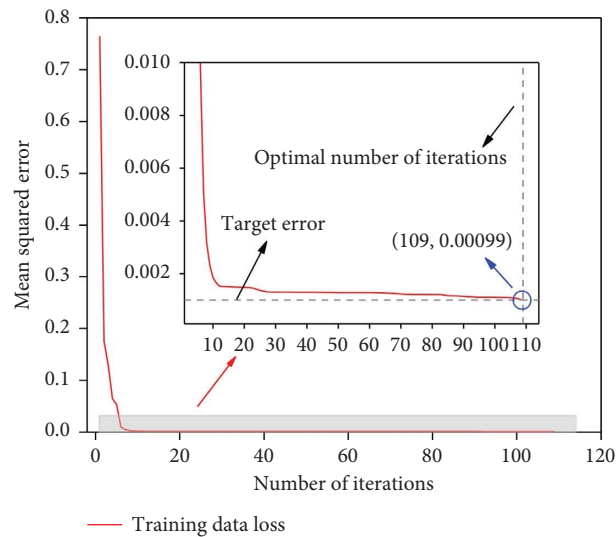


FIGURE 14: Continued.



(c)

FIGURE 14: Schematic diagram of the BP neural network training process: (a) C-0, (b) C-12, and (c) C-48.

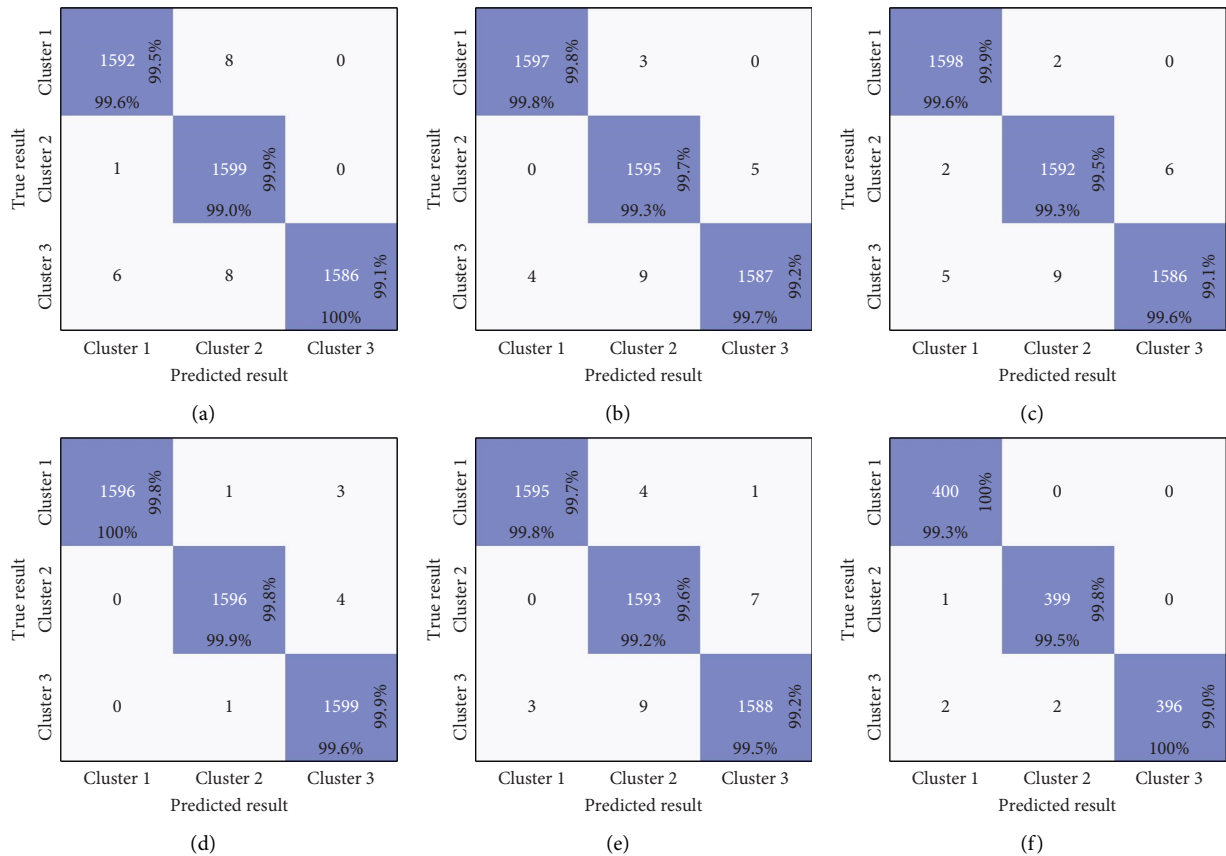


FIGURE 15: Continued.

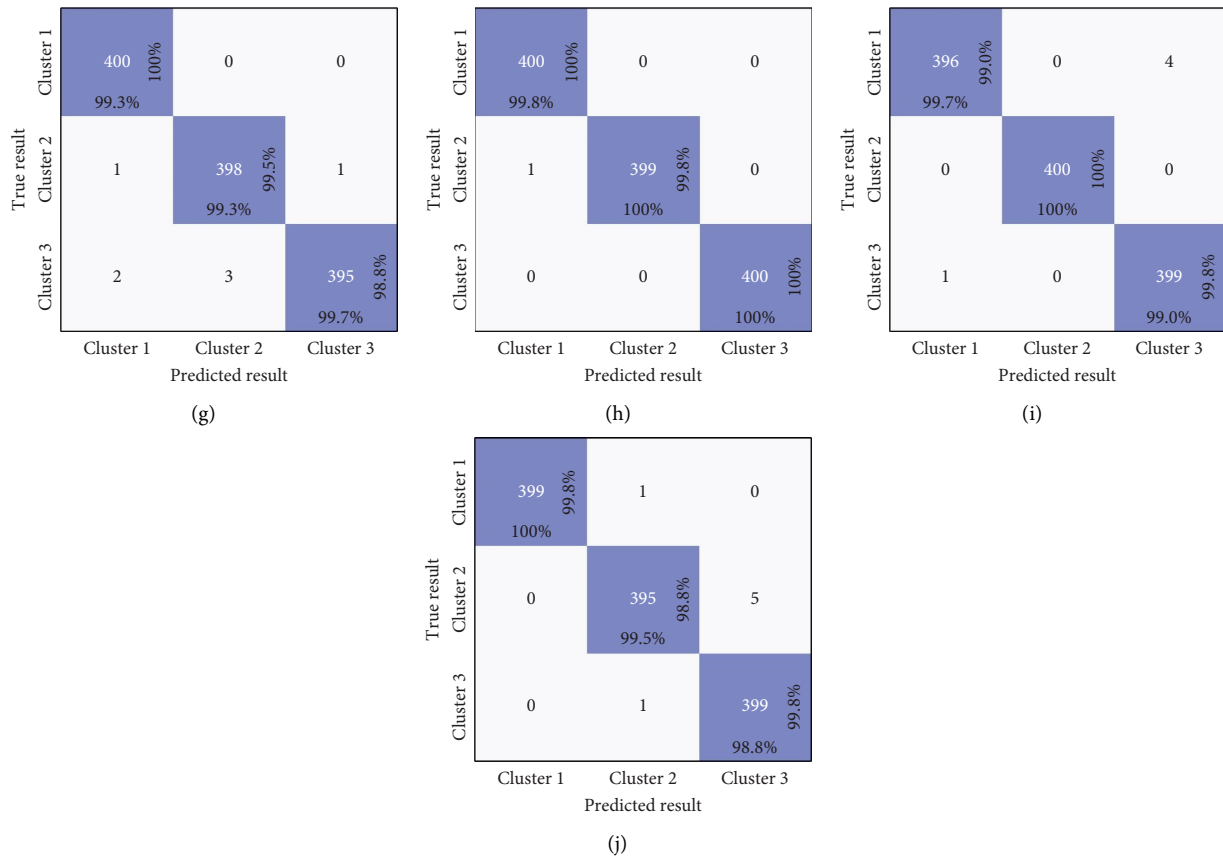


FIGURE 15: Confusion matrix of training data and test data: (a) C-0 train data, (b) C-6 train data, (c) C-12 train data, (d) C-24 train data, (e) C-48 train data, (f) C-0 test data, (g) C-6 test data, (h) C-12 test data, (i) C-24 test data, and (j) C-48 test data.

factors. When ITZ is destroyed, the energy released is weak and the amplitude is small, corresponding to the part with the lowest energy and the smallest amplitude. Due to the varying degrees of cracking of cement mortar, the energy range generated is large, corresponding to the part between the coarse aggregate cracking signal and the ITZ cracking signal.

Each cluster is divided into the three stages distinguished in Chapter 3.1.1, and the proportion of clusters in each stage is shown in Figure 13. It is observed that the three clusters are predominantly distributed in Stages II and III, indicating that the three damage mechanisms primarily occur during the stable and unstable fracture development stages. In Stage II, the frictional motion of dry concrete compression surface, fracture surface activity, and aggregate cracking are predominant. As moisture content increases, the three signal modes of water-bearing concrete exhibit a trend of transitioning from dominance in Stage III to dominance in Stage II. This suggests that increased moisture content may accelerate the development speed and intensity of microcracks in concrete, leading to a shortened stable crack development stage, premature cracking, and ultimately a reduction in strength.

3.2.5. Cracking Signal Pattern Prediction Based on the BP Neural Network. Although the clustering algorithm is capable of identifying distinct signal patterns during the

concrete failure process, it fails to achieve real-time prediction of signals generated during the monitoring process. To address this limitation, a real-time signal category prediction model based on the BP neural network is proposed for predicting signal patterns in the real-time monitoring process. The clustering algorithm's recognition output serves as the data label for the BP neural network. The BP neural network [55] is an artificial neural network model that relies on the backpropagation algorithm, and it stands as one of the most extensively utilized neural networks in various applications. The fundamental mechanism of this network encompasses both the forward propagation of signals and the backward propagation of errors. Structurally, a BP neural network consists of three layers: the input layer, hidden layer, and output layer. It refines its performance by adjusting weights and thresholds through the backpropagation algorithm. Specifically, the input layer receives raw data inputs, the hidden layer processes these inputs through complex transformations, and the output layer delivers the final prediction outcomes. Throughout this process, neurons exert their influence progressively layer by layer. In cases where the initial predictions do not meet expected standards, the network initiates the error backpropagation phase. During this phase, the discrepancies between predicted and actual results are used to fine-tune the network's weights and thresholds, thereby enhancing the overall accuracy of the BP neural network.

A substantial volume of training data can prolong the neural network's training duration and lead to issues such as inadequate generalization capability. Consequently, for each sample category, only 6000 instances are selected for model training and evaluation. The 6000 sets of data contain 3 signal modes, and each set of signals contains 4 features, namely, PC1, PC2, PC3, and PC4. The data amount of each signal mode is 2000 groups, and the data label is provided by the clustering result. We take 80% of the data for training the model and 20% for testing the model. In this study, the number of hidden layers was set at 0.5 to 2 times that of the input layer. The performance of the test set was monitored as the number of hidden layers was gradually increased, thereby determining the optimal number of hidden layers. Taking specimens C-0 and C-6 as examples, Table 4 shows the accuracy rates of the test set when the number of hidden layers was 2–8. When the number of hidden layers was 7 and 8, the accuracy rate of the test set reached over 99%, so the number of hidden layers was set to 7. We set the maximum number of iterations for the model to 1000, the target training error to 0.001, and the learning rate to 0.01. The performance variation of the mean square error during the training process is illustrated in Figure 14. As the number of iterations increases, the mean square error of the BP neural network gradually decreases. The model can basically achieve the set target error within 200 iterations. The result analysis is represented by a confusion matrix. The confusion matrix shows how the predicted label compares to the real label, helping to identify how the model performs in different categories. Figure 15 shows the confusion matrix for the training set and the test set. As can be seen from Figure 15, the BP neural network model is fully trained, and the accuracy rate of the test set is basically above 99%, which can accurately identify different signal modes in the whole process of concrete compression. This further demonstrates the accuracy of the clustering outcomes, as the training of neural networks inherently demands a high-quality dataset. When a new signal generated during the compression process is input to the model, the BP neural network prediction model can give the correct prediction.

4. Conclusions

In this study, uniaxial compression tests and AE monitoring were conducted on concrete specimens with varying moisture contents. The temporal and frequency-domain changes in AE signal parameters were systematically analyzed. Furthermore, the tension and shear cracks were distinguished using the AE RA-AF method, allowing for an in-depth analysis of the crack development mechanisms. With moisture content as the main line, the ML algorithm is used to identify and predict the AE signal pattern of the whole process of concrete compression. The study has led to several key conclusions, which are summarized as follows:

- (1) The AE activity in dry concrete is more pronounced and exhibits a gradual reduction as the moisture content increases. The free water in the void inside the concrete can absorb the medium- and high-

frequency signal above 100 kHz. The higher the moisture content, the stronger the absorption effect, and it tends to be stable when it reaches saturated water absorption.

- (2) Based on the AE analysis conducted by RA-AF, the predominant failure mode of concrete is tensile failure, with shear failure comprising only 1.44%–14.09% of all failures. The filling of the void in the concrete and the bearing effect of water reduce the possibility of diagonal shear failure of the specimen.
- (3) The correlation analysis of AE parameters reveals that the relationship between these parameters is a strong nonlinear correlation. KPCA was employed to extract the principal components PC1, PC2, PC3, and PC4. These components effectively capture the majority of AE parameters, with a cumulative variance explanation rate exceeding 95%.
- (4) Based on variations in moisture content, an unsupervised learning algorithm was employed to identify three distinct signal modes during the entire process of concrete compression fracture: frictional motion signals of the compression surface, fracture surface activity signals, and aggregate cracking signals. These three signal modes predominantly manifest during the middle and late stages of concrete failure. With the increase of moisture content, the activity of the fracture surface is gradually enhanced, and the cracking degree of aggregate is gradually weakened.
- (5) The BP neural network was employed to predict the three signal modes, achieving a test set accuracy exceeding 99%. This model can reliably distinguish between different signal patterns throughout the entire process of concrete compression.

Data Availability Statement

Some or all data, models, or codes that support the findings of this study are available from the authors upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions

Aiping Yu and Feng Fu designed the research methodology. Tao Liu performed the analysis. Tianjiao Miao and Aiping Yu drafted the manuscript. Xuandong Chen, Xuelian Deng, and Feng Fu reviewed the manuscript. All authors read and approved the final manuscript.

Funding

This research was funded by the National Natural Science Foundation of China (Grant No. 51968014), the Guangxi Natural Science Foundation (Grant No. 022GXNSFAA035553),

and the Innovation Project of Guangxi Graduate Education (Grant No. YCSW2025413).

References

- [1] S. Pereira, F. Magalhães, J. P. Gomes, Á. Cunha, and J. V. Lemos, "Dynamic Monitoring of a Concrete Arch Dam During the First Filling of the Reservoir," *Engineering Structures* 174 (2018): 548560.
- [2] X. Niu, Q. Li, W. Liu, and Y. Hu, "Effects of Ambient Temperature, Relative Humidity and Wind Speed on Interlayer Properties of Dam Concrete," *Construction and Building Materials* 260 (2020): 119791, <https://doi.org/10.1016/j.conbuildmat.2020.119791>.
- [3] Y. Zhu, J. Zhou, B. Zhang, H. Wang, and M. Huang, "Statistical Analysis of Major Tunnel Construction Accidents in China From 2010 to 2020," *Tunnelling and Underground Space Technology* 124 (2022): 104460, <https://doi.org/10.1016/j.tust.2022.104460>.
- [4] J. Sun, J. Zhang, W. Huang, L. Zhu, Y. Liu, and J. Yang, "Investigation and Finite Element Simulation Analysis on Collapse Accident of Heyuan Dongjiang Bridge," *Engineering Failure Analysis* 115 (2020): 104655, <https://doi.org/10.1016/j.engfailanal.2020.104655>.
- [5] B. Liu, P. Li, and C. Gao, "Seismic Response Analysis of Bridge Pier in Water by Different Water Depth," *China Civil Engineering Journal* 43, no. S2 (2010): 199–203.
- [6] Z. Xing, L. Zhang, M. Tan, M. Zhou, and H. Huang, "Research on the Influence of Tide on the Stability of Subsea Tunnel Face," *Chinese Journal of Underground Space and Engineering* 20, no. S1 (2024): 381–390.
- [7] K. C. Hover, "The Influence of Water on the Performance of Concrete," *Construction and Building Materials* 25, no. 7 (2011): 3003–3013, <https://doi.org/10.1016/j.conbuildmat.2011.01.010>.
- [8] H. Wang, L. Wang, Y. Song, and J. Wang, "Influence of Free Water on Dynamic Behavior of Dam Concrete Under Biaxial Compression," *Construction and Building Materials* 112 (2016): 222–231, <https://doi.org/10.1016/j.conbuildmat.2016.02.090>.
- [9] X. H. Vu, Y. Malecot, L. Daudeville, and E. Buzaud, "Experimental Analysis of Concrete Behavior Under High Confinement: Effect of the Saturation Ratio," *International Journal of Solids and Structures* 46, no. 5 (2009): 1105–1120, <https://doi.org/10.1016/j.ijsolstr.2008.10.015>.
- [10] H. Li, S. Meng, D. Shi, Q. Wei, Z. Xu, and W. Zhao, "Influence of Moisture on Ultrasonic Propagation, Acoustic Emission Activity, and Failure Mechanism in Concrete Media," *Construction and Building Materials* 386 (2023): 131499, <https://doi.org/10.1016/j.conbuildmat.2023.131499>.
- [11] P. G. Ranjith, D. Jasinge, J. Y. Song, and S. K. Choi, "A Study of the Effect of Displacement Rate and Moisture Content on the Mechanical Properties of Concrete: Use of Acoustic Emission," *Mechanics of Materials* 40, no. 6 (2008): 453–469, <https://doi.org/10.1016/j.mechmat.2007.11.002>.
- [12] B. D. Liu, W. J. Lv, L. Li, and P. F. Li, "Effect of Moisture Content on Static Compressive Elasticity Modulus of Concrete," *Construction and Building Materials* 69 (2014): 133–142, <https://doi.org/10.1016/j.conbuildmat.2014.06.094>.
- [13] X. Sun, Y. Tian, W. Yin, and H. Wang, "Effect of Free Water on Fatigue Performance of Concrete Subjected to Compressive Cyclic Load," *Construction and Building Materials* 318 (2022): 125995, <https://doi.org/10.1016/j.conbuildmat.2021.125995>.
- [14] Z. Zheng, H. Liu, M. Xiao, J. D. He, H. Q. Xie, and L. Zhuo, "A Creep Model Coupling Moisture and Mechanical Damage for Water-Bearing Concrete," *Construction and Building Materials* 326 (2022): 126598, <https://doi.org/10.1016/j.conbuildmat.2022.126598>.
- [15] Z. Zhou, J. Lu, X. Cai, Y. Rui, and L. Tan, "Water Saturation Effects on Mechanical Performances and Failure Characteristics of Rock-Concrete Disc With Different Interface Dip Angles," *Construction and Building Materials* 324 (2022): 126684, <https://doi.org/10.1016/j.conbuildmat.2022.126684>.
- [16] A. Fahim, S. De Carufel, P. Ghods, A. R. Alizadeh, and M. Salehi, "Practical Model for Predicting Internal Relative Humidity of Concrete Exposed to Drying," *Journal of Materials in Civil Engineering* 31, no. 8 (2019): 04019143, [https://doi.org/10.1061/\(asce\)mt.1943-5533.0002785](https://doi.org/10.1061/(asce)mt.1943-5533.0002785).
- [17] L. Dong, Y. Zhang, S. Bi, J. Ma, Y. Yan, and H. Cao, "Uncertainty Investigation for the Classification of Rock Micro-Fracture Types Using Acoustic Emission Parameters," *International Journal of Rock Mechanics and Mining Sciences* 162 (2023): 105292, <https://doi.org/10.1016/j.ijrmms.2022.105292>.
- [18] C. U. Grosse and F. Finck, "Quantitative Evaluation of Fracture Processes in Concrete Using Signal-Based Acoustic Emission Techniques," *Cement and Concrete Composites* 28, no. 4 (2006): 330–336, <https://doi.org/10.1016/j.cemconcomp.2006.02.006>.
- [19] K. Ohno, K. Uji, A. Ueno, and M. Ohtsu, "Fracture Process Zone in Notched Concrete Beam Under Three-Point Bending by Acoustic Emission," *Construction and Building Materials* 67 (2014): 139–145, <https://doi.org/10.1016/j.conbuildmat.2014.05.012>.
- [20] Y. Gao and H. Sun, "Influence of Initial Defects on Crack Propagation of Concrete Under Uniaxial Compression," *Construction and Building Materials* 277 (2021): 122361, <https://doi.org/10.1016/j.conbuildmat.2021.122361>.
- [21] P. R. Prem, M. Verma, and P. S. Ambily, "Damage Characterization of Reinforced Concrete Beams Under Different Failure Modes Using Acoustic Emission," *Structures* 30 (2021): 174–187, <https://doi.org/10.1016/j.istruc.2021.01.007>.
- [22] A. Yu, Z. Chen, L. Zhang, X. Li, J. Shi, and F. Fu, "Study on AE Characteristics of Concrete With Different w/c Ratio Under Uniaxial Compression," *Structures* 58 (2023): 105443, <https://doi.org/10.1016/j.istruc.2023.105443>.
- [23] Z. Chen, T. Miao, T. Liu, X. Chen, and A. Yu, "Active-Passive Joint Acoustic Emission Monitoring Test Considering the Heterogeneity of Concrete," *Materials* 16, no. 24 (2023): 7694, <https://doi.org/10.3390/ma16247694>.
- [24] B. Chen and J. Liu, "Investigation of Effects of Aggregate Size on the Fracture Behavior of High Performance Concrete by Acoustic Emission," *Construction and Building Materials* 21, no. 8 (2007): 1696–1701, <https://doi.org/10.1016/j.conbuildmat.2006.05.030>.
- [25] J. Lee, H. Kim, and T. Oh, "Acoustic Emission Characteristics During Uniaxial Compressive Loading for Concrete Specimens According to Sand Content Ratio," *KSCE Journal of Civil Engineering* 24, no. 9 (2020): 2808–2823, <https://doi.org/10.1007/s12205-020-5697-0>.
- [26] Z. Shi, X. Chen, Y. Ning, and H. Tian, "Study on Crack Propagation of Rubber Self-Compacting Concrete Based on RA-AF Characteristics," *Journal of Civil & Environmental Engineering* 46, no. 05 (2024): 175–183.
- [27] J. Yang, Z. Mu, and S. Yang, "Experimental Study of Acoustic Emission Multi-Parameter Information Characterizing Rock Crack Development," *Engineering Fracture Mechanics* 232 (2020): 107045, <https://doi.org/10.1016/j.engfracmech.2020.107045>.
- [28] Z. Zhang, X. Liu, Y. Zhang, X. Qin, and M. Khan, "Comparative Study on Fracture Characteristics of Coal and Rock Samples Based on Acoustic Emission Technology,"

- Theoretical and Applied Fracture Mechanics* 111 (2021): 102851, <https://doi.org/10.1016/j.tafmec.2020.102851>.
- [29] X. Luo, H. Haya, T. Inaba, and T. Shiotani, "Seismic Diagnosis of Railway Substructures by Using Secondary Acoustic Emission," *Soil Dynamics and Earthquake Engineering* 26, no. 12 (2006): 1101–1110, <https://doi.org/10.1016/j.soildyn.2006.03.002>.
- [30] H. Liu, F. Briffod, T. Shiraiwa, M. Enoki, and S. Emura, "Clustering Analysis of Acoustic Emission Signals During Compression Tests in Mille-Feuille Structure Materials: Mechanics of Materials," *Materials Transactions* 63, no. 3 (2022): 319–328, <https://doi.org/10.2320/matertrans.mt-m2021105>.
- [31] H. Heidary, N. Z. Karimi, M. Ahmadi, A. Rahimi, and A. Zucchelli, "Clustering of Acoustic Emission Signals Collected During Drilling Process of Composite Materials Using Unsupervised Classifiers," *Journal of Composite Materials* 49, no. 5 (2015): 559–571, <https://doi.org/10.1177/0021998314521258>.
- [32] J. Li, G. Du, C. Jiang, and S. Jin, "The Classification of Acoustic Emission Signals of 304 Stainless Steel During Stress Corrosion Process Based on K-Means Clustering," *Anti-Corrosion Methods & Materials* 59, no. 2 (2012): 76–80, <https://doi.org/10.1108/00035591211210848>.
- [33] V. Soltangharai, R. Anay, L. Assi, M. Bayat, J. R. Rose, and P. Ziehl, "Analyzing Acoustic Emission Data to Identify Cracking Modes in Cement Paste Using an Artificial Neural Network," *Construction and Building Materials* 267 (2021): 121047, <https://doi.org/10.1016/j.conbuildmat.2020.121047>.
- [34] V. Radhika and J. M. Chandra Kishen, "A Comparative Study of Crack Growth Mechanisms in Concrete Through Acoustic Emission Analysis: Monotonic Versus Fatigue Loading," *Construction and Building Materials* 432 (2024): 136568, <https://doi.org/10.1016/j.conbuildmat.2024.136568>.
- [35] P. Gao, J. Liu, X. Wang, Y. Jiao, and W. Shan, "Damage Evaluation and Failure Mechanism Analysis of Axially Compressed Circular Concrete-Filled Steel Tubular Column via AE Monitoring," *Structural Health Monitoring* 23, no. 2 (2024): 701–713, <https://doi.org/10.1177/14759217231174697>.
- [36] C. Barile, C. Casavola, G. Pappaletta, and V. P. Kannan, "Interpreting the Lempel–Ziv Complexity of Acoustic Emission Signals for Identifying Damage Modes in Composite Materials," *Structural Health Monitoring* 22, no. 3 (2023): 1708–1720, <https://doi.org/10.1177/14759217221112831>.
- [37] K. He, X. Liu, Q. Yang, and Y. Chen, "An Extraction Method of Welding Crack Acoustic Emission Signal Using Harmonic Analysis," *Measurement* 103 (2017): 311–320, <https://doi.org/10.1016/j.measurement.2017.02.026>.
- [38] Z. Liu, X. Chen, T. Ji, and Z. Peng, "Fracture Properties of Full-Graded Dam Concrete Under Discontinuous Cyclic Loading Based on Acoustic Emission," *Journal of Materials in Civil Engineering* 37, no. 4 (2025): 18138, <https://doi.org/10.1061/jmcee7.mteng-18138>.
- [39] Y. Bi, Z. Yao, Y. Qin, Q. Chen, C. Yang, and A. Kali, "Experimental Study on Acoustic Emission Damage in Precast Reinforced Concrete Interior Joints Containing Disc Springs," *Structures* 69 (2024): 107271, <https://doi.org/10.1016/j.istruc.2024.107271>.
- [40] Z. Lv, C. Chen, R. Li, J. Jin, and B. Hu, "Multifractal of Acoustic Emission for the Multi-Scale Fracture Behavior of Diatomite Modified Concrete," *Construction and Building Materials* 445 (2024): 137951, <https://doi.org/10.1016/j.conbuildmat.2024.137951>.
- [41] W. Shen, F. Wang, C. Li, Z. Zhang, and M. Hou, "Revealing Evolution Law and Failure Mechanism of Interface Damage in Concrete-Encased CFST Columns by Acoustic Emission Technology," *Journal of Building Engineering* 93 (2024): 109792, <https://doi.org/10.1016/j.job.2024.109792>.
- [42] S. Du, B. Liang, Y. Zhang, et al., "Mechanical Properties and Damage Characteristics Analysis on Recycled Aggregate Concrete With Glazed Hollow Beads After High Temperatures by Acoustic Emission Method," *Journal of Building Engineering* 90 (2024): 109429, <https://doi.org/10.1016/j.job.2024.109429>.
- [43] Y. Ma, M. Liu, L. Yang, P. Dai, J. Fan, and E. Tsangouri, "Acoustic Emission for Monitoring the Damage Progress on Concrete-Filled Stainless-Steel Tubes (CFSST) Under Bending Test," *Engineering Structures* 314 (2024): 118316, <https://doi.org/10.1016/j.engstruct.2024.118316>.
- [44] K. Zhao, H. Ma, C. Yang, and J. J. K. Daemen, "The Role of Prior Creep Duration on the Acoustic Emission Characteristics of Rock Salt Under Cyclic Loading," *International Journal of Rock Mechanics and Mining Sciences* 157 (2022): 105166, <https://doi.org/10.1016/j.ijrmms.2022.105166>.
- [45] S. Li, S. Chang, P. Li, X. Zhang, and N. Jiang, "Optimization of Concrete Surface Sensor Arrangement for Acoustic Emission Monitoring of Prestressed Steel Strand Damage in T-Beams," *Applied Acoustics* 223 (2024): 110082, <https://doi.org/10.1016/j.apacoust.2024.110082>.
- [46] B. Hu, S. Chen, D. Yang, P. Yu, and Z. Liao, "Study on the Evolution of Acoustic Signals and Strain Evolution During Large-Scale Tensile Failure Test of Limestone," *Water Resources and Hydropower Engineering* (2024): 1–20.
- [47] J. Guo, S. Liang, H. Zhang, F. Sun, and B. He, "Experimental Study on Rockburst of Granite With Prefabricated Folded Fissure Under True-triaxial With Single Face Unloading," *Journal of Basic Science and Engineering* (2024): 1–20.
- [48] C. Barile, C. Casavola, G. Pappaletta, and V. Paramasamy Kannan, "Damage Assessment of Carbon Fibre Reinforced Plastic Using Acoustic Emission Technique: Experimental and Numerical Approach," *Structural Health Monitoring* 20, no. 3 (2021): 1090–1101, <https://doi.org/10.1177/1475921720946438>.
- [49] P. Lin, X. Feng, W. Guo, J. Wu, X. Tang, and X. Wang, "Plastic Deformation Defect Detection of TiB/TC4 Composites Based on Acoustic Emission," *Journal of Netshape Forming Engineering* 16, no. 07 (2024): 96–108.
- [50] C. Wu, G. Ma, H. J. Hwang, and D. J. Kim, "Machine Learning-based Identification of the Relationship Between Bond Strength of GFRP Bars and AE Hits in Concrete Beams," *Engineering Structures* 329 (2025): 119845, <https://doi.org/10.1016/j.engstruct.2025.119845>.
- [51] Y. Wang, X. Li, Y. Sun, J. Zhang, P. Guo, and R. Wang, "A Monitoring and Warning Method for Hydroelectric Units Based on KPCA and LSTM," *Journal of Vibration and Shock* 43, no. 24 (2024): 287–294.
- [52] S. L. Li, S. H. Chen, M. Li, et al., "Separation Method for Acoustic Emission Signals of Prestressing Steel Strand Fracture and Concrete Cracking," *Measurement* 245 (2025): 116607, <https://doi.org/10.1016/j.measurement.2024.116607>.

- [53] M. Solarte-Sanchez, D. Marquez-Viloria, A. E. Castro-Ospina, E. Reyes-Vera, N. Guerrero-Gonzalez, and J. Botero-Valencia, "m-QAM Receiver Based on Data Stream Spectral Clustering for Optical Channels Dominated by Nonlinear Phase Noise," *Algorithms* 17, no. 12 (2024): 0553, <https://doi.org/10.3390/a17120553>.
- [54] L. Yin, Y. Wang, H. Chen, and W. Deng, "An Improved Density Peak Clustering Algorithm for Multi-Density Data," *Sensors* 22, no. 22 (2022): 8814, <https://doi.org/10.3390/s22228814>.
- [55] A. Jierula, S. Wang, T. Oh, J. W. Lee, and J. H. Lee, "Detection of Source Locations in RC Columns Using Machine Learning With Acoustic Emission Data," *Engineering Structures* 246 (2021): 112992, <https://doi.org/10.1016/j.engstruct.2021.112992>.