International Scientific Technical and Economic Research (ISSN: 2959-1309)
Vol. 4, No. 2, (June 2026), pp. 206-231
Journal homepage: https://www.istaer.online

RESEARCH ARTICLE DOI: 10.71451/ISTAER2622

Comparison and Analysis of Multiple Entropy Feature Tests for Abnormal Signals in Fixed-Point Deformation Monitoring

Sirui Liu[1], Cong Pang[1,2]*, Xin Wang[3], Meiping Song[4]

[1]*Institute of Seismology, China Earthquake Administration, Wuhan, Hubei, China*

[2]*Wuhan Gravitation and Solid Earth Tides, National Observation and Research Station, Wuhan, Hubei, China*

[3]*Shenyang Earthquake Monitoring Center, Liaoning Earthquake Administration, Shenyang, Liaoning, China*

[4]*Datong Earthquake Monitoring Center, Shanxi Earthquake Administration, Datong, Shanxi, China*

Abstract: Fixed-point deformation monitoring is essential for geological hazard early warning, and identifying abnormal signals remains a key challenge. To evaluate the effectiveness of five entropy methods for this purpose, a quantitative comparison was conducted. Abnormal deformation data were preprocessed to extract sample entropy (SE), fuzzy entropy (FE), distribution entropy (DE), permutation entropy (PE), and adaptive weighted multi-scale fusion entropy (AWM-FE). Their ability to distinguish abnormal signals was compared using the T-test, followed by the Kruskal-Wallis test and post hoc multiple comparisons. Simulation experiments showed that AWM-FE exhibited stable, reliable performance and was well-suited for complex field environments with multi-scale analysis needs. Real deformation data analysis revealed that SE had an average T-test p-value of 0.0016, indicating significant distinction across six category pairs, while DE achieved an F-value of 74.7205 in ANOVA, reflecting the largest overall inter-group variation. This study provides a reference for feature selection in identifying abnormal signals in fixed-point deformation observations.

Keywords: Sample entropy; Fuzzy entropy; Distribution entropy; Permutation entropy; Adaptive weighted multi-scale fusion entropy

How to Cite: Liu, S., Pang, C., Wang, X., & Song, M. (2026). Comparison and Analysis of Multiple Entropy Feature Tests for Abnormal Signals in Fixed-Point Deformation Monitoring. *International Scientific Technical and Economic Research, 4*(2), 206–231. **https://doi.org/10.71451/ISTAER2622**

Article history: Received: 26 Mar 2026; **Revised:** 23 Apr 2026; **Accepted:** 28 Apr 2026; **Published:** 29 May 2026

1. INTRODUCTION

1.1 Research background

Point deformation monitoring captures minute crustal deformation characteristics,

*Corresponding author: Cong Pang, Institute of Seismology, China Earthquake Administration, Wuhan, Hubei, China. Email: ponspc@hubdzj.gov.cn

providing critical data support for active tectonic research, seismic incubation pattern analysis, and potential seismic hazard risk assessment. It constitutes an essential component of earthquake monitoring and early warning systems [1],[2]. Seismic observation requires precise identification of anomalous signals, which serve as critical indicators for detecting precursory crustal deformation and predicting seismic activity trends [3],[4],[5],[6],[7]. Anomaly detection methods often rely on technicians' manual experience, potentially leading to misjudgments due to strong subjectivity. With the ongoing demand for enhanced monitoring accuracy, real-time capabilities, and intelligence, constructing datasets through anomaly signal feature analysis and integrating artificial intelligence technologies can meet the technical requirements for high-precision monitoring [8],[9],[10],[11].

Entropy features can accurately capture signal irregularities and exhibit excellent noise robustness, overcoming the limitations of traditional linear analysis in characterizing the complexity of signal dynamics. They demonstrate significant advantages in nonlinear signal analysis [12],[13],[14]. Leveraging their robust signal analysis capabilities, entropy-based methods are now widely applied in mechanical and circuit fault detection. Consequently, the advantages of entropy methods in nonlinear signal analysis can be harnessed to construct feature sets for anomaly signals in fixed-point deformation anomaly detection. These feature sets can then be integrated into artificial intelligence algorithms to achieve automated anomaly signal recognition.

Currently, there is limited research on the comparative analysis and quantitative evaluation of entropy features for abnormal signals in fixed-point deformation monitoring. This study not only selects commonly used entropy features such as SE, FE, DE, and PE, but also introduces AWM-FE as a research subject. It employs t-tests, ANOVA, and multiple comparison techniques to conduct comparative experiments on entropy features of abnormal signals. SE quantifies the probability of signal repetition patterns in phase space based on Takens' embedding theorem and is suitable for low-noise, stationary continuous signals; FE replaces the hard thresholds in sample entropy with fuzzy membership functions and is suitable for continuous signals with strong noise; DE quantifies distribution uncertainty through kernel density estimation or histogram partitioning, and is suitable for non-stationary signals with significant variations in distribution; PE is highly effective at resisting noise based on signal ordinal patterns, and is suitable for nonlinear, non-stationary, short-burst, and anomalous signals; AWM-FE integrates adaptive scaling and weighted fusion mechanisms, and is suitable for complex multiscale signals. By integrating statistical testing, it achieves quantitative evaluation of these five entropy features, systematically comparing the applicability of five entropy measures in fixed-point deformation scenarios.

2. ENTROPY CHARACTERISTICS OF ABNORMAL SIGNALS IN FIXED-POINT DEFORMATION MONITORING

2.1 Characteristics of anomalous signals in fixed-point deformation monitoring

Abnormal signals generated during fixed-point deformation monitoring may originate from four scenarios: power supply failure, lightning anomalies, atmospheric pressure and rainfall interference, and collapse earthquakes. The time-domain diagrams for signals under these four conditions are shown below.

Power supply failures caused by insufficient power or poor contact at soldered power cable joints can result in small data gaps and significant waveform jumps in the time-domain waveform. Following lightning strikes, strong lightning current signals induced or coupled through various lines can disrupt normal data collector operation, generating lightning-induced anomalies that cause partial data loss and dense sudden jumps in the time-domain waveform. Significant atmospheric pressure fluctuations can cause substantial distortion in deformation

observation data. Short-duration heavy rainfall may form runoff along mountain slopes, accumulating in low-lying areas to create small dammed lakes. This increases pressure in the affected zone, causing the vertical pendulum NS observation curve to decline. Consequently, the time-domain waveform exhibits large low-frequency fluctuations and localized abrupt drops. Localized collapse vibrations triggered by geological instability and concentrated stress release produce progressive fluctuations with distinct peaks in the time-domain waveform.

While these anomalies exhibit distinct characteristics in the time-domain waveform, they fundamentally represent nonlinear, non-stationary dynamic processes with continuously evolving statistical properties. Traditional techniques based on signal amplitude or fixed frequency fail to effectively capture their intrinsic patterns.

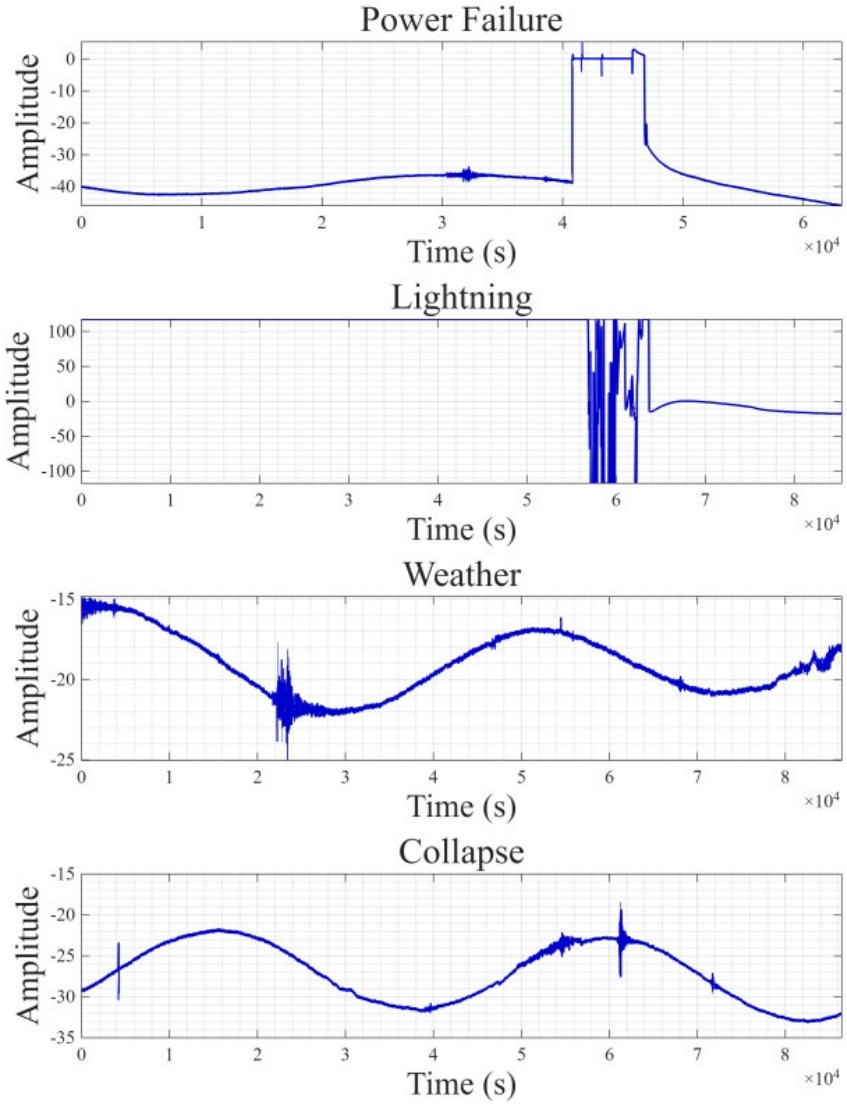

Figure 1. Abnormal signal time domain plot

The entropy method bypasses direct measurement of specific signal amplitudes or frequency components, and instead directly quantifying the complexity of time series. It yields low entropy values for periodic and linear signals, while producing high entropy values for nonlinear, non-stationary complex signals. Different complex structures correspond to distinct entropy values, providing a viable quantitative basis for distinguishing various anomaly signals.

2.2 Sample entropy

SE measures the complexity of a time series by quantifying the probability of new patterns emerging in the signal. However, it exhibits dependence on data length and parameter selection, making it suitable for assessing overall complexity [15],[16].

The time series $\{x(n)\}$ consists of N data points: $\{x(n)\} = \{x(1), x(2), \cdots, x(N)\}$. Reconstruct m-dimensional vectors $X_m(1), X_m(2), \cdots, X_m(N - m + 1)$, where $X_m(i) = \{x(i), x(i + 1), \cdots, x(i + m - 1)\}, 1 \leq i \leq N - m + 1$. The embedding dimension m primarily determines the level of detail in the phase space reconstruction; if the value is too small, it is difficult to adequately capture the signal structure, whilst if it is too large, it is prone to introducing redundant information into the calculations. Compute the Chebyshev distance between $X_m(i)$ and $X_m(j)$, denoted as d_{ij}^m. For a given $X(i)$ count the number $j(1 \leq j \leq N - m, i \neq j)$ of d_{ij}^m values less than or equal to the similarity tolerance threshold r to obtain the approximate count, denoted as B_i. The similarity tolerance r primarily determines the threshold for identifying similar patterns; if set too high, the entropy value tends to stabilise but sensitivity decreases; if set too low, the system becomes susceptible to noise interference. Calculate the ratio of approximate count to total count to obtain the approximation ratio, denoted as $B_i^m(r)$.

$$B_i^m(r) = \frac{B_i}{N - m - 1} \tag{1}$$

Calculate $B^{(m)}(r)$:

$$B^{(m)}(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^m(r) \tag{2}$$

Increase the dimension to m+1. Compute the approximate quantity by counting the number $j(1 \leq j \leq N - m, i \neq j)$ such that the Chebyshev distance between $X_{m+1}(i)$ and $X_{m+1}(j)$ is less than or equal to the similarity tolerance threshold r. Denote this quantity as A_i.

Calculate the ratio of the approximate quantity to the total quantity to obtain the approximate proportion, denoted as $A_i^m(r)$.

$$A_i^m(r) = \frac{A_i}{N - m - 1} \tag{3}$$

Calculate $A^{(m)}(r)$:

$$A^{(m)}(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} A_i^m(r) \tag{4}$$

The SE formula is:

$$SE(m, r, N) = -\ln\left[\frac{A^{(m)}(r)}{B^{(m)}(r)}\right] \tag{5}$$

2.3 Fuzzy entropy

FE addresses the problem of sample entropy being susceptible to baseline drift by using an exponential function in the fuzzy similarity measure and removing the mean [17],[18],[19]. Compared with sample entropy, fuzzy entropy shows stronger parameter robustness and produces more stable calculation results.

Unlike sample entropy, FE subtracts the mean value of each vector's elements when reconstructing the m-dimensional vector. The reconstructed vectors are expressed as $X_m(1)$,

$X_m(2), \cdots, X_m(N\text{-}m+1)$, where $X_m(i) = \{x(i), x(i+1), \cdots, x(i+m-1)\} - x_m^I$, with i ranging from $1 \le i \le N - m + 1$.

Compute the Chebyshev distance between $X_m(i)$ and $X_m(j)$, denoted as d_{ij}^m.

By introducing the fuzzy membership degree n, the similarity degree $D_{ij}^{m,n,r}$ between $X_m(i)$ and $X_m(j)$ is calculated, where r represents the similarity tolerance threshold. When $1 \le j \le N - m$ and $i \ne j$:

$$D_{ij}^{m,n,r} = e^{-\frac{\left(d_{ij}^m\right)^n}{r}} \tag{6}$$

Calculate $\Phi^{m,r,n}$:

$$\Phi^{m,r,n} = \frac{1}{N-m}\sum_{i=1}^{N-m}\left[\frac{1}{N-m-1}\sum_{j=1,j\ne i}^{N-m} D_{ij}^{m,n,r}\right] \tag{7}$$

Increase the dimension to m+1. Calculate the distance d_{ij}^{m+1} between $X_{m+1}(i)$ and $X_{m+1}(j)$. Determine the similarity $D_{ij}^{m+1,n,r}$ between $X_{m+1}(i)$ and $X_{m+1}(j)(1 \le j \le N - m, i \ne j)$:

$$D_{ij}^{m+1,n,r} = e^{-\frac{\left(d_{ij}^{m+1}\right)^n}{r}} \tag{8}$$

Calculate $\phi^{m+1,r,n}$:

$$\Phi^{m+1,r,n} = \frac{1}{N-m}\sum_{i=1}^{N-m}\left[\frac{1}{N-m-1}\sum_{j=1,j\ne i}^{N-m} D_{ij}^{m+1,n,r}\right] \tag{9}$$

The FE formula is:

$$FE(m,r,n) = -\ln\left[\frac{\phi^{m+1,r,n}}{\phi^{m,r,n}}\right] \tag{10}$$

2.4 Distribution entropy

DE addresses the issue of high parameter dependency in sample entropy by utilizing the distribution of distances between vectors in phase space, effectively capturing complexity determined by the statistical properties of signal amplitude [20],[21],[22].

Unlike sample entropy, DE constructs a distance matrix $D = \{d_{ij}\}_{(N-m)(N-m)}$, $1 \le i, j \le N - m$ after calculating the Chebyshev distance d_{ij} between $X_m(i)$ and $X_m(j)$.

The empirical probability density is calculated from the histogram and denoted as f_k. With q histogram bins and a frequency distribution $H=\{h_1, h_2, \cdots, h_q\}$, the empirical probability density function of D is:

$$F_k = \frac{h_k}{\sum_{k=1}^{q} h_k} \tag{11}$$

The DE formula is:

$$DE = -\sum_{k=1}^{q} f_k \log_2 f_k \tag{12}$$

2.5 Permutation entropy

PE is calculated based on the distribution of permutations of adjacent values in a time series, offering fast computation speed and suitability for capturing transient events [23],[24],[25].

If the time series $\{x(n)\}$ consists of N data points, $\{x(n)\} = \{x(1), x(2), \cdots, x(N)\}$. Given an embedding dimension m and a time delay t, the time series undergoes phase space reconstruction to generate k sub-sequences, where $k = N - (m-1)t$. The sub-sequences are defined as $X(i) = \{x(i), x(i)+t, \cdots, x(i)+(m-1)t\}$. The time delay t controls the correlation among embedding vectors and affects the effectiveness of signal feature extraction. Arrange the components of each sub-sequence in ascending order of magnitude. There are m! possible permutations, and the probability $P_j (1 \leq j \leq m!)$ for each permutation.

The formula for permutation entropy is:

$$PE = -\sum_{j=1}^{m!} P_j \ln P_j \tag{13}$$

2.6 AWM_FE

AWM_FE calculates and adaptively weights signal fusion based on signal energy across different timescales, yielding a comprehensive complexity metric that characterizes both local transients and overall evolution. By integrating multiscale analysis with multiple entropy measures, AWM_FE reveals complex signal behavior across diverse temporal scales.

The original time series is coarsened to generate multiple coarse-grained sequences $Y_j^{(s)}$.

$$Y_j^{(s)} = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x_I, j = 1,2, \ldots, \left[\frac{N}{s}\right] \tag{14}$$

$Y_j^{(s)}$ represents the j-th data point of the coarsened sequence, where s is the scaling factor that determines the degree of coarsening, N is the total length of the original time series, and x_I is the i-th data point of the original time series.

For each coarsened sequence, compute the SE, FE, and DE.

Weighted averaging of these three entropies yields the fusion entropy FE_s:

$$FE_s = \alpha \cdot SE\left(y^{(s)}\right) + \beta \cdot FE\left(y^{(s)}\right) + \gamma \cdot DE\left(y^{(s)}\right) \tag{15}$$

Weight the fusion entropy at all scales using energy as the scale weighting factor to obtain the final AWM−FE.

$$AWM - FE = \sum_{s \in S} w_s \cdot FE_s \tag{16}$$

S is the set of all segments, w_s is the weight of the s-th scale, and FE_s is the fusion entropy of the s-th scale.

The formula for calculating w_s is as follows:

$$W_s = \frac{E_s}{\sum_{s' \in S} E'_s} \tag{17}$$

E_s is the original indicator value of the s-th component, calculated as follows:

$$E_s = \frac{1}{T_s} \sum_{t=1}^{T_s} \left[y_t^{(s)} \right]^2 \tag{18}$$

T_s is the sequence length of scale s.

As AWM-FE integrates multi-scale information with various types of entropy features, it exhibits relatively greater robustness to parameter variations; however, the calculation of basic entropy at each scale remains subject to the combined influence of r, m and t.

3. KEY TECHNOLOGIES

3.1 T-test

The t-test determines whether the difference between the means of two sample groups is statistically significant by constructing statistical hypotheses, thereby distinguishing systematic differences from random fluctuations [26],[27].

Assuming two independent sample groups X_1 and X_2 have sample sizes n_1 and n_2 respectively, with means $\bar{X}_1$ and $\bar{X}_2$ and variances s_1^2 and s_2^2 respectively. Under the assumption of equal variances, the degrees of freedom for the t-statistic in an independent samples t-test correspond to:

$$d_f = n_1 + n_2 - 2 \tag{19}$$

The combined variance s_p^2 for two samples is:

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{d_f} \tag{20}$$

The independent samples t-test statistic is:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \tag{21}$$

Based on the calculated degrees of freedom (df) and the t-statistic, the significance of the p-value is determined. The corresponding calculation formula is defined as follows:

$$P = 2 \times [1 - F_t(|t|; df)] \tag{22}$$

Set the significance level α to 0.05. When $p > 0.05$, there is no statistically significant difference between the means of the two sample groups; when $p \leq 0.05$, there is a statistically significant difference between the means of the two sample groups.

3.2 Multiple Comparison Methods

When the Kruskal-Wallis nonparametric analysis of variance reveals a significant overall difference, Dunn's test is used for post hoc multiple comparisons to accurately identify groups with significant differences, thereby addressing the problem of accumulated false positives in multiple testing [28],[29],[30].

Suppose there are k mutually independent samples; the rank sum of the i-th sample is R_i,

the sample size of the i-th sample is n_i, and the total sample size is N.

Calculate the mean rank for each group of samples:

$$\bar{R}_i = \frac{R_i}{n_i} \tag{23}$$

Calculate the mean rank difference between two sample groups:

$$D_{ij} = \bar{R}_i - \bar{R}_j \tag{24}$$

Calculate the standard error of the rank difference:

$$SE_{ij} = \sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \tag{25}$$

Calculate the number of multiple comparisons:

$$C = \frac{k(k-1)}{2} \tag{26}$$

Calculate the corrected significance level:

$$\alpha^* = \frac{\alpha}{C} \tag{27}$$

Based on the significance level α and the number of multiple comparisons C, determine the corrected significance level α^*, and consult a standard normal distribution table to determine the critical value $Z_{\alpha^*/2}$.

Calculate the statistic for Dunn's test:

$$Z_{ij} = D_{ij}/SE_{ij} \tag{28}$$

When $|Z_{ij}| > Z_{\alpha^*/2}$, the distributions of the two groups are judged to be significantly different; when $|Z_{ij}| \leq Z_{\alpha^*/2}$, there is no significant difference between the distributions of the two groups.

4. EXPERIMENTS COMPARING MULTIPLE ENTROPY FEATURES

4.1 Evaluating multiple entropy properties of simulated signals

The experiment was implemented on the MATLAB 2025a platform and Windows 10 operating system. Seven distinct types of test signals were generated and evaluated for their entropy performance across five dimensions: computational efficiency, discrimination capability, robustness, stability, and parameter sensitivity. As shown in **Figure 2**, the test signals comprised pure periodic signals, periodic signals with noise, chaotic signals, white noise, pink noise, fractal Brownian motion signals, and composite signals.

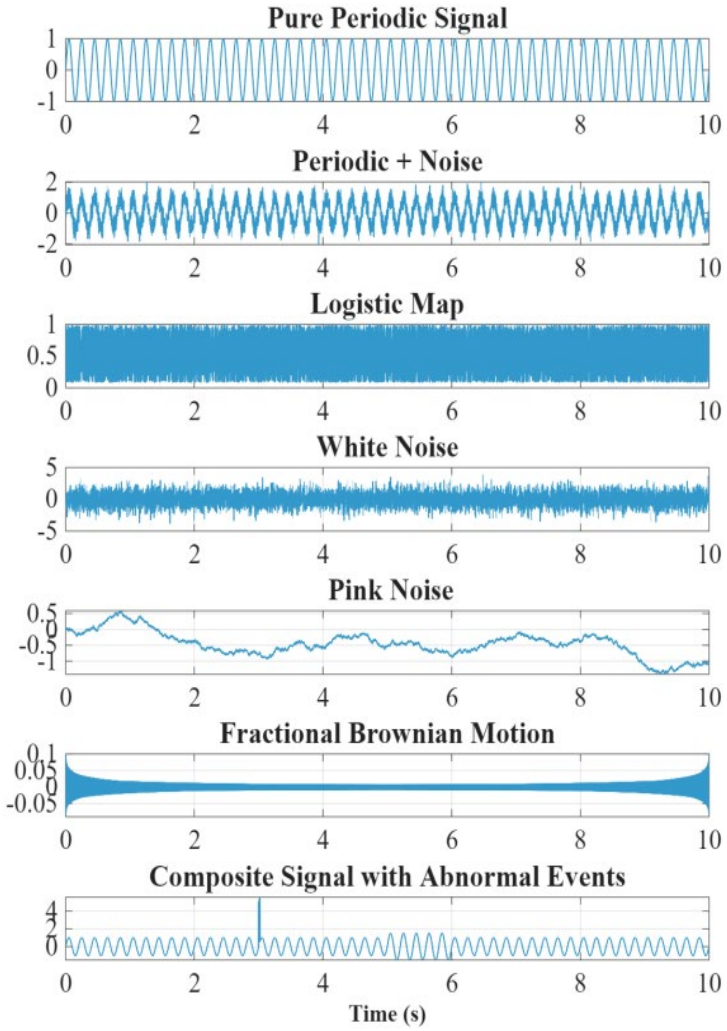

Figure 2. Simulation signal diagram

The computational speed of each entropy method was measured across different signal types to quantify computational efficiency. All tests were conducted under the following conditions: a signal length of 5001, an embedding dimension of 2, and a time delay of 1. The computation times were calculated as the average of five independent runs to ensure fairness and stability. The results are shown in Table 1. As indicated in Table 1, permutation entropy exhibits the shortest computation time and highest computational efficiency, while AWM-FE demonstrates the longest computation time and lowest computational efficiency.

Table 1. Statistical results of computational efficiency for various entropy algorithms

Entropy algorithm	Mean (s)	Std (s)
SE	3.3231	0.4134
FE	5.1367	0.7213
DE	16.7646	0.9104
PE	0.1292	0.0492
AWM-FE	44.1536	12.3107

SE, FE, and DE all require the calculation of pairwise Chebyshev distances between all embedding vectors, and their time complexity is $O(N^2)$. However, unlike SE and FE, which do not require storing intermediate results for all calculated distances, DE must maintain a complete distance distribution histogram, which introduces additional storage and updating overhead. Consequently, DE exhibits a significantly longer computation time than SE and FE, despite sharing the same order of time complexity. PE only requires performing sorting operations on each of the N windows, without the need to traverse all pairs of vectors; its time complexity is $O(N \cdot m \log m)$. When N is large, the growth rate of $O(N \cdot m \log m)$ is far slower than that of $O(N^2)$; therefore, PE is the fastest to compute among these entropies. AWM-FE adds multi-scale decomposition and adaptive weighting steps to the basic fuzzy entropy framework. Its complexity grows linearly with the number of scales, and its overall time complexity is $O(k \cdot N^2)$. Furthermore, the weighting calculations introduce additional statistical and multiplicative overhead, resulting in an overall execution time that is significantly longer than that of the basic single-scale entropies. The entropy value differences across various signal types are illustrated in Figure 3.

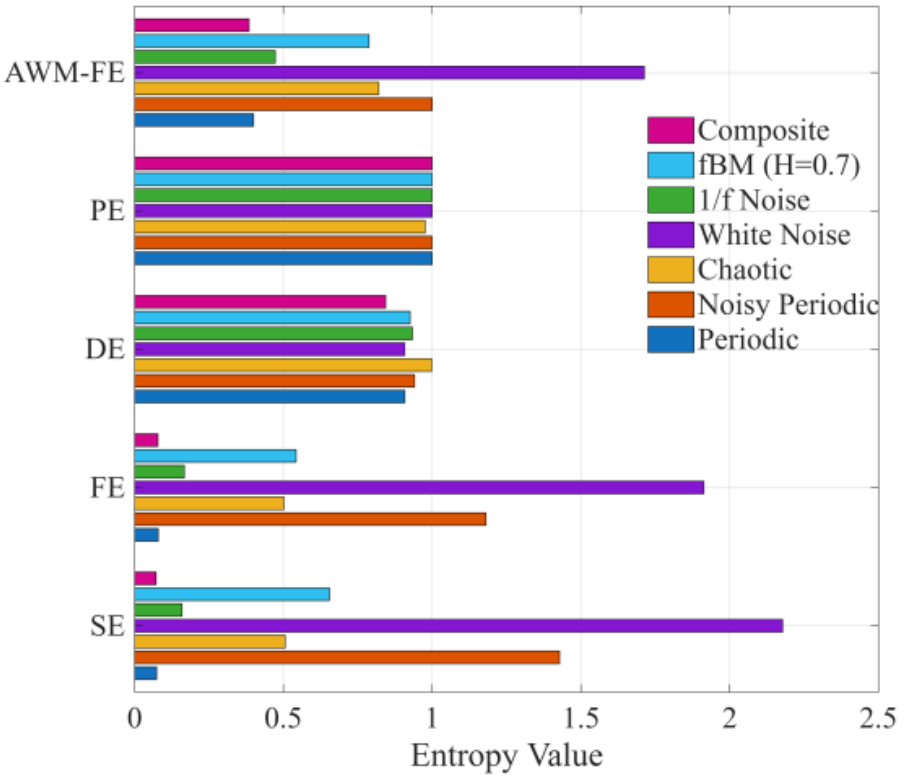

Figure 3. Comparison chart of signal entropy values

As shown in Figure 3, except for PE, the remaining entropy methods exhibit certain amplitude differences under different conditions. By calculating the standard deviation of entropy values for each method, their quantitative discrimination capabilities are quantified, as illustrated in Figure 4.

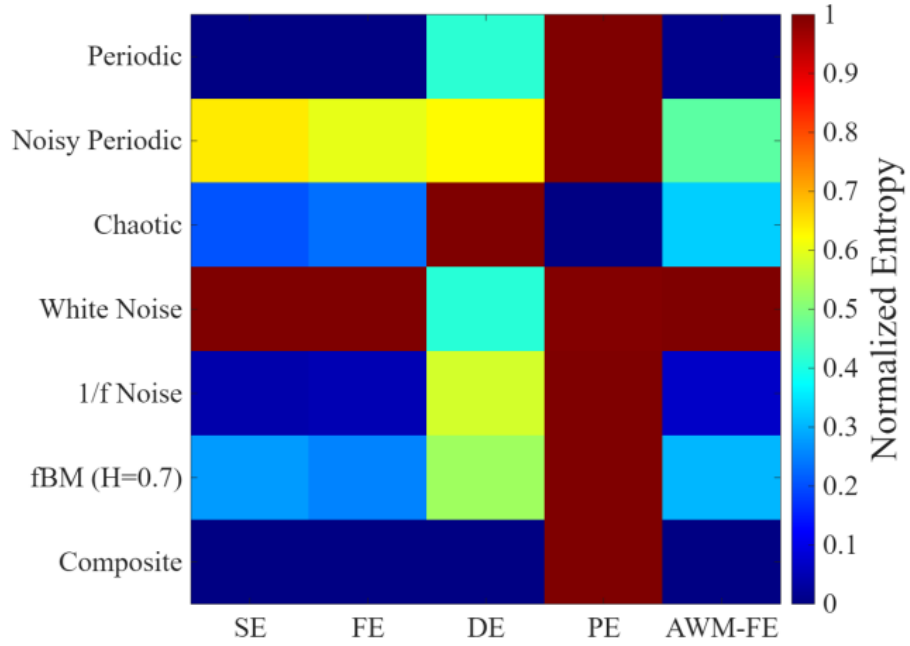

Figure 4. Entropy method and heatmap of standard entropy values for each signal

As shown in Figure 4, DE exhibits significant differences in color gradients compared with AWM-FE, showing high distinguishability among entropy values for different signals.

Gaussian white noise of varying intensities was successively added to a pure periodic signal, causing the signal-to-noise ratio to decrease gradually from 50 dB to –10 dB, in order to simulate the transition from low-level to high-level noise interference. The entropy values were extracted, and the difference between the maximum and minimum entropy values was calculated to quantify the entropy. The robustness of this method varies, as shown in Figure 5 and 6.

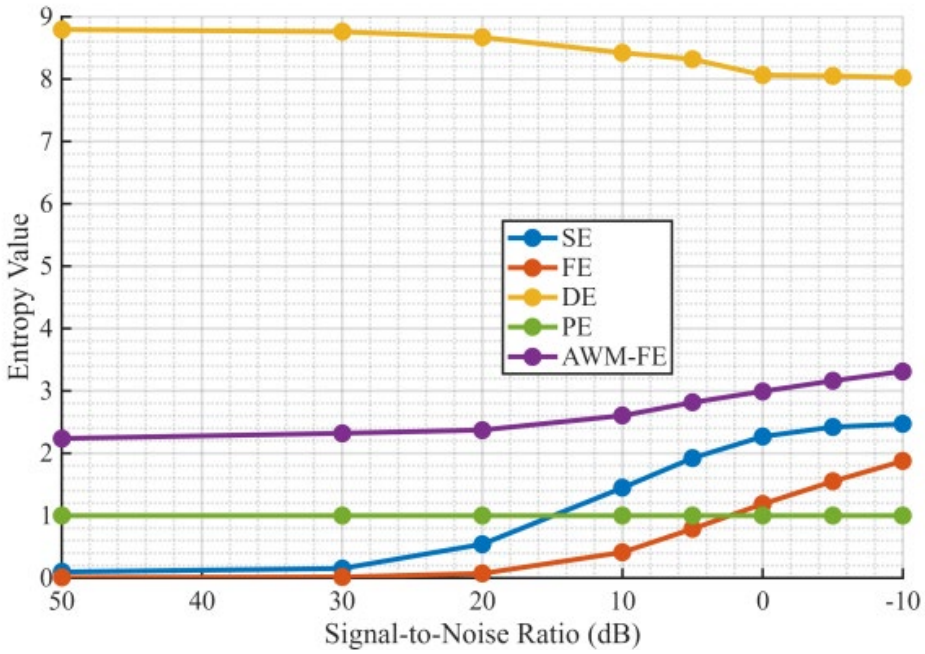

Figure 5. Entropy value variation with noise reduction

As shown in Figure 5, SE and FE are most sensitive to changes in noise, whilst DE and PE demonstrate exceptional robustness but lack sufficient discriminatory power. In contrast, AWM-FE exhibits superior overall performance; its entropy value shows a smooth and monotonic variation as noise increases, ensuring both an effective representation of signal randomness and good resistance to noise interference.

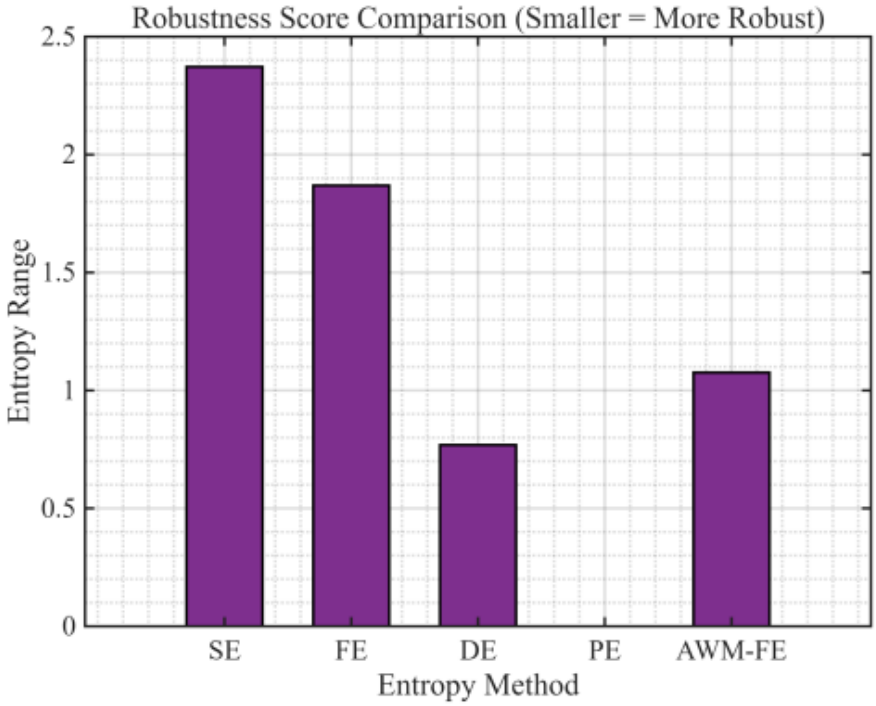

Figure 6. Comparison of noise robustness among different entropy methods

As shown in Figure 6, the difference between the PE, DE, and AWM-FE are minimal, indicating the best robustness.

The stability of the entropy quantification entropy method was assessed by repeatedly applying small perturbations to noisy periodic signals and calculating the variation coefficient of each entropy sequence, as illustrated in Figure 7.

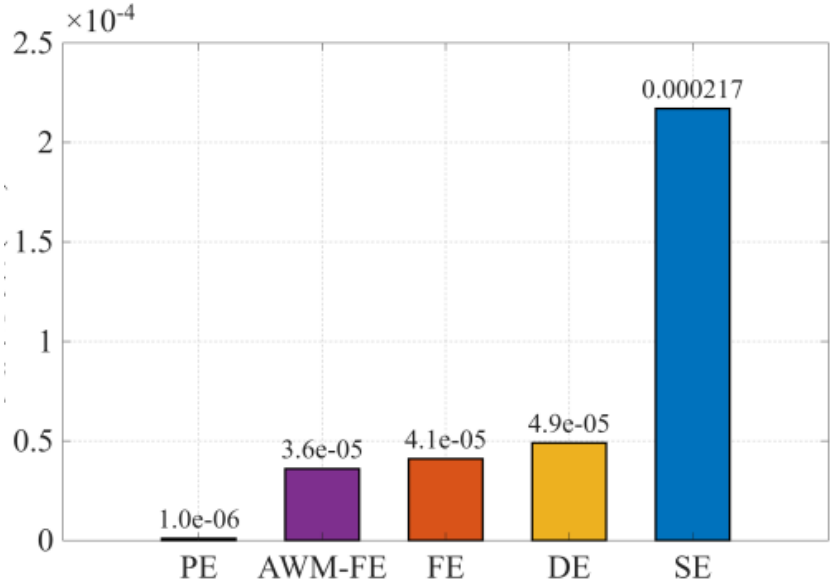

Figure 7. Coefficient of variation for each entropy value

As shown in Figure 7, the coefficient of variation for permutation entropy is the smallest, indicating the highest stability. The coefficients of variation for AWM-FE, FE, and DE are relatively low, suggesting good stability.

Using chaotic signals, we constructed parameter gradient scenarios for the key parameters of various entropy methods and calculated the entropy values, observing the extent to which the various entropies varied with changes in the parameters. The parameter sensitivity heatmap of SE is presented in Figure 8.

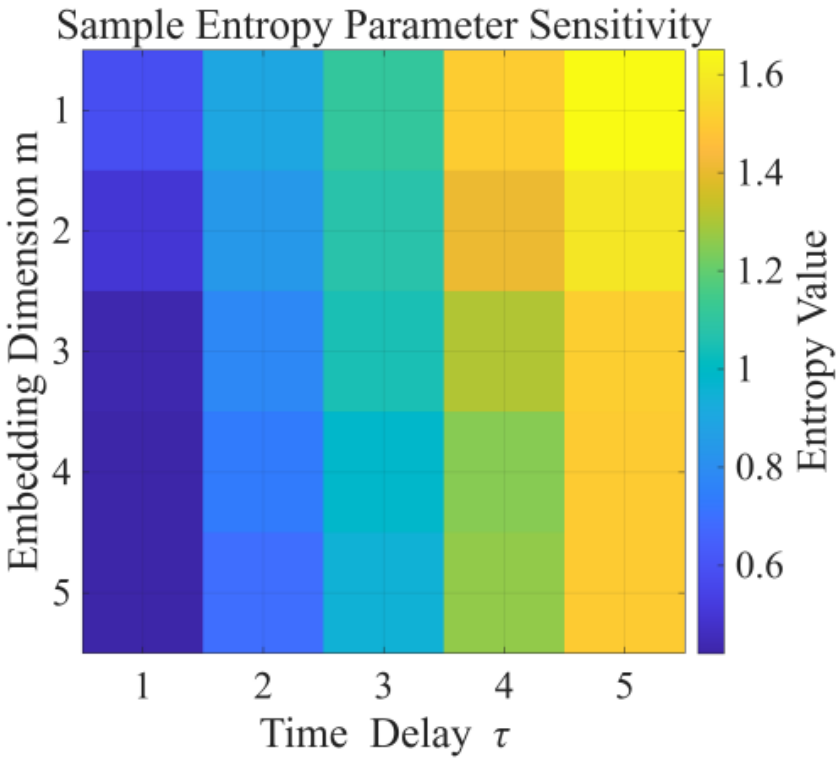

Figure 8. SE parameter sensitivity heatmap

As shown in Figure 8, the SE value increases significantly with increasing time delay and decreases with increasing embedding dimension, with the time delay exerting a more dominant effect. The parameter sensitivity heatmap of FE is presented in Figure 9.

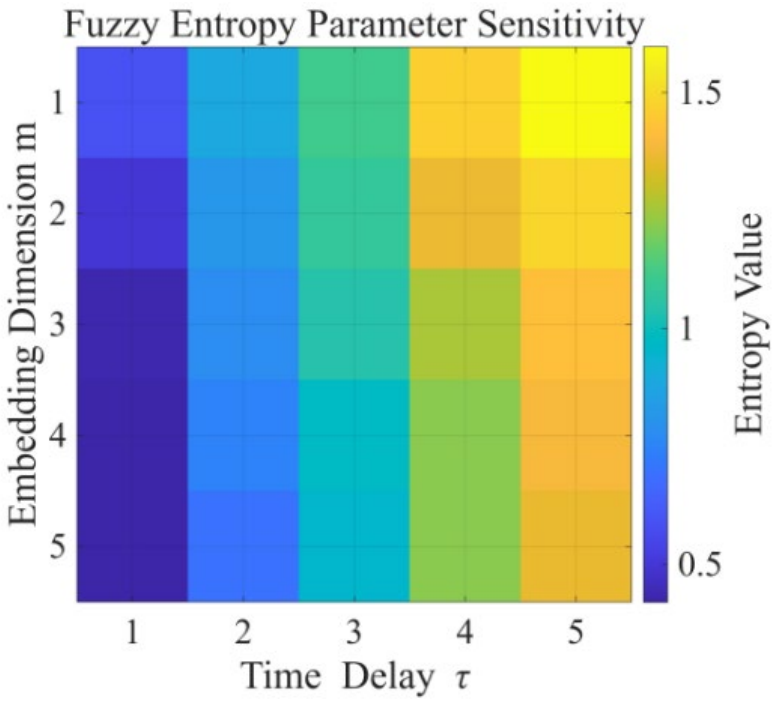

Figure 9. FE parameter sensitivity heatmap

As shown in Figure 9, both SE and FE exhibit excessive sensitivity to their parameters and lack sufficient stability. The parameter sensitivity heatmap of DE is presented in Figure 10.

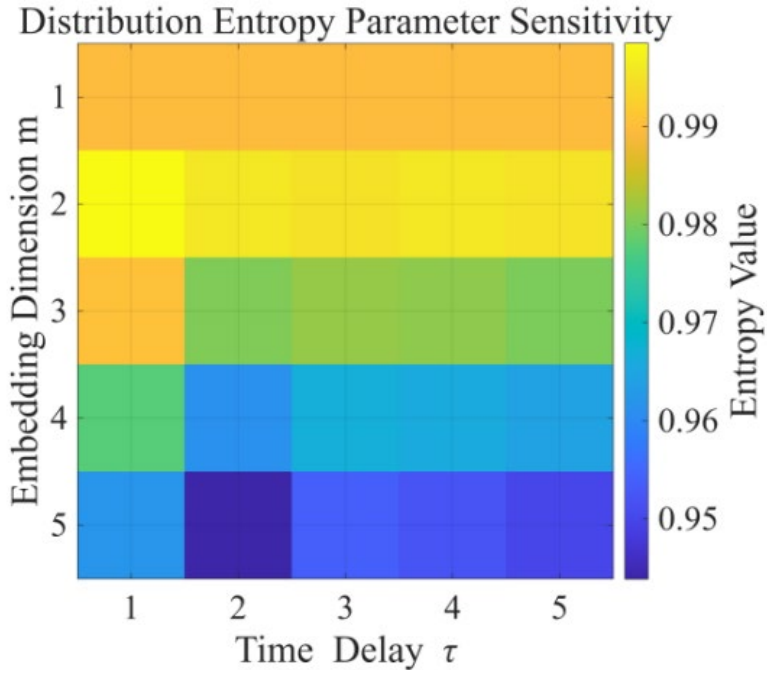

Figure 10. DE parameter sensitivity heatmap

As shown in Figure 10, DE exhibits low sensitivity to time delay, with its entropy value mainly affected by the embedding dimension, showing good overall parameter stability. The parameter sensitivity heatmap of PE is presented in Figure 11.

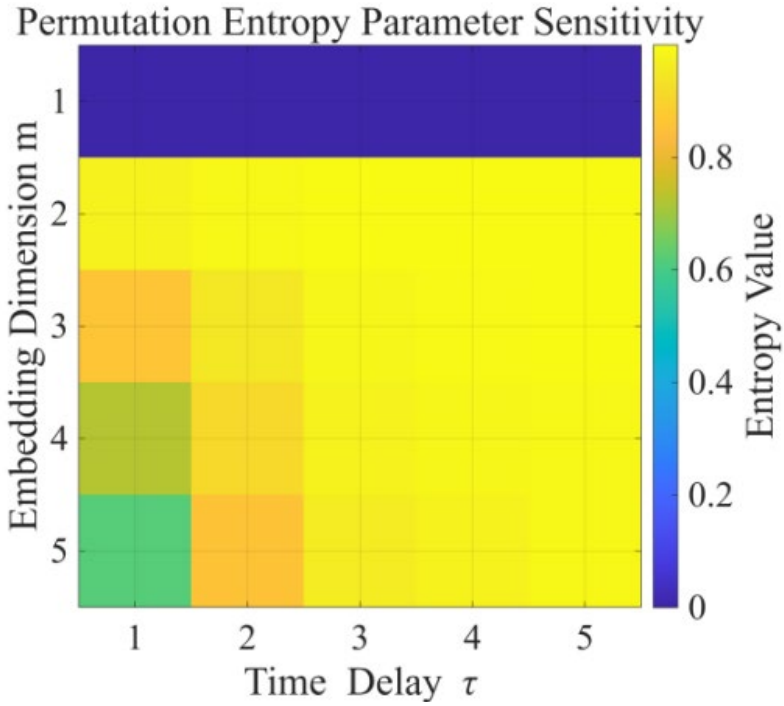

Figure 11. PE parameter sensitivity heatmap

As shown in Figure 11, PE is insensitive to time delay when the embedding dimension is greater than or equal to 2, showing good overall parameter stability, yet it suffers from low-dimensional failure at small embedding dimensions. The parameter sensitivity heatmap of AWM-FE is presented in Figure 12.

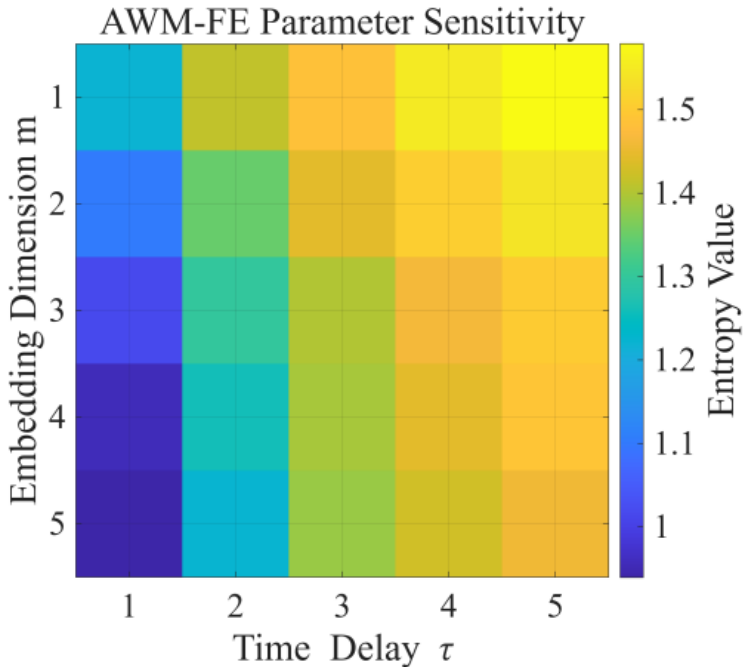

Figure 12. AWM-FE parameter sensitivity heatmap

As shown in Figure 12, AWM-FE exhibits a balanced and smooth response to m and τ; it neither completely loses its discriminatory power, as seen with DE, nor suffers from reduced stability due to excessively steep gradients, as observed with SE or FE. The results quantifying the parameter sensitivity of the entropy methods by calculating the global range of variation for

each entropy value are shown in Figure 13.

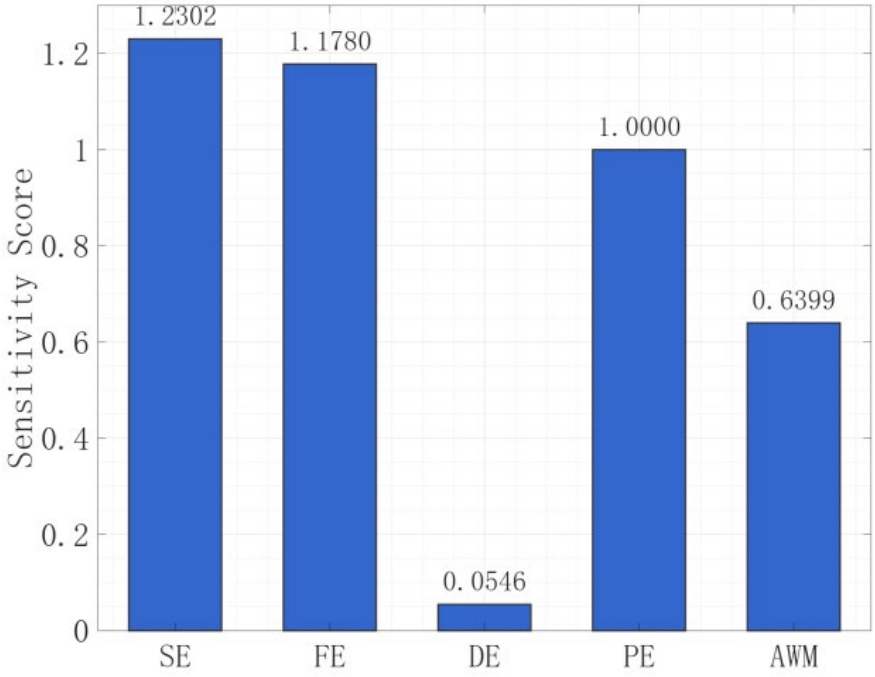

Figure 13. Comparison of parameter sensitivity across entropy measures

As can be seen from Figure 13, DE and AWM-FE exhibit low parameter sensitivity, whereas PE, SE and FE exhibit high parameter sensitivity.

4.2 Analysis of anomalous features in fixed-point deformation data

The raw data for the experiment on the classification of anomalous signals using various entropy measures was labelled by maintenance staff at the Yixian Seismological Observatory of the Hebei Provincial Seismological Bureau, based on the actual operational status of the VP inclinometer. The experiment was conducted using MATLAB 2025a on a Windows 10 system. After pre-processing the raw data, time-domain plots were generated, analysed, and signal segments were selected. After standardizing segment lengths, resampling techniques were applied to expand these equal-length segments. The final experimental dataset comprised 49 power failure instances, 50 lightning anomaly instances, 48 barometric rainfall interference instances, and 56 collapse/earthquake instances, totaling 203 samples. The dataset underwent category calibration and feature extraction using multiple entropy measures. Power failure signals were categorized as Class 1, lightning anomaly signals as Class 2, barometric rainfall interference signals as Class 3, and collapse signals as Class 4. A statistical plot of the entropy values for the four signal categories, as shown in Figure 14.

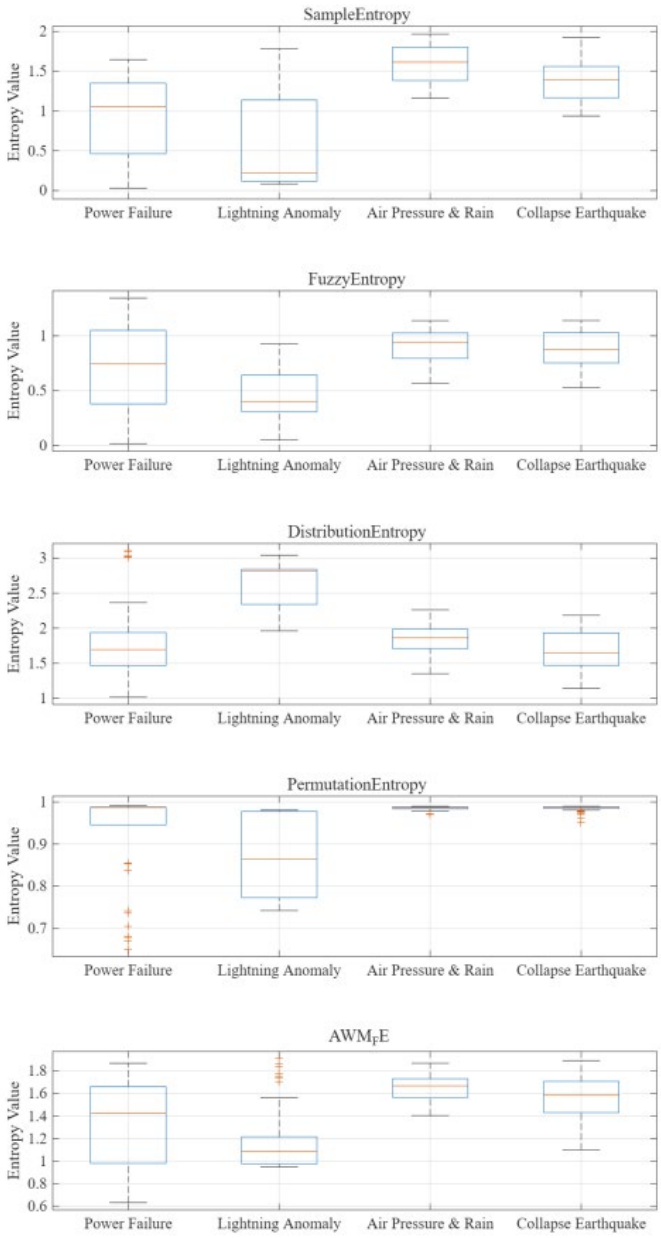

Figure 14. Statistical chart of entropy values for four types of signals

As shown in Figure 14, the entropy features of the resampled feature segments exhibit differences. Specifically, there are significant differences in sample entropy, distribution entropy, and the median of AWM_FE. The box overlap is relatively low, suggesting good discriminatory power; however, none of these measures satisfy the assumption of homogeneity of variances, necessitating the use of the Welch t-test.

We performed a test for homogeneity of variances on the set of multi-entropy features we constructed. For samples with heterogeneous variances, we conducted statistical analysis using the Welch t-test. For the four types of anomalous signals under each entropy feature, we performed non-repeated pairwise comparisons and conducted t-tests for each pair; the test results are shown in Figure 15.

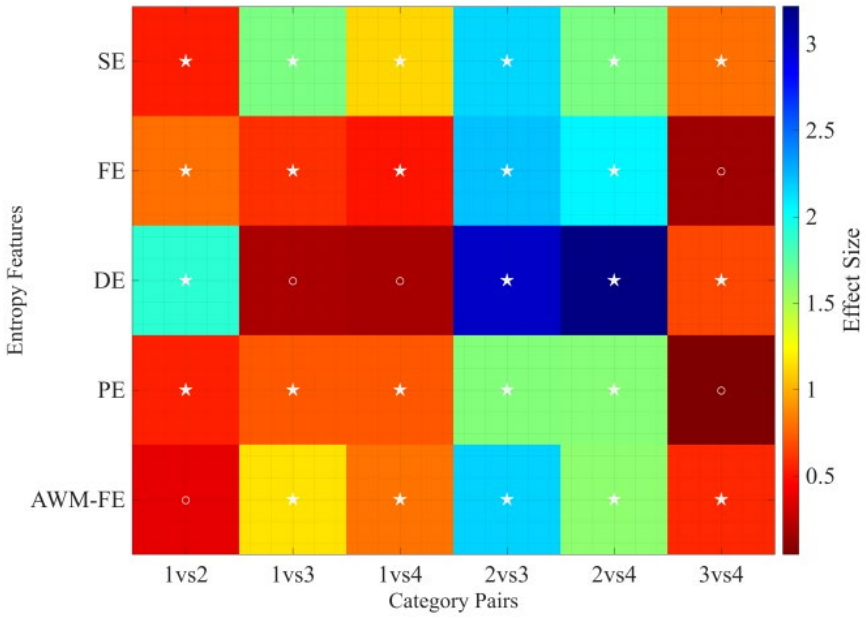

Figure 15. Significance heatmap of entropy differences

As shown in Figure 15, the SE and AWM-FE exhibit the most pronounced color gradient changes when comparing different signal groups, demonstrating superior color discrimination capabilities. Statistical analysis of the p-values provides an initial quantitative assessment of the discrimination performance across various entropy features, with results presented in Figure 16.

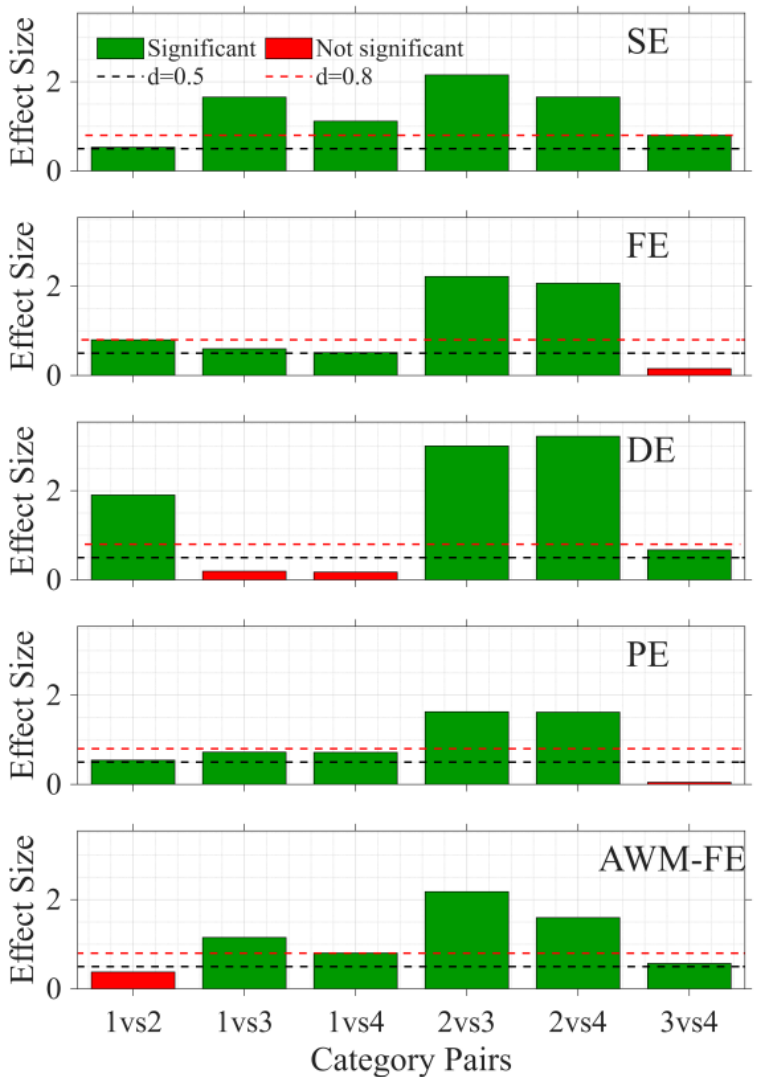

Figure 16. Quantitative diagram of distinguishing performance for various entropy features

As shown in Figure 16, SE exhibits significant discrimination among all six comparison groups, followed by FE, PE and AWM-FE. To mitigate the risk of cumulative false positives from multiple t-tests, this study further employed Kruskal-Wallis to examine overall differences across each entropy feature. The results confirmed significant differences in all entropy features between different signal types. Post-hoc multiple comparison tests were then conducted to control the family-wise error rate and determine significant differences between each pair of categories. The corresponding significance heat-map is shown in Figure 17.

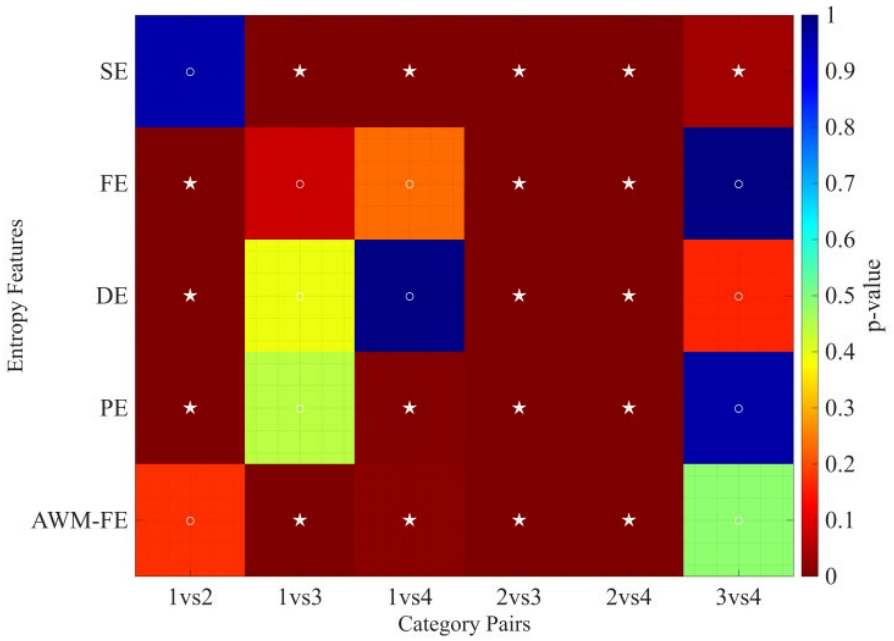

Figure 17. Multi-entropy significance heatmap

The five-pointed star symbols within the heatmap cells indicate that the entropy feature exhibits statistically significant discriminatory power for the two signal types labeled in the column header. As shown in Figure 17, SE demonstrate the strongest discriminatory capabilities, achieving significant differentiation for 5 groups; AWM_FE and PE exhibits relatively good discriminatory power, achieving significant differentiation for 4 groups; while DE and FE show weaker discriminatory capabilities. The differences in mean entropy values across different signals are shown in Figure 18.

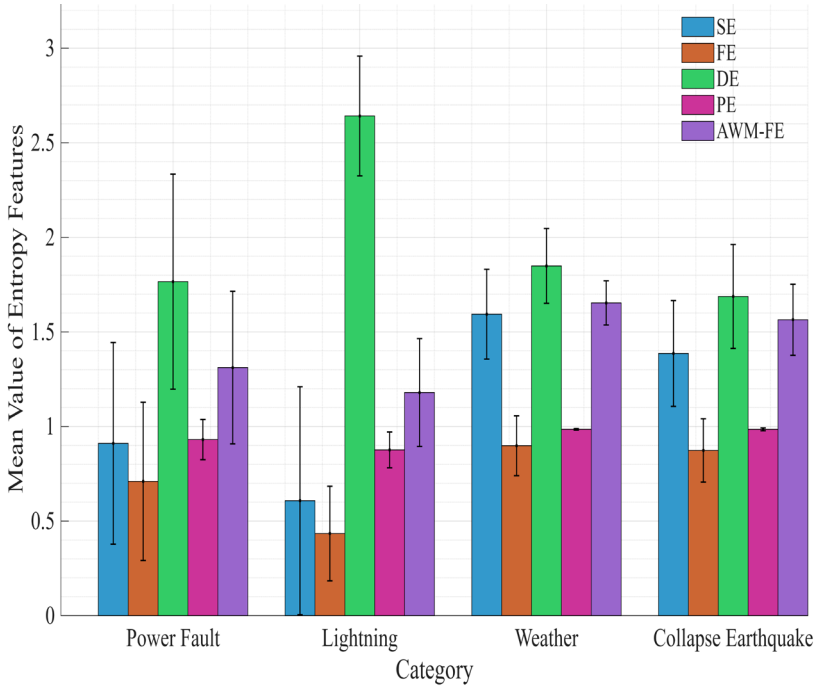

Figure 18. Differences in mean values of multiple entropies across different signal categories

As shown in Figure 18, the intra-category standard deviations of PE are extremely small,

exhibiting high consistency among similar anomaly signals and minimal influence from individual signal fluctuations.

The intra-category standard deviations of AWM_FE and DE mostly range between 0.1 and 0.3, ensuring consistency among similar samples while allowing for some individual variation. SE and FE exhibit larger standard deviations in Category 1, indicating weaker stability.

SE exhibits the largest mean span across categories, with pronounced mean differences in all category pairs. AWM_FE and FE show good mean differences between categories, though some pairs exhibit minor variations and certain groups show insignificant differences.

DE and PE exhibit extremely small mean differences between categories, yielding insignificant discrimination results; all category mean differences for PE are less than 0.03, making it nearly impossible to distinguish between different signal types.

Additionally, the overall mean, standard deviation, median, maximum, and minimum values for each entropy category are summarized in the Table 2 below.

Table 2. Core statistics table for entropy features

Feature	SE	FE	DE	PE	AWM_FE
Mean	1.1287	0.7317	1.9795	0.9452	1.4296
Std	0.5817	0.3231	0.5274	0.0833	0.3271
Mdn	1.2556	0.8040	1.9081	0.9852	1.5205
Min	0.0314	0.0129	1.0158	0.6505	0.6314
Max	1.9671	1.3449	3.1042	0.9907	1.9090

As shown in Table 2, the overall standard deviation of PE is 0.0833, exhibiting good stability but the narrowest numerical range. AWM-FE and FE demonstrate relatively good stability but average discrimination performance. SE shows weaker stability yet a symmetrical distribution, delivering optimal discrimination performance. DE exhibits the widest range but contains extreme values. Figure 19 shows the heatmap obtained from the correlation analysis of the five entropies.

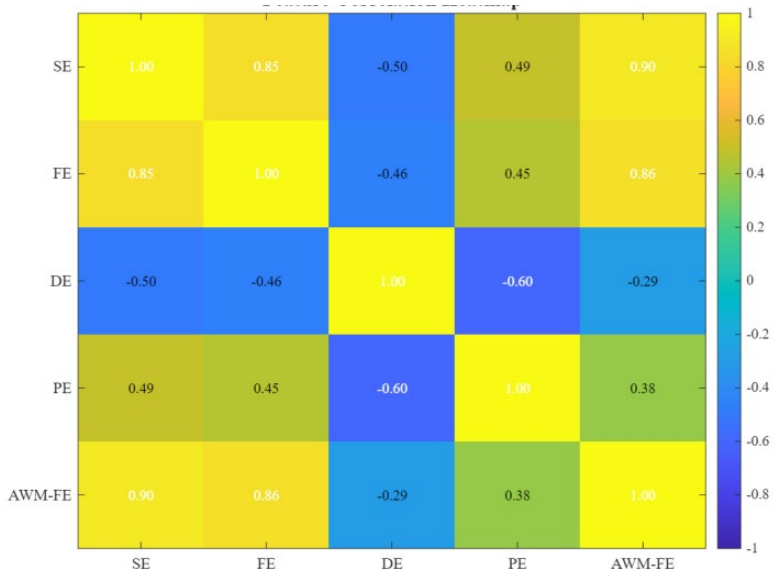

Figure 19. Heatmap of correlations between entropies

As shown in Figure 19, the SE, FE, and AWM-FE features exhibit strong correlations and significant dimensional overlap when describing signal complexity, which may indicate information redundancy. In contrast, DE and PE show weaker correlations with other entropy features and demonstrate notable complementary information, enabling them to capture the intrinsic characteristics of signals from different dimensions. In subsequent multi-feature fusion modeling, DE and PE can be prioritized as core complementary features to be fused with other entropy features, thereby reducing feature redundancy and enhancing the model's ability to distinguish among various types of abnormal signals.

In addition, to avoid the bias introduced by subjective weighting in AWM-FE, we conducted a stability analysis of the AWM-FE weight w_s and verified the validity of the fusion of SE, FE, and DE at different scales. Using perturbation experiments, we added Gaussian noise with standard deviations of 0.05, 0.10, and 0.15 to the central weights, recalculated the features, and evaluated performance using an SVM classifier. The results show that performance metrics remained stable when the perturbation step size Δw is set to 0.05; when Δw is set to 0.10 and 0.15, the relative fluctuations in all metrics were less than 5%, and the mean classification performance remained essentially stable, indicating that AWM-FE is insensitive to weight perturbations.

A comparison was made across four scale ranges: 1:5, 1:10, 1:15, and 1:20. The results showed that the p-value was smallest and the discriminatory ability was strongest for the 1:5 scale range; as the scale range increased, the p-value gradually rose, and the discriminatory ability showed a downward trend. Smaller scale ranges are better at capturing local details, whereas larger scales are prone to introducing noise or smoothing out anomalous features.

In the comparison between adaptive weighting and equal weighting, the p-value of the adaptive weighting model was 2.73×10^{-21}, significantly outperforming the equal-weighting model. The energy-adaptive weight allocation strategy is far superior to subjectively set equal weights: it better aligns with the physical laws governing the multi-scale energy distribution of signals, offers good interpretability, and effectively enhances feature discrimination capability, while avoiding the arbitrariness and limitations of subjective weighting.

4. CONCLUSION

This paper focuses on the feature extraction problem for anomaly signal identification in fixed-point deformation monitoring. Through quantitative evaluation of simulated signals and statistical analysis of actual measurement data, the performance of five entropy features was compared across multiple dimensions. Based on a comprehensive assessment of discrimination capability and stability, SE and AWM_FE demonstrated superior overall performance among the five entropy features. Experimental results from simulated signals reveal that sample entropy not only exhibits strong discrimination capability but also benefits from low parameter sensitivity. However, it demonstrates poor noise resistance and stability. AWM-FE shows moderate discrimination capability but good stability. Future algorithmic parameter optimization could enhance its discrimination ability, offering promising potential. FE and distribution entropy perform at moderate levels across all dimensions.

In the point-deformation anomaly signal experiment, SE achieved the best discrimination performance with an average p-value of only 0.0016 in the t-test, effectively distinguishing all category pairs. AWM-FE and FE could distinguish 5 groups, while permutation entropy and distribution entropy showed weaker discrimination performance with multiple category pairs being non-significant. In the analysis of variance, DE yielded the highest F-value and greatest intergroup variation, while PE produced the lowest F-value. In multiple comparisons, SE showed non-significant classification for only one group, while the remaining features exhibited non-significant results for 2–3 category pairs. Overall data analysis indicates that SE exhibits the smallest overall standard deviation and strongest discrimination capability. AWM-

FE and FE perform well, while distribution entropy shows high inter-group variation but limited discrimination. PE demonstrates the weakest discrimination capability but the most stable distribution.

Abbreviations

SE, Sample Entropy;
FE, Fuzzy Entropy;
DE, Distribution Entropy;
PE, Permutation Entropy;
AWM-FE, Adaptive Weighted Multi-scale Fusion Entropy;
Std, Standard Deviation;
Mdn, Median;
Min, Minimum;
Max, Maximum.

Supplementary Material

Not applicable.

Appendix

Not applicable.

Ethics approval and consent to participate.

This study did not involve human participants, animal subjects, or any data requiring ethical approval. Therefore, ethics approval and consent to participate are not applicable.

Acknowledgements

The authors would like to thank the editors of this journal and all the anonymous reviewers who provided valuable comments on this work.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Author contributions

All authors have read and agreed to the published version of the manuscript. The author's contributions are specified as follows: **C.P.:** Conceptualization, Writing – Review & Editing, Methodology. **S.L.:** Validation, Writing – Original draft, Visualization. **X.W.:** Formal analysis, Investigation. **M.S.:** Supervision, Project administration.

Funding information

This work was supported by the Open Fund of Wuhan Gravitation and Solid Earth Tides, National Observation and Research Station, No. WHYWZ202406; Earthquake Monitoring and Early Warning Task No.CEA-JCYJ-202601025; Research grants from National Institute of Natural Hazards, Ministry of Emergency Management of China (Grant Number: ZDJ2024-31, ZDJ2025-58).

Data availability

The data that support the findings of this study are available upon request from the corresponding authors, **C.P.**

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

Declaration of AI and AI-assisted Technologies in the Writing Process

During the writing of this article, the author used DeepSeek for spelling and grammar checking. After using this tool, the author reviewed and edited the content as needed and assumes full responsibility for the final published content.

REFERENCES

[1] Peng, Z., & Shao, Y. (2024, November). Deep Learning Seismic Detection on Fixed-Point Deformation Data: A Case in Tianjin, China. In 2024 5th International Conference on Artificial Intelligence and Computer Engineering (ICAICE) (pp. 598-604). IEEE. DOI: **https://doi.org/10.1109/icaice63571.2024.10864303**

[2] Wu, L., Huang, Y., & Li, D. (2021). Tilt active vibration isolation using vertical pendulum and piezoelectric transducer with parallel controller. *Applied Sciences*, *11*(10), 4526. DOI: **https://doi.org/10.3390/app11104526**

[3] Zhao, Y. (2020). The background noise level of vertical pendulum tiltmeter in China. *Geodesy and Geodynamics*, *11*(2), 143-150. DOI: **https://doi.org/10.1016/j.geog.2020.01.002**

[4] Klügel, T., & Wziontek, H. (2009). Correcting gravimeters and tiltmeters for atmospheric mass attraction using operational weather models. *Journal of Geodynamics*, *48*(3-5), 204-210. DOI: **https://doi.org/10.1016/j.jog.2009.09.010**

[5] Bailun, C., Yue, Y., Fujiang, K., Zhengxian, Z., Yong, D., & Wang, L. (2023). Research on tidal disturbance phenomenon of DSQ water tube tiltmeter at Yunnan Yanziyan observation station. *Progress in Earthquake Sciences*, *54*(4), 265-271. DOI: **https://doi.org/10.19987/j.dzkxjz.2023-093**

[6] Ordoñez, M., Idárraga, J., Adamo, R., & Battaglia, M. (2024). Geodetic monitoring of the recent activity and the dome forming eruption at Nevado del Ruiz (Colombia), 2010–2023. *Scientific Reports*, *14*(1), 21441. DOI: **https://doi.org/10.1038/s41598-024-72058-y**

[7] Yang, F., Guan, D., Li, X., & Dou, C. (2025). Application Research on High-Precision

Tiltmeter with Rapid Deployment Capability. *Sensors*, *25*(5), 1559. DOI: **https://doi.org/10.3390/s25051559**

[8] Kim, S. Y., & Mukhiddinov, M. (2023). Data anomaly detection for structural health monitoring based on a convolutional neural network. *Sensors*, *23*(20), 8525. DOI: **https://doi.org/10.3390/s23208525**

[9] Choe, H. O., & Lee, M. H. (2025). Artificial Intelligence-Based Anomaly Detection Technology for Equipment Condition Monitoring in Smart Farms. *Applied Sciences*, *15*(23), 12843. DOI: **https://doi.org/10.3390/app152312843**

[10] Jung, Y., Park, E. G., Jeong, S. H., & Kim, J. H. (2024). AI-based anomaly detection techniques for structural fault diagnosis using low-sampling-rate vibration data. *Aerospace*, *11*(7), 509. DOI: **https://doi.org/10.3390/aerospace11070509**

[11] Nong, X., Luo, X., Lin, S., Ruan, Y., & Ye, X. (2023). Multimodal deep neural network-based sensor data anomaly diagnosis method for structural health monitoring. *Buildings*, *13*(8), 1976. DOI: **https://doi.org/10.3390/buildings13081976**

[12] Pang, C., Li, C., Ma, L., Jiang, Y., Chen, J., Ma, W., & Xiang, Y. (2023, February). Automatic Identification of Seismic Signal Based on Entropy Feature and Genetic Algorithm Optimized BP Neural Network. In *2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (Vol. 6, pp. 778-782). IEEE. DOI: **https://doi.org/10.1109/itnec56291.2023.10082096**

[13] Li, Y., Wang, S., Yang, Y., & Deng, Z. (2022). Multiscale symbolic fuzzy entropy: An entropy denoising method for weak feature extraction of rotating machinery. *Mechanical Systems and Signal Processing*, *162*, 108052. DOI: **https://doi.org/10.1016/j.ymssp.2021.108052**

[14] Huo, Z., Martínez-García, M., Zhang, Y., Yan, R., & Shu, L. (2020). Entropy measures in machine fault diagnosis: Insights and applications. *IEEE Transactions on Instrumentation and Measurement*, *69*(6), 2607-2620. DOI: **https://doi.org/10.1109/tim.2020.2981220**

[15] Zhuang, D., Liu, H., Zheng, H., Xu, L., Gu, Z., Cheng, G., & Qiu, J. (2023). The IBA-ISMO method for rolling bearing fault diagnosis based on VMD-sample entropy. *Sensors*, *23*(2), 991. DOI: **https://doi.org/10.3390/s23020991**

[16] Shaikh, M. M., Butt, R. A., Khawaja, A., & Jarboui, S. (2025). Optimized TEC prediction with the CEEMDAN-SE-LSTM framework: integrating Sample Entropy for reducing processing time. *Advances in Space Research*. DOI: **https://doi.org/10.1016/j.asr.2025.09.051**

[17] Liang, F., Hong, J., & Zhang, X. (2025). Fault diagnosis of real-scenario battery systems in electric vehicles using the multi-scale fuzzy entropy algorithm. *Engineering Applications of Artificial Intelligence*, *151*, 110670. DOI: **https://doi.org/10.1016/j.engappai.2025.110670**

[18] Gao, S., Wang, Q., & Zhang, Y. (2021). Rolling bearing fault diagnosis based on CEEMDAN and refined composite multiscale fuzzy entropy. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-8. DOI: **https://doi.org/10.1109/tim.2021.3072138**

[19] Liang, F., Hong, J., & Zhang, X. (2025). Fault diagnosis of real-scenario battery systems in electric vehicles using the multi-scale fuzzy entropy algorithm. *Engineering Applications of Artificial Intelligence*, *151*, 110670. DOI: **https://doi.org/10.1016/j.engappai.2025.110670**

[20] Wang, A., Pang, C., Chen, G., Li, C., & Zhao, T. (2025). Optimization of the SOM neural network model using CEEMDAN distribution entropy and ALO for seismic and blasting

identification. *J Seismic Explor*, *34*(2), 28-43. DOI: **https://doi.org/10.36922/jse025280033**

[21] Ma, Y., Cheng, J., Wang, P., Wang, J., & Yang, Y. (2023). A new rotating machinery fault diagnosis method for different speeds based on improved multivariate multiscale fuzzy distribution entropy. *Nonlinear Dynamics*, *111*(18), 16895-16919. DOI: **https://doi.org/10.1007/s11071-023-08609-1**

[22] Aung, S. T., & Wongsawat, Y. (2021). Prediction of epileptic seizures based on multivariate multiscale modified-distribution entropy. *PeerJ Computer Science*, *7*, e744. DOI: **https://doi.org/10.7717/peerj-cs.744**

[23] Myers, A., & Khasawneh, F. A. (2020). On the automatic parameter selection for permutation entropy. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(3). DOI: **https://doi.org/10.1063/1.5111719**

[24] Li, X., Chen, Z., Xu, Z., Long, Y., & Wang, Y. (2025). Research on gearbox fault diagnosis based on the CHO-VMD-BiGRU algorithm. *Engineering Research Express*, *7*(3), 035528. DOI: **https://doi.org/10.1088/2631-8695/adf0c2**

[25] Jiao, S., Zhu, Y., Zhang, Q., Wang, Y., Li, Y., Li, C., & Wang, X. (2025). CEEMDAN permutation entropy based statistical complexity measure: A new stochastic resonance metric for enhanced detection of feature-unknown weak signals. *Chinese Journal of Physics*. DOI: **https://doi.org/10.1016/j.cjph.2025.08.036**

[26] Sumith, K. V., & S, B. (2026). Comparative Analysis of Deep Learning and Machine Learning Models for Evapotranspiration Prediction in Semi-Arid Regions: Statistical Model Evaluation with a Paired t-Test and Bootstrap Resampling. *Water Resources Management*, *40*(4), 158. DOI: **https://doi.org/10.1007/s11269-025-04411-3**

[27] Abdulmohsin, H. A., Wahab, H. B. A., & Hossen, A. M. J. A. (2021). A new hybrid feature selection method using T-test and fitness function. *Computers, Materials, & Continua*, *68*(3), 3997. DOI: **https://doi.org/10.32604/cmc.2021.014840**

[28] Harish, S., & Sathyakam, P. U. (2025). Dunn's method for distinguishing charge storage mechanisms in supercapacitors: a status quo review. *Journal of Electronic Materials*, 1-15. DOI: **https://doi.org/10.1007/s11664-025-12481-7**

[29] Hasan, R., Altaf, Y., Jabeen, N., Hassan, N. U., Ahmed, F., Hussain, S., ... & Al-Asbahi, B. A. (2024). ZnS@ Fe2O3 core–shell nanorod arrays for supercapattery applications; theoretical evaluation of faradic and non-faradic behavior using Dunn's model. *Journal of Electroanalytical Chemistry*, *966*, 118411. DOI: **https://doi.org/10.1016/j.jelechem.2024.118411**

[30] Afzal, A. M., Muzaffar, N., Iqbal, M. W., Dastgeer, G., Manzoor, A., Razaq, M., ... & Eldin, S. M. (2024). Exploring the charge storage mechanism in high-performance Co@ MnO2-based hybrid supercapacitors using Randles–Ševčík and Dunn's models. *Journal of Applied Electrochemistry*, *54*(1), 65-76. DOI: **https://doi.org/10.1007/s10800-023-01939-3**