ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

A multi-class anomaly signal classification method for fixed-point deformation combining PSO–VMD multiscale adaptive sample entropy and ALO–SOM model

Cong Pang[1,2], Tianwen Zhao[3], Guoqing Chen[4,5], Wugang Ma[1], Chawei Li[1], Yong Jiang[1], and Piyapatr Busababodhin[5]*

[1]Key Laboratory of Earthquake Geodesy, 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]Department of Trade and Logistics, Daegu Catholic University, Gyeongsan, Daegu, Republic of Korea

[4]Mathematical Modeling Research Center, Chengdu Jincheng College, Chengdu, Sichuan, China

[5]Department of Mathematics, Faculty of Science, Mahasarakham University, Kantharawichai, Maha Sarakham, Thailand

***Corresponding author:**
Piyapatr Busababodhin
(piyapatr.b@msu.ac.th)

Citation: Pang C, Zhao T, Chen G, *et al*. A multi-class anomaly signal classification method for fixed-point deformation combining PSO–VMD multiscale adaptive sample entropy and ALO–SOM model. *J Seismic Explor.*
doi: 10.36922/JSE026290141

Received: July 16, 2026

Revised: September 30, 2026

Accepted: October 1, 2026

Published online: November 4, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract

To address the challenge of identifying five predefined categories of disturbance-related abnormal signals (lightning interference, artificial blasting, power-supply failure, collapse-related events, and meteorological anomalies) in fixed-point deformation monitoring, this paper proposes an identification framework integrating parameter-optimized variational mode decomposition (VMD), multiscale entropy analysis, and an intelligent classification model. Firstly, particle swarm optimization is adopted to adaptively determine the mode number and penalty factor for VMD to realize adaptive signal decomposition. Then, an adaptive sample entropy algorithm with self-adjustable similarity thresholds is applied to extract multiscale complexity-oriented features from each decomposed modal component. Finally, the ant lion optimizer optimizes the competitive-layer topology and training hyperparameters of a self-organizing map network to build the target classifier. Comparative experiments against 18 widely-used classification models (including AdaBoost ensemble learners, generalized regression neural networks, long short-term memory networks, and gated recurrent units) are carried out on the real-world observation dataset from the Yixian seismic station. The experimental results show that the proposed framework achieves competitive classification performance for the five anomaly categories. The mean test-set accuracy over 100 independent repeated-run experiments reaches 99.62%, with a standard deviation of 0.6361. This study offers an effective technical reference for identifying deformation-related abnormal waveforms under complex observational conditions.

Keywords: Fixed-point deformation multi-class anomaly classification; Self-organizing feature mapping; Ant lion optimization; Variational mode decomposition; Adaptive sample entropy; Particle swarm optimization

1. Introduction

Crustal deformation is one of the most direct physical phenomena in the process of earthquake preparation and occurrence. Since Reid put forward the "elastic rebound theory" in the early 20th century, the precise observation of crustal deformation has become the core way to explore the earthquake precursor and seismogenic mechanism.[1-4] Fixed-point deformation observation, as a technology that can continuously and accurately record weak signals such as crustal tilt and strain, has built a national monitoring network consisting of about 275 stations and nearly 600 sets of instruments in China after more than half a century of development. The high sampling rate (minute-value, second-value) time series data produced by the network not only contain rich information such as solid tide and coseismic response, but also record transient or slow-moving abnormal signals related to earthquake preparation, which play an irreplaceable role in earthquake prediction research and geophysical exploration.[5] In this context, the diversity of abnormal signals recorded by fixed-point deformation instruments has attracted increasing attention, and entropy-based feature tests have been used to characterize the complexity differences among abnormal signals in fixed-point deformation monitoring.[6]

However, with the massive growth of observational data, the disadvantages of the traditional "manual sorting" anomaly classification mode relying on expert experience are increasingly prominent. At present, the identification of coseismic response and other events in fixed-point deformation data still mainly depends on manual visual judgment. This method is inefficient and has poor consistency, and is seriously restricted by the subjective experience of analysts, which makes it difficult to meet the needs of real-time monitoring and big data analysis.[7] Although previous studies have attempted to use wavelet analysis, information entropy, and other methods for automatic detection, most of these methods can only achieve "abnormal trigger" and cannot accurately identify and classify the physical types of abnormal events (such as lightning interference, artificial blasting, collapse earthquake). Related work on interference and precursor signal analysis has shown that machine-learning models such as random forest and improved logistic regression can be used to separate interference from precursory signals, which provides a useful reference for the present multi-class classification task.[8] In actual observations, signals are often submerged by various kinds of noise (such as instrument noise and environmental interference), and different types of abnormal signals (such as transient lightning interference and slowly varying tectonic deformation) are intertwined in time-frequency characteristics, forming a complex sequence with low signal-to-noise ratio and multi-source mixing.[9-14] Therefore, the development of a technology that can automatically, accurately, and reliably identify and classify multiple types of abnormal events in fixed-point deformation data has become an urgent need to improve the ability of earthquake monitoring and prediction and deepen the understanding of geophysical phenomena.

Early fixed-point deformation anomaly analysis focuses on extracting precursory information from solid tidal changes, such as detecting changes in the state of crustal media through loading and unloading response ratio and other methods.[15-17] For a wider range of abnormal signal detection, the evolution of methodology can be roughly divided into two stages: (i) Traditional signal processing stage: mainly using methods based on thresholds, waveform correlation (such as the zero-delay multiplication algorithm), or time-frequency analysis (such as wavelet transform). These methods are simple to calculate and have a certain effect in suppressing noise and highlighting specific signals. However, their performance depends heavily on preset parameters and prior knowledge, and their generalization ability is weak. (ii) Exploration stage of intelligent classification: In recent years, machine learning and deep learning technologies have been introduced.[18,19] For example, deep-learning-based fault diagnosis and prediction models have demonstrated strong feature-learning capability in complex signal classification tasks, and feature extraction combined with machine learning classification algorithms has also been widely used in classification.[20,21] For instance, some studies have used transfer learning to apply a deep learning model trained on seismic data to the detection of fixed-point deformation coseismic response, and achieved better efficiency and consistency than manual methods. This indicates that the field is moving towards automation and intelligence. However, most existing research focuses on detecting a single event type, and no mature solution exists for the multi-category, diverse characteristics of abnormal mixtures composed of lightning, blasting, power failure, collapse, earthquake, and natural anomalies.[22-24] This requires that the model should not only have strong feature extraction ability, but also have fine classification and discrimination ability.

Feature extraction is the cornerstone of anomaly classification. For non-stationary and nonlinear fixed-point deformation signals, the effect of direct analysis on the original data is limited. As an adaptive signal decomposition method, variational mode decomposition (VMD) can effectively decompose complex signals into a series of intrinsic mode functions with practical physical

significance, which is highly suitable for subsequent feature extraction. However, the performance of VMD is strongly dependent on its two key parameters: modal number k and penalty factor α. Inappropriate parameters can lead to modal aliasing or over-decomposition.[25-29] Traditionally, parameters are set by experience and lack universality. For this reason, scholars have introduced intelligent optimization algorithms for parameter self-adaptive optimization, such as using the particle swarm optimization (PSO) algorithm to optimize the superparameters of backpropagation (BP) neural networks or XGBoost models, and its advantages have been proved in the fields of tunnel deformation prediction and investigation of the deformation mechanism.[30,31] This provides a direct methodological reference for the optimization of VMD using the PSO algorithm in this paper.

In terms of feature quantification, sample entropy is a classic index to measure the complexity and irregularity of time series. However, its analysis effect is subject to fixed similarity tolerance parameters.[32,33] The improved PSO algorithm has also been successfully applied to complex optimization problems, further confirming its effectiveness in parameter adaptation.[34] Developing a kind of sample entropy that can adapt to the characteristics of samples, and automatically constructing parameters such as similarity tolerance, can more robustly describe the dynamic characteristics of signals at different scales. By combining it with the optimized VMD, that is, calculating the adaptive sample entropy for each intrinsic mode function (IMF) component obtained by optimal decomposition, the feature vector that can comprehensively represent the multi-scale complexity information of the original signal can be constructed, which lays the foundation for the subsequent accurate classification.

In terms of pattern classification, from the traditional support vector machine (SVM) and random forest to the deep learning models emerging in recent years, they have been widely used in various fields of fault diagnosis and anomaly classification. In the literature, stacked denoising autoencoders and optimized stacked denoising autoencoder depth models are often used to deal with complex features. However, for structured feature input, the self-organizing map (SOM) network has unique advantages in classification because of its powerful unsupervised clustering and visual dimensionality reduction ability. SOM maps high-dimensional features to low-dimensional topological grids, and similar samples are adjacent on the graph, intuitively showing category relationships. However, its performance is significantly affected by the dimensions of the competitive layer, learning rate, training time, and other parameters.[35,36]

At the same time, research shows that a single-depth model or traditional model is often difficult to achieve the best balance between accuracy and efficiency. The integrated or optimized hybrid model shows greater potential. For example, the grey wolf optimization algorithm is used to optimize the parameters of the SVM to improve the classification rate, or the hybrid intelligent algorithm is used to optimize the network hyperparameters. As a new bionic optimization algorithm, the ant lion optimizer (ALO) algorithm has excellent performance in optimizing the parameters of complex problems because of its strong global search and convergence ability. It is expected to automatically find the best configuration of the SOM network by using it to optimize the competitive dimension and training time of the SOM, so as to improve the stability and accuracy of the overall classification performance.[37-39] In addition, Wang *et al.*[40] proposed an ALO–SOM identification scheme based on complete ensemble empirical mode decomposition (EMD) with adaptive noise and distribution entropy for seismic-blasting discrimination. Compared with their work, this paper adopts PSO-optimized VMD for signal decomposition instead of complete ensemble EMD with adaptive noise, constructs adaptive sample-entropy multiscale features, and builds an end-to-end collaborative optimization pipeline targeting five mixed anomaly categories in fixed-point deformation monitoring, which covers more disturbance types in station observations.

To sum up, the current research has the following gaps and deficiencies:

(i) Lack of pertinence: Most advanced feature extraction and classification methods (such as optimized VMD, adaptive entropy, intelligent optimized SOM) focus on mechanical fault diagnosis and other fields, and have not been systematically introduced and integrated into the specific and important geophysical problem of multi-class anomaly classification of fixed-point deformation.

(ii) The technology chain is fragmented: The existing research focuses on feature extraction or classification algorithm optimization, lacking a complete end-to-end collaborative optimization technology framework from signal adaptive decomposition → multi-scale robust feature extraction → classifier intelligent optimization.

(iii) Insufficient comparison and verification: After the new method is proposed, it is often compared with a few baseline models and lacks systematic performance evaluation with a wide range of model systems, including classical machine learning, mainstream deep learning, and integration methods.

In view of the above research background and current situation, this paper aims to build an automatic and high-precision classification framework for multi-class abnormal signals of fixed-point deformation. The core research content is to propose a "fixed-point deformation anomaly classification method based on PSO–VMD multi-scale adaptive sample entropy and ALO–SOM model." The main innovations of this paper are as follows:

(i) This work systematically integrates PSO–VMD, multiscale adaptive sample entropy, and ALO–SOM into multi-class anomaly-signal classification for fixed-point deformation monitoring, constructing a complete collaborative-optimized processing pipeline.
(ii) End-to-end collaborative optimization framework: A series optimization process of "parameter optimization decomposition adaptive multi-scale feature extraction clustering network structure optimization" is constructed. PSO ensures that the signal is decomposed into the most physically meaningful modes; Adaptive sample entropy automatically extracts robust multi-scale complexity features; ALO ensures that SOM learns feature space topology with optimal structure. The three are linked together to achieve competitive overall performance.
(iii) Comprehensive system performance verification: First, the input construction method of random sampling of the training set; second, to carry out the identification of multiple rounds of non-repetitive experiments; Finally, the proposed method is compared to 18 representative models such as BP, general regression neural network (GRNN), SVM, random forest, long short-term memory (LSTM), and stacking integration. The dataset contains five types of actual data, including lightning interference, artificial blasting, power failure, collapse earthquake, and natural anomaly. The superiority and practicability of the proposed method are comprehensively verified from the accuracy, robustness, efficiency, and other dimensions.

This paper is divided into six chapters, and the specific structure is as follows:

(i) Section 1: Introduction. This paper expounds the importance of fixed-point deformation observation, the necessity of abnormal signal classification, and the limitations of traditional methods, leading to the research topic. This paper summarizes the main research content, innovation, and the structure of the full text.
(ii) Section 2: PSO–VMD multi-scale adaptive sample entropy feature extraction process. The basic principles of VMD, sample entropy, and adaptive improvement are introduced in detail, which lays a theoretical foundation for the construction of subsequent methods.
(iii) Section 3: Anomaly classification method based on ALO–SOM. The overall architecture of the proposed method is described in detail, and the specific steps and processes of ALO optimizing SOM parameters are detailed in the sections.
(iv) Section 4: Experimental design and result analysis. The composition and preprocessing process of the fixed-point deformation anomaly data set used in the experiment are introduced. The experimental setup, comparison model, and evaluation index are described in detail. The experimental results of the proposed method and 18 other comparative models were displayed and analyzed, and the effectiveness of each module was verified by ablation experiments.
(v) Section 5: Discussion. The experimental results are discussed in depth, the reasons for the advantages of the proposed method are analyzed, and the possible limitations of the method and the challenges in practical business applications are discussed.
(vi) Section 6: Conclusion and prospect. This paper summarizes the work of the full text, summarizes the main conclusions, and prospects the future research direction.

2. Fixed point deformation feature extraction process based on PSO–VMD and mixed multiscale entropy

2.1. Construction of adaptive sample entropy

Sample entropy (SampEn) is a time series complexity measurement method proposed by Richman *et al.*,[42] which reflects the probability of the system generating new patterns. The larger the value, the higher the complexity of the system, and the greater the probability of generating new patterns. Compared with other entropy measures (such as approximate entropy), the SampEn algorithm has better consistency; that is, the influence of dimension m and threshold r on sample entropy is equivalent, and the calculation efficiency is relatively high.

Because the original SampEn calculation relies on fixed parameters (embedding dimension m, similarity tolerance R), it is easy to suffer from the shortcomings of similar entropy overlapping and dissimilar entropy approaching for non-stationary vectors (such as noise and mutation feature vectors), and the discrimination of various features is insufficient. Therefore, a kind of adaptive sample entropy (ADA-sampen) with adaptive similarity tolerance R is designed, which discards the fixed r preset in the experiment, and combines the dynamic

adjustment of local standard deviation σ local and kurtosis K of the vector to highlight the difference between abrupt components and stationary vectors, so as to automatically adapt to the fluctuation characteristics of the vector, reduce parameter sensitivity, and deal with non-stationary/abrupt vectors (such as pulse and noise signals). The specific calculation steps are shown in Figure 1. The main steps of signal decomposition are as follows:

(i) Data preprocessing includes two parts: Detrending and normalization. First-order linear detrending is performed on the input original vector $x(t)$ to eliminate the influence of linear drift; after trend removal, apply Min–Max normalization to each individually extracted waveform, linearly mapping the time-series amplitudes to the interval [−1, +1] to eliminate differences in units and amplitudes between different deformation records.

$$x_{detread}(t) = x(t) - \left(\hat{a} \cdot t + \hat{b}\right) \tag{1}$$

$$x_{norm}(t) = 2 \times \frac{x_{detread}(t) - x_{min}}{x_{max} - x_{min}} - 1 \tag{2}$$

In **Equations 1** and **2**, $\hat{a},\hat{b}$ are linear fitting coefficients, and $\sum_{t=1}^{N}\left[x(t) - (\hat{a} \cdot t + \hat{b})\right]^2$ is the minimum; x_{min} is the minimum value, x_{max} is the maximum value; $x_{detread}(t)$ is the waveform after trend removal.

(ii) The key statistic, kurtosis KS, is calculated to characterize the data mutation. To calculate the global mean μ_{global} and global standard deviation σ_{global} of $x_{detread}(t)$, the kurtosis KS is calculated to characterize the spike/mutation characteristics of the data (the mutation vector k is larger, and the stationary vector k is close to 0). To calculate the local standard deviation by sliding window: Traverse the vector with the window size w, calculate the standard deviation in each window, and then calculate the mean value σ_{local_mean} of the standard deviation of all windows to reflect the local fluctuation characteristics of the data.

$$\sigma_{local}(i) = \sqrt{\frac{1}{w-1}\sum_{k=i}^{i+w-1}\left[x_{norm}(k) - \mu_{win}(i)\right]^2}, i = 1,2,\ldots,N-w+1 \tag{3}$$

$$\sigma_{local_mean} = \frac{1}{N-w+1}\sum_{i=1}^{N-w+1}\sigma_{local}(i) \tag{4}$$

$$KS = \frac{\frac{1}{N}\sum_{t=1}^{N}[x_{norm}(t) - \mu_{global}]^4}{\sigma_{global}^4} - 3 \tag{5}$$

In **Equations 3–5**, $\mu_{win}(i)$ is the mean value of the data in the i-th sliding win*d*ow; $\mu_{global} = \frac{1}{N}\sum_{t=1}^{N} x_{\text{detread}}(t)$, that is, the global mean, $\sigma_{global} = \sqrt{\frac{1}{N\text{-}1}\sum_{t=1}^{N}\left[x_{detread}(t) - \mu_{global}\right]^2}$, that is, the global standard deviation.

(iii) The similarity tolerance is calculated adaptively. The adaptive tolerance formula is introduced, where the tolerance threshold is dynamically adjusted by integrating the global standard deviation, mean local standard deviation, and kurtosis. A lower bound constraint is imposed on the tolerance parameter to prevent the failure of similar-pair counting caused by an excessively small tolerance value.

$$r_{adapt} = \alpha \cdot \sigma_{global} + \beta \cdot \frac{\sigma_{local_mean}}{|KS| + 1} \tag{6}$$

In **Equation 6**, α is the global reference coefficient, preset as 0.15; β is the local adjustment coefficient, preset as 0.05. The r_{adapt} achieved this way will be greater than or equal to 0.05, avoiding statistical deviation. The two coefficients α = 0.15 and β = 0.05 are determined based on the characteristics of fixed-point deformation observation waveforms and repeated preliminary tests. The global reference coefficient α controls the baseline tolerance magnitude, which is set to 0.15 to match the overall fluctuation range of tiltmeter time series after sample-wise Min–Max normalization. The local adjustment coefficient β = 0.05 is responsible for adjusting tolerance according to local waveform variability; a relatively small value prevents excessive sensitivity to local transient noise in deformation signals. The minimum tolerance bound of 0.05 is selected from exploratory trials: smaller thresholds would lead to insufficient matched vector pairs and unstable entropy estimation for low-amplitude, slow-varying deformation disturbances, whereas overly large tolerance reduces the capability to distinguish subtle differences between similar anomaly waveforms. The sliding window size w used for local standard deviation calculation is set to 120, which balances temporal resolution and statistical stability for our 4,000-point intercepted deformation waveform samples. This window length is capable of capturing short-term local fluctuations while avoiding over-smoothing that erases transient anomaly features.

(iv) Construct the embedded vector matrix: Based on the embedding dimension m, the m-dimensional embedding vector matrix is constructed (the number of rows is N - m + 1, and the number of columns is m). An m + 1-dimensional embedded vector matrix (N - m rows and m + 1 columns) is constructed to prepare for subsequent similar pair statistics (**Equations 7** and **8**).

$$X_m(i) = \left[x_{norm}(i), x_{norm}(i+1), \ldots, x_{norm}(i+m-1)\right], i = 1,2,\ldots,N-m+1 \tag{7}$$

$$X_{m+1}(i) = \left[x_{norm}(i), x_{norm}(i+1), \ldots, x_{norm}(i+m)\right], i = 1,2,\ldots,N-m \tag{8}$$

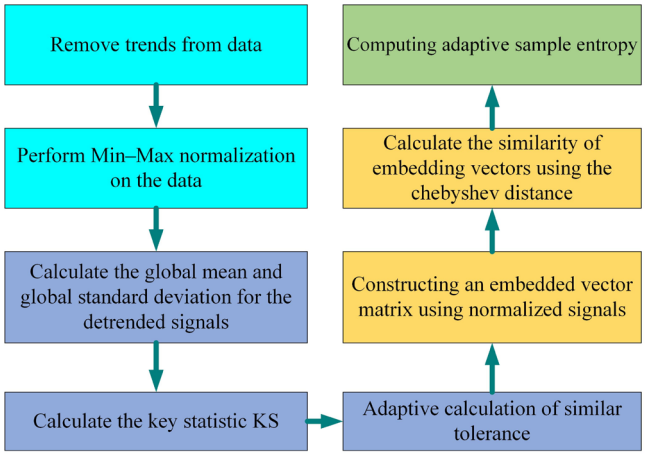


Figure 1. The process of constructing adaptive sample entropy
Abbreviation: KS: Kurtosis.

The embedding dimension m = 2 is selected in this study. For time-series entropy analysis, m = 2 is widely adopted for short geophysical observation sequences. Preliminary parameter tests show that m = 1 cannot adequately reflect the dynamic correlation of deformation time series, while m ≥ 3 sharply reduces the number of available embedded vectors for our 4000-point samples, leading to unreliable entropy estimates and degraded feature discrimination between different anomaly categories.

(v) Statistics of similar proportions. All pairwise combinations of m-dimensional embedded vectors are traversed. Chebyshev distance is used to judge vector similarity; we count the number of similar pairs and calculate the matching proportion C_m. Similarly, we traverse (m + 1)-dimensional embedded vectors and compute the similar-pair proportion C_{m+1}.

$$C_m = \frac{2}{(N-m+1)(N-m)} \sum_{i=1}^{N-m+1} \sum_{j=i+1}^{N-m+1} I\left(max\left|X_m(i) - X_m(j)\right| < r_{adapt}\right) \quad (9)$$

$$C_{m+1} = \frac{2}{(N-m)(N-m-1)} \sum_{i=1}^{N-m} \sum_{j=i+1}^{N-m} I\left(max\left|X_{m+1}(i) - X_{m+1}(j)\right| < r_{ada\,pt}\right) \quad (10)$$

In **Equations 9** and **10**, $I(\cdot)$ is an indicator function, which takes 1 if it meets the conditions; otherwise, it takes 0.

(vi) Calculate the adaptive sample entropy. Substitute the above proportions into the core formula of adaptive sample entropy to obtain the final entropy value (**Equation 11**).

$$\text{AdaSampEn} = -\ln\left(\frac{C_{m+1}}{C_m}\right) \quad (11)$$

It should be emphasized that adaptive sample entropy is a purely statistical metric that reflects the temporal complexity of time-series waveforms. It improves feature separability for classification tasks, but cannot independently infer geophysical genesis (blasting, power failure, tectonic deformation). The ground-truth physical category labels in this paper are determined by station logs, blasting archives, power-failure records, and manual waveform inspection, rather than derived from entropy values.

For methodological clarity, we further clarify the signal variables and statistical quantities used in the calculation pipeline. All intermediate statistical quantities, including global mean μ_{global}, global standard deviation σ_{global}, kurtosis KS, and the average of sliding-window standard deviations σ_{local_mean} are computed from the detrended signal $x_{detread}(t)$ before normalization. The normalized signal $x_{\text{norm}}(t)$ is only adopted for the subsequent adaptive sample-entropy main computation and vector embedding. Specifically, kurtosis KS is calculated using the global statistics of the full-length detrended waveform and characterizes overall impulsive features of the whole signal; σ_{local_mean} represents the average magnitude of local fluctuations obtained from sliding-window analysis, which serves as an extra correction term for adaptive tolerance and is not equivalent to the global standard deviation σ_{globa}.

2.2. Improved variational mode decomposition method for particle swarm optimization

The VMD is a frequency domain signal separation technology under the variational framework proposed by Zosso *et al.*[41] It has the computational advantages of strict mathematical constraints, no modal aliasing, and no significant end effects. It aims to adaptively decompose complex non-stationary and nonlinear signals into a series of quasi-orthogonal eigenmode functions with specific center frequency and bandwidth. The modal number k and

penalty factor α are two of the important preset parameters required for VMD calculation, which determine the number of variational mode functions (VMFs) decomposed and the bandwidth quality balance. Conventional empirical manual settings will bring unpredictable effects.

Particle swarm optimization optimizes the two key parameters of VMD, namely the mode number K and penalty factor α. The PSO algorithm searches for the optimal K and α matched to the noise level and spectral characteristics of the dataset for the batch of deformation anomaly signals. These optimized parameters are then adopted for the decomposition of all signals within this dataset. The advantages of this method are: (i) Stable decomposition results: It directly defines the fixed number of output modes of VMD and guarantees consistent feature dimensions for the subsequent feature extraction process; (ii) Controllable component quality: A properly selected α constrains the bandwidth of each variational mode function, balancing mode aliasing and over-smoothing and enabling each decomposed mode to contain primarily one dominant frequency component; (iii) Adaptive capability: The optimized parameters are automatically adapted to the signal characteristics of the target dataset. When migrated to data from other fixed-point deformation stations with different geological backgrounds and noise environments, the PSO–VMD framework can re-calculate and generate new parameter combinations by re-running the PSO optimizer on local station data, so as to adapt to varied signal features and avoid mode aliasing or over-decomposition caused by directly reusing parameters from another station. The key steps for PSO to optimize VMD are shown in Figure 2.

(i) Define the fitness function of PSO: Determining whether the decomposition effect is good or not requires a reliable evaluation index, and the minimum envelope entropy is usually used as the fitness function. The smaller the envelope entropy, the better the sparsity of VMF and the better the decomposition effect. This function will be called by PSO. The input is a particle position (two-dimensional vector), and the output is a fitness value (scalar).

- Given a particle position [k, α], perform VMD on each signal × (t) (signal length n) to obtain a total of M VMFs, and then perform a Hilbert transform on the m-th modal function $u_m(t)$, to obtain the analytical signal $\hat{u}_m(t)$ and its instantaneous amplitude (such as the envelope $e_m(t) = |\hat{u}_m(t)|$), then calculate the envelope normalization function $e_m^{'}(t)$, and finally calculate the entropy H_m.

$$\hat{u}_m(t) = u_m(t) + j\Gamma\left[u_m(t)\right] \tag{12}$$

$$H_m = -\sum_{t=1}^{N} e_m^{'}(t) \cdot ln\left(e_m^{'}(t)\right) \tag{13}$$

In **Equations 12** and **13**, $\Gamma[u_m(t)]$ is a Hilbert transform operator, which is a nonparametric transformation operation of modal functions. Its core is to construct orthogonal imaginary components for real-valued signals, so as to generate complex-valued analytical signals, and finally realize the extraction of instantaneous amplitude/frequency.

- Calculate the envelope entropy for all VMFs, and then sum them. However, since K will change, we take the sum of the envelope entropy of all VMFs as the fitness value $F(K, \alpha)$. Theoretically, the smaller the value, the better; that is, the optimization target is $minF(K, \alpha)$.

$$F(K,\alpha) = \sum_{m=1}^{M} H_m(K,\alpha) \tag{14}$$

In **Equation 14**, $H_m(K, \alpha)$ is the objective function value of the m-th signal under the parameter (k, α), that is, the sum of envelope entropy.

(ii) Set PSO parameters and variable boundaries, and use PSO to optimize K and α at the same time, where K is an integer and α is a continuous value. Since K is an integer, K is rounded when the particle position is updated. It should be noted that in the optimization process, K is a variable, so the number of VMFs decomposed each time is different.

(iii) When the PSO iteration stops and the optimal K and α are obtained, these two parameters are fixed for VMD decomposition of all signals; that is, K VMF components can be obtained for each fixed-point deformation anomaly signal.

The VMD algorithm calculates and updates the center-frequency estimate for each VMF during its solving procedure. After decomposition, VMF_1 yields the highest center frequency w_m, VMF_2 has the second-highest center frequency, and so forth, such that the last component VMF_M possesses the lowest center frequency. It should be noted that VMFs are mathematically derived narrow-band decomposition outputs rather than components with uniquely defined geophysical origins. For certain signal categories, VMF_1 (highest-frequency component) can exhibit fast-varying waveforms that resemble features of high-frequency noise or transient impulse events. Intermediate-order VMFs may display oscillatory behaviors associated with intermediate-frequency signal content. The last component, VMF_M (lowest-frequency component), often presents slow-varying temporal behavior, which can resemble trend-like

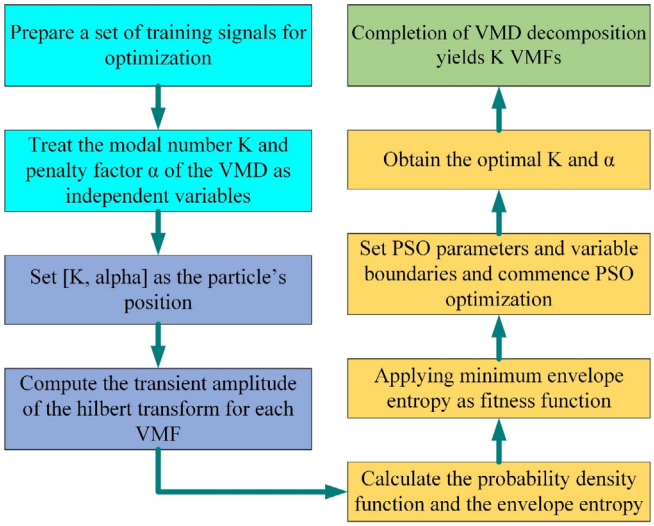


Figure 2. Computational workflow for optimizing the VMD model using PSO
Abbreviations: PSO: Particle swarm optimization; VMD: Variational mode decomposition; VMF: Variational mode functions.

terms or low-frequency oscillations observed in many time series. Analogously, such low-frequency components may correspond to rigid-body modes in vibration-signal analysis or long-term tendencies within financial time series; however, such correspondences are case-dependent and cannot be generalized arbitrarily.

Unlike EMD, the ordering of VMFs output by VMD is a priori and stable. This arises because the pre-defined mode number K is set at initialization, and the algorithm optimizes all K modes and their corresponding center frequencies simultaneously. Consequently, the input signal is decomposed into a set of quasi-orthogonal narrow-band modes with specific center frequencies. Under ideal decomposition conditions, the spectral support sets (frequency ranges containing major signal energy) of individual VMFs show little or no mutual overlap. Meanwhile, the output VMF sequence is automatically sorted according to the optimized center-frequency values. This avoids two well-known drawbacks of EMD: Unstable ordering of intrinsic mode functions induced by mode aliasing, and residual-component retention after decomposition.

3. Principle of fixed-point deformation anomaly classification based on ALO–SOM

3.1. Ant lion optimization algorithm

Ant lion optimization is a heuristic algorithm proposed by Australian scholars Seyedali Mirjalili and others in 2015, which is derived from the behavior of ant lions in capturing prey in nature.[43] Its inspiration comes from the unique hunting behavior of ant lion larvae, which build funnel-shaped traps in the sand and prey on ants. ALO algorithm searches for the optimal solution of the complex optimization problem by simulating the whole process of ant lions setting traps, trapping, killing, and reconstructing traps. ALO has been used in many engineering optimization fields because of its novel concept, few parameters, easy implementation, and good global search ability. ALO algorithm process is generally:

(i) Ant random walk: The set of steps of the ant random walk is defined as X_i^t, and its initial number of steps is 0. At the same time, the ant random walk has a limited range, and the ant position should be normalized according to **Equation 15**.

$$X_i^t = \frac{\left(X_i - a_i\right)\times\left(d_i^t - c_i^t\right)}{\left(b_i - a_i\right)} + c_i^t \tag{15}$$

In **Equation 15**, X_i^t is the normalized result of the ant random walk step set X_i in the feasible region; a_i and b_i are the minimum and maximum values of the i-dimensional ant position vector preset by the algorithm, respectively; c_i^t and d_i^t are the minimum and maximum values of the ant position in the i-dimensional under the t-th iteration, respectively.

(ii) Ant lion traps: When ants enter the pit dug by the ant lion by mistake, the action of ants will be limited, and the walking formula in the trap is:

$$c_i^t = A_j^t + c^t \tag{16}$$

$$d_i^t = A_j^t + d^t \tag{17}$$

In **Equations 16** and **17**, c^t is the minimum value of all variables at iteration t, d^t is the maximum value of all variables at iteration t, and A_j^t is the ant lion j selected at iteration t.

(iii) Ant lion trap target: When the ant lion finds the ant, it will throw sand at the edge of the sand pit to prevent the ant from escaping from the trap and make it continuously slide to the ant lion at the bottom of the pit. ALO algorithm uses roulette to find the appropriate location of the ant lion and dynamically reduces the trap range to speed up the capture of ants.

$$c^t = \begin{cases} \left|c^t\right| & ,t \le 0.1T \\ T \times c^t / \left(10^w \times t\right), & t > 0.1T \end{cases} \tag{18}$$

$$d^t = \begin{cases} d^t & ,t \le 0.1T \\ T \times d^t / \left(10^w \times t\right), & t > 0.1T \end{cases} \tag{19}$$

$$w = \begin{cases} 2, 0.10T < t \le 0.50T \\ 3, 0.50T < t \le 0.75T \\ 4, 0.75T < t \le 0.90T \\ 5, 0.90T < t \le 0.95T \\ 6, 0.95T < t \le T \end{cases} \tag{20}$$

In **Equations 18–20**, t is the preset upper limit of algorithm iteration times; w is the dynamic weight factor related to the current iteration times T, and the value conditions are shown in **Equation 20**, showing a stepwise increasing trend.

(iv) Ant hunting: When the prey falls to the bottom of the pit, the ant lion will move quickly and catch the prey. The biological phenomenon is reflected in the comparison of the fitness values of the ant and the ant lion in the algorithm. When the fitness value of the ant is better, the latest position of the ant lion is changed to the current position of the ant.

$$Antlion_j^t = Ant_i^t, f_{obj}^{ant}\left(t\right) < f_{obj}^{antlion}\left(t\right) \tag{21}$$

In **Equation 21**, $Antlion_j^t$ is the ant lion position at iteration T; Ant_i^t is the ant position obtained in the t-th iteration; $f_{obj}^{ant}(t)$ is the fitness function value obtained by the ant at the T iteration; $f_{obj}^{antlion}(t)$ is the fitness function value of the ant lion at the t iteration.

(v) Elite ant lion strategy: Ant wandering will be affected by the current elite ant lion position and the specified ant lion position of roulette. The ant position is defined as **Equation 22**.

$$Ant_i^{t+1} = \frac{R_A^t\left(l\right) + R_E^t\left(l\right)}{2} \tag{22}$$

In **Equation 22**, Ant_i^{t+1} represents the position of the i-th ant under iteration t + 1; $R_A^t\left(l\right)$ is the latest position determined by roulette under the t-th iteration to walk randomly around the ant lion for l step; $R_E^t\left(l\right)$ is the ant position of the t-generation elite ant lion (the best ant lion position calculated by each iteration) that randomly walks l steps nearby.

3.2. Basic principles of self-organizing map

An SOM is an unsupervised artificial neural network, also known as a Kohonen network. Its core function is to map high-dimensional and complex input data to a low-dimensional (usually one-dimensional or two-dimensional) discrete grid, while preserving the original topology and feature distribution of the data. In short, SOM is like an intelligent and self-learning "map plotter." It can organize a large number of disorderly data into an orderly "map," on which similar data points are close to each other, while dissimilar data points are far away from each other. The calculation steps of the SOM network are as follows:

(i) Competition process: Calculate the Euclidean distance d_j between the input vector x and the neuron weight vector W_j of the competition layer (**Equation 23**).

$$d_j = \left\|X - W_j\right\| = \sqrt{\sum_{i=1}^{m}\left(x_i\left(t\right) - w_{ij}\left(t\right)\right)^2} \tag{23}$$

where $\boldsymbol{X} = (x_1, x_2, \cdots, x_m)^T$ is the input vector; W_j is the weight vector of the competitive layer; w_{ij} is the weight between the input layer neuron I and the competition layer neuron J.

The best matching unit (BMU) is defined as the neuron with the smallest d_j, that is, $BMU = arg\ min_j \|X\text{-}W_j\|$.

(ii) Cooperation process: In this process, the weight of the BMU and adjacent neurons will change, in which the adjacent spatial range is defined by the Gaussian function h, and σ will decrease with time t, so as to dynamically adjust the adjacent spatial range of the BMU (**Equation 24**).

$$h\left(j, BMU, t\right) = exp\left(-\frac{distance(j, BMU)^2}{2\sigma(t)^2}\right) \tag{24}$$

(iii) Adaptation process: To achieve BMU approaching the input vector, the process updates the weights of the BMU and neurons in its field by introducing a learning rate η and a Gaussian function h, where the learning rate η will decrease with the number of iterations T based on the initial learning rate η_0, that is, $\eta(t) = \eta_0 exp(-t/\tau)$, where τ is a time constant, which determines the decay rate of the learning rate (**Equation 25**).

$$W_j(t+1) = W_j(t) + \eta(t) \cdot h(t) \cdot (x - W_j(t)) \quad (25)$$

(iv) Training: initialize weights, randomly select samples to calculate BMU, and then update BMU and domain weights; At the same time, the domain radius and learning rate will decrease with the increase in the number of iterations. As the iterations progress, the two gradually decrease, and the network is adjusted locally in a small range in a fine-grained manner (fine-tuning stage), and finally tends to be stable.

The SOM itself is an unsupervised model, and its training process does not need any label information. According to the clustering results and topological map, each neuron (or region) can be assigned a category label by using the generated topological map, and then the "label map" can be used to classify new samples.

3.3. Fixed-point deformation anomaly identification process

Based on the detailed discussion of the adaptive sample entropy calculation method and the detailed steps of PSO optimizing VMD in the previous chapters, combined with the ALO–SOM prediction model, an accurate identification process for a variety of fixed-point deformation abnormal signals is constructed, as shown in Figure 3.

Step 1: Perform data preprocessing such as null value replacement, fixed-length center interception, and data normalization on the original data of N fixed-point deformation anomalies to obtain the effective data set $X_i(i = 1,2,\cdots,N)$ involved in subsequent calculations;

In step 2, the abnormal signal X_i is decomposed into six intrinsic mode functions $VMF_{i,j}(j = 1,2,\cdots,10)$ using PSO–VMD technology, and can be linearly reconstructed into **Equation 26**:

$$X_i = \sum_{j=1}^{10} \mathrm{V}MF_{i,j} \quad (26)$$

Step 3 calculates the adaptive sample entropy of each VMF component to obtain the multi-scale adaptive sample entropy vector of the signal, and repeats steps 1 to 2 to successively calculate the adaptive sample entropy of each abnormal signal X_i, so as to form the input matrix S_0 of the SOM neural network.

In step 4, the complete dataset is partitioned via a two-stage, physical-event-oriented splitting strategy instead of simple individual-sample random sampling at a fixed ratio of 8:2. At the first-level split, all waveform fragments derived from the same real-world physical anomaly event are wholly assigned either to the 1st-generation training set or to the fully held-out 1st-generation test set. This guarantees that waveform fragments belonging to one physical anomaly event will never be distributed across both partitions. The held-out 1st-generation test set S_2 is completely isolated and is not involved in any parameter-optimization procedures. Subsequently, the 1st-generation training set S_1 is further divided at an 8:2 ratio through the second-level inner split to produce the 2nd-generation training set and the 2nd-generation validation set solely for hyperparameter tuning. Taking the 2nd-generation training set as the optimization object, the ALO algorithm is applied to optimize critical network parameters, including the competitive-layer dimension and training epochs of the SOM model. The identification error-rate expression (**Equation 27**) serves as the fitness function during optimization. After optimization completes, the original parameter values of the SOM model are substituted with the optimized ones to construct the final ALO–SOM diagnostic model. The held-out 1st-generation test set is only adopted for final performance assessment after all network hyperparameters are fully fixed.

During each iteration of ALO optimization, the fitness value is calculated using only samples from the 2nd-generation validation set. For each sample in this validation subset, the SOM network outputs a clustering result, which is mapped to the predefined physical category label of its activated competitive-layer neuron. This predicted category is then compared against the known ground-truth label of the anomaly signal. If the two labels match, the sample is marked as correctly identified; otherwise, it is marked as misclassified. The classification accuracy (ACC) is calculated as the ratio of correctly identified samples to the total number of samples within the 2nd-generation validation set. The corresponding error rate, defined as 1-ACC, is fed into **Equation 27** as the fitness metric to guide the search of the ALO algorithm. No data from the held-out 1st-generation test set participates in this fitness calculation at any iteration.

$$y = \frac{|100 - Acc|}{100} \quad (27)$$

where Acc denotes the classification accuracy percentage value of the SOM model, and y denotes the error rate of the SOM model.

In step 5, the ALO–SOM model is applied to the first-generation training set and first-generation test set. Based on the first-generation training set S_1, the SOM clustering tag value $l_p(p = 1,2,\cdots,P)$ of P abnormal signals is obtained. Based on the first-generation test set S_2, the SOM clustering tag value $L_q(q = 1,2,\cdots,Q)$ of Q abnormal signals is obtained. The L_q value and l_p value are matched one by one. If the tag values are the same, they are regarded as the same type of anomaly by the ALO–SOM model, and if L_q and l_p represent

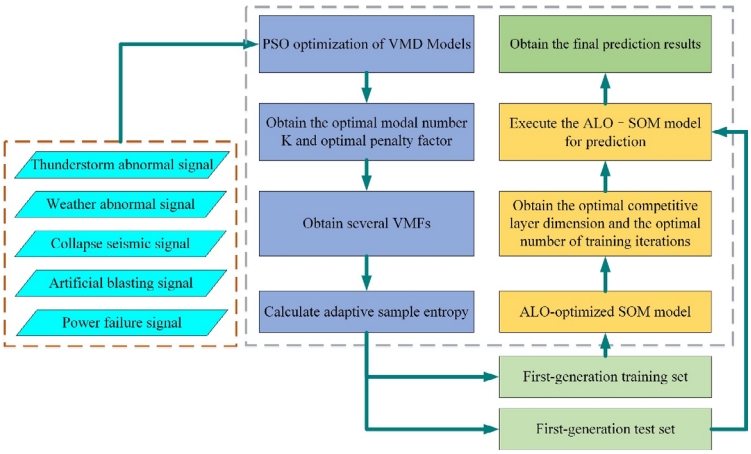


Figure 3. A process for fixed-point deformation anomaly classification combining PSO–VMD adaptive sample entropy and the ALO–SOM model
Abbreviations: ALO: Ant lion optimizer; PSO: Particle swarm optimization; SOM: Self-organizing map; VMD: Variational mode decomposition; VMF: Variational mode functions.

the true attributes of abnormal signals, if the tag values T_q and T_p are also identical, it is judged that the diagnosis is correct; otherwise, the classification is wrong. After all the Q abnormal signals to be tested in the test set have completed the above diagnostic steps, the identification is completed, and the final overall identification result is obtained.

In this study, a rigorous two-stage event-level dataset partitioning strategy is adopted to strictly isolate training optimization data and final evaluation data, effectively eliminating potential information leakage:

(i) First-level global partitioning: The entire dataset is divided at an 8:2 ratio by independent physical anomaly events, generating the first-generation training set and the fully held-out first-generation test set. All waveform samples derived from a single physical event are entirely assigned to one subset to avoid event-level data mixing. Critically, the first-generation held-out test set is completely isolated throughout all parameter optimization, feature tuning, and model training procedures, and is only used for final generalization performance verification.

(ii) Second-level internal partitioning (for hyperparameter optimization): The first-generation training set is further subdivided at an 8:2 ratio into a second-generation training subset and a second-generation validation subset. Within the PSO–VMD and ALO–SOM optimization loop, the second-generation training subset is utilized for model training and iterative parameter searching, while the second-generation validation subset (split entirely from the first-generation training data pool) serves to monitor overfitting and calculate fitness values for ALO hyperparameter optimization. No data from the held-out first-generation test set participates in any optimization iteration.

(iii) Final testing pipeline: After all PSO–VMD decomposition parameters and ALO–SOM network hyperparameters are fully fixed and finalized, standardized VMD decomposition and adaptive sample entropy feature calculation are performed uniformly on the isolated first-generation held-out test set. The finalized ALO–SOM model is then applied to these unseen test features to obtain objective and reliable classification accuracy metrics.

It is worth noting that both PSO and ALO are stochastic heuristic optimization algorithms. They can find high-quality near-optimal solutions within the predefined parameter-search space, but cannot provide mathematical guarantees for true global optimality.

4. Experiment and analysis

This paper adopts observational datasets from the Yixian seismic station, Hebei Province, China. The tilt-meter instrument is deployed inside an observational cave near the fault boundary between the North China Plain and the Taihang Mountains. The site lithology mainly consists of diorite, limestone, and shale. Long-term continuous observations have captured abundant anomalous waveform records belonging to five typical categories: thunderstorm interference, power-supply failure, artificial blasting, collapse-related events, and meteorological interference. Brief physical definitions for the five anomaly categories are given as follows:

(i) Thunderstorm interference: Short-duration impulse-like distortion induced by nearby lightning activities;

(ii) Power-supply failure: Step-shaped or abrupt waveform shifts caused by power supply fluctuation and power outage;
(iii) Artificial blasting: Transient oscillation signals triggered by man-made quarry or construction blasting;
(iv) Collapse-related events: Waveforms generated by local rock collapse and small-scale collapse earthquakes;
(v) Meteorological interference: Slow-varying disturbances originating from environmental temperature, humidity, and atmospheric pressure changes.

The dataset spans from April 15, 2016, to November 20, 2021. There are 200 original real-world abnormal event records, which are augmented to form 1,000 valid waveform samples used for subsequent experiments. Ground-truth labels are determined through multi-source joint verification to guarantee dataset reliability: (i) station operation and maintenance logs; (ii) local archives of blasting activities and power-failure incidents; (iii) manual visual waveform inspection performed by geophysical observers; (iv) historical observation records from the local meteorological office. Samples with ambiguous or mixed features that cannot be reliably categorized are excluded before formal experiments.

For preprocessing, missing-value points in raw waveforms are assigned zero values. Each waveform fragment is center-intercepted with a window length of 4,000 points centered at the absolute maximum anomaly amplitude. Each intercepted waveform sample is further processed with individual-sample Min–Max normalization to eliminate dimensional effects. In the stage of experimental data collection and preprocessing, continuous annual observation data are managed as independent daily files. Each raw daily data file is labeled with only one primary physical event, and only one experimental sample is extracted via the fixed time window according to the position of the anomalous waveform. Physical anomalous events are thus adopted as the minimum unit for dataset partitioning: all derived samples from the same physical anomaly event are entirely assigned to either the first-generation training set or the reserved first-generation test set. This rule ensures that data from one physical event will not appear in both subsets simultaneously, fundamentally avoiding time-series data leakage and over-optimistic classification results. The second-level dataset partition is performed only inside the first-generation training set for model hyperparameter optimization, without breaking the boundaries of physical events. The complete logic of dataset partitioning is supplemented in Section 4 of the manuscript.

4.1. PSO–VMD multiscale decomposition and adaptive sample entropy calculation

Before performing specific parameter calculations, some parameters need to be preset. For example, the convergence tolerance of VMD is 1×10^{-7}, the noise tolerance is set to 0, and the number of PSO particles and the maximum number of iterations are 20 and 30, respectively. The lower bound (LB) and upper bound (UB) for PSO searching are defined as LB = [6, 100] and UB = [12, 5,000], where the first element corresponds to the search range for mode number K and the second corresponds to penalty factor α. The optimal mode number obtained by PSO optimizing VMD is 10, the optimal penalty factor is 2,000, and the embedding dimension, adaptive threshold ratio, and local window length of adaptive sample entropy are 2, 0.15, and 10, respectively. The fixed-point deformation anomaly characteristics of five types of deformation anomaly signals after VMD decomposition and multi-scale adaptive sample entropy calculation are shown in Figure 4 and Table 1, where the ordinate of Figure 4 is the normalized amplitude and the abscissa is the time series(s).

After PSO–VMD decomposition for five categories of fixed-point deformation signals, namely lightning interference, power failure, artificial blasting, collapse-related events, and meteorological interference, 10 VMF components are yielded. Each VMF approximates an amplitude-modulation–frequency-modulation signal with quasi-zero mean, where the counts of extreme points and zero-crossing points are comparable. Figure 4 presents the temporal morphology of each decomposed VMF, which are ordered from high center frequency to low frequency. Comparing VMF_3 to VMF_9 across the five signal categories, we observe that the waveforms of lightning interference, collapse events, and power failure share similar morphological features: impulse-like waveforms appear near the center of the time axis, while the remaining segments of these VMF traces remain relatively flat. In contrast, the medium-frequency and high-frequency components (VMF_1–VMF_9) of artificial blasting and meteorological interference exhibit sustained anomalous fluctuations, rapid waveform variations, and abundant energy distribution; only the lowest-frequency VMF_{10} displays slowly varying features and spectral peak patterns. Such morphological distinctions imply that these two signal types contain relatively strong environmental noise or continuous impact-related signal components. Based on the PSO–VMD decomposition results, adaptive sample entropy values are computed for each VMF of the 1,000 abnormal waveform samples, constructing the 10-dimensional feature set used for the subsequent identification experiment (Figure 5).

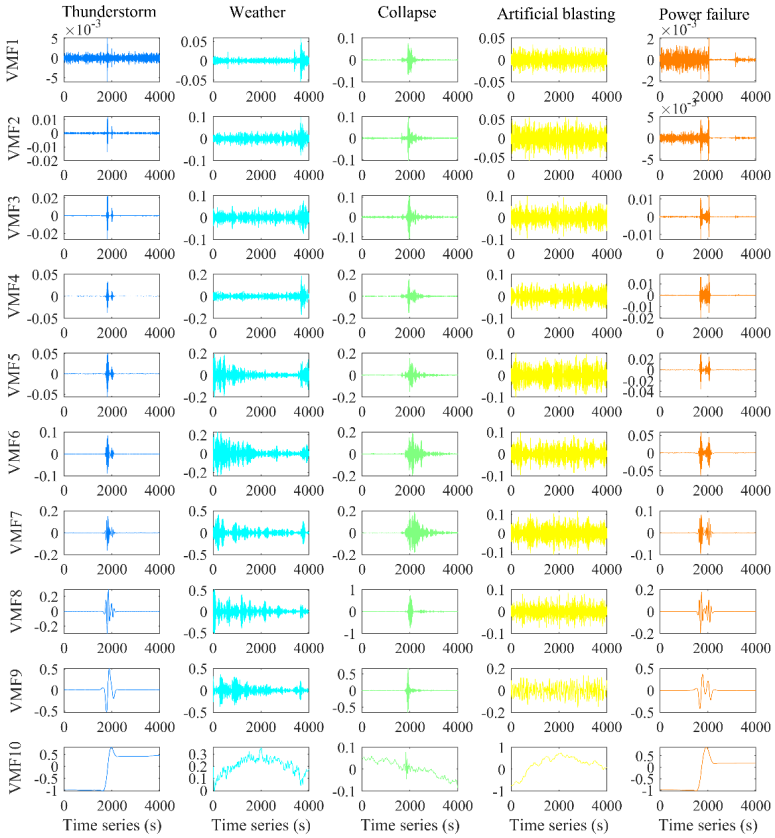


Figure 4. Comparison of PSO–VMD decomposition results for different fixed-point deformation anomaly signals
Abbreviations: PSO: Particle swarm optimization; VMD: Variational mode decomposition; VMF: Variational mode functions.

Table 1. Adaptive sample entropy statistics table for different deformation anomaly signals on each variational mode function (VMF)

VMF	Lightning interference	Abnormal weather	collapse earthquake	Manual blasting	Power failure
VMF_1	1.4901 ± 0.2831	1.0090 ± 0.2712	1.0674 ± 0.0341	0.9668 ± 0.0359	1.5510 ± 0.3541
VMF_2	1.3819 ± 0.1632	0.8302 ± 0.0891	0.7946 ± 0.0147	0.7789 ± 0.0139	1.2714 ± 0.3188
VMF_3	1.1349 ± 0.0928	0.7613 ± 0.0788	0.7199 ± 0.0134	0.7715 ± 0.0133	1.0822 ± 0.2622
VMF_4	0.9403 ± 0.0816	0.6788 ± 0.0782	0.7008 ± 0.0224	0.6734 ± 0.0204	0.9013 ± 0.1286
VMF_5	0.8246 ± 0.0785	0.6413 ± 0.0945	0.7503 ± 0.0344	0.6745 ± 0.0211	0.8204 ± 0.0574
VMF_6	0.8852 ± 0.0996	0.6231 ± 0.1327	0.8541 ± 0.0452	0.6369 ± 0.0504	0.8330 ± 0.0789
VMF_7	0.9617 ± 0.0930	0.6019 ± 0.1675	0.9530 ± 0.0491	0.7314 ± 0.0324	0.8757 ± 0.1041
VMF_8	0.9374 ± 0.1006	0.6522 ± 0.1930	0.9445 ± 0.0416	0.7577 ± 0.0327	0.9044 ± 0.0950
VMF_9	0.8624 ± 0.1041	0.7576 ± 0.1692	0.8312 ± 0.0395	0.8298 ± 0.0375	0.9163 ± 0.1524
VMF_{10}	0.8061 ± 0.1225	0.8836 ± 0.1346	0.7208 ± 0.0342	0.7686 ± 0.0245	0.8133 ± 0.1124

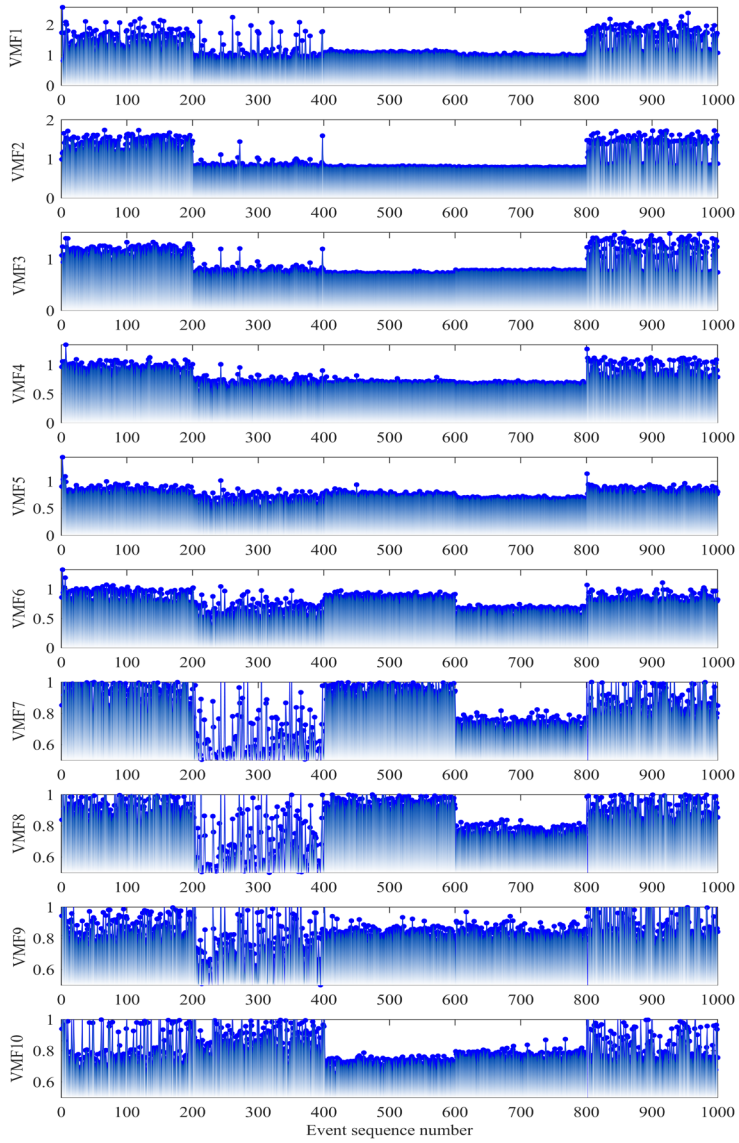


Figure 5. Comparison of multiscale adaptive sample entropy calculation results for 10 variational mode functions (VMFs)

In the adaptive sample-entropy waveform analysis, each VMF exhibits different discriminative power for the five anomaly categories (lightning interference, meteorological anomalies, collapse-related events, artificial blasting, and power failure). For instance, the entropy values of VMF_1–VMF_4 show relatively good separation for lightning interference and power-failure signals, yet their discriminative capacity is weaker for meteorological anomalies, collapse-related events, and artificial blasting. The entropy distributions from VMF_5–VMF_6 yield only modest inter-class differences across all five signal types, whereas VMF_7–VMF_{10} provide better discriminative performance for meteorological-anomaly and artificial-blasting signals. Therefore, the entropy features derived from individual VMFs show inconsistent distinguishing capabilities across categories, which necessitates a robust nonlinear anti-noise machine-learning model to achieve accurate discrimination among these five classes of abnormal signals. To quantitatively analyze the statistical differences in adaptive sample entropy across anomaly types and VMF components, the mean value and standard deviation of entropy for each anomaly category under every VMF are calculated, as listed in Table 1.

Similar to the visual-based observation derived from Figure 5, the entropy values for meteorological anomalies, collapse-related events, and artificial blasting show relatively close distributions across multiple VMFs (VMF_1–VMF_4). Comparatively, VMF_{10} exhibits stronger

discriminative capability for the five abnormal-signal categories, while the remaining VMFs (VMF_5–VMF_9) yield relatively poorer separation performance compared with VMF_{10}. Nevertheless, certain abnormal-signal types still present highly similar entropy distributions. To further quantify the discriminative power of these entropy-based features for the five deformation-anomaly categories, we adopt the Kruskal–Wallis non-parametric test together with analysis of variance (ANOVA). A composite-score metric is introduced as the final evaluation indicator, which is computed via weighted summation of five standardized metrics: *ANOVA*(*P*), Kruskal–Wallis *p*-value (*K-W*(*P*)), *F*-statistic, between-to-within-class variance ratio, and decision-tree-based information gain (*InfoGain*) (**Equation 28**).

$$Score = w_1 \cdot Z_{ANOVA(P)} + w_2 \cdot Z_{K\text{-}W(P)} + w_3 \cdot Z_{F-stat} + w_4 \cdot Z_{VarRatio} + w_5 \cdot Z_{InfoGain} \quad (28)$$

The corresponding weights are set as $w_1 = 0.25$, $w_2 = 0.20$, $w_3 = 0.15$, $w_4 = 0.20$, $w_5 = 0.20$. These weight coefficients are assigned according to the relative discriminative contribution of each metric for multi-class waveform separation on the fixed-point deformation dataset: higher weights are allocated to hypothesis-testing indicators that directly reflect inter-class statistical differences, while relatively lower weights are given to statistics that mainly describe the magnitude of feature differences. Here, *Z* denotes the standardized value of each individual indicator. This composite score evaluates the overall hypothesis-testing performance of adaptive sample entropy from each VMF in distinguishing five categories of deformation anomalies, and the corresponding results are summarized in Table 2.

According to the calculation results of the composite score, the ranking of the adaptive sample entropy test effect of each VMF is: 2 > 4 > 7 > 3 > 6 > 5 > 1 > 8 > 10 > 9. According to the score range, eight features (2, 4, 7, 3, 6, 5, 1, and 8) have high discrimination (composite score ≥ 0), and features 10 and 9 have relatively weak discrimination (composite score < 0). From the secondary test index, ANOVA analysis will identify the features with $p < 0.05$ as significant, and the *p*-value results show that the number of significant features is 10 (100.0%); that is, the intergroup differences of these 10 criteria are far greater than the intragroup differences, which is suitable for use as multiclassification criteria. Similarly, Kruskal–Wallis will identify the features with $p < 0.05$ as significant, and the *p*-value results show that the number of significant features is 10 (100.0%); that is, the differences between the 10 criteria are much greater than the differences within the group, which is also suitable for use as a multiclassification criterion. Finally, the correlation between the evaluation indexes was calculated. The results showed that the *p*-value correlation between ANOVA and Kruskal–Wallis was 0.945, the correlation between *F*-statistics and variance ratio was 1.000, and the average correlation between information gain and other indexes was 0.784.

Table 2. *T*-test results for each VMF adaptive sample entropy in identifying fixed-point deformation anomalies

Adaptive sample entropy	ANOVA(P)	K–W(P)	*F*-stat	Variance radio	Info-Gain	Composite score	Rank
2	2.3251×10^{-268}	1.9317×10^{-161}	619.9075	2.4920	0.0029	0.7409	1
4	3.6514×10^{-254}	3.7567×10^{-155}	564.6231	2.2698	0.0031	0.6882	2
7	7.0620×10^{-228}	1.7733×10^{-140}	471.3645	1.8949	0.0036	0.6361	3
3	2.7726×10^{-226}	1.1143×10^{-140}	466.0659	1.8736	0.0034	0.5479	4
6	4.5403×10^{-210}	1.3496×10^{-137}	414.3298	1.6656	0.0031	0.3417	5
5	7.1412×10^{-187}	7.9212×10^{-151}	346.7447	1.3939	0.0030	0.1723	6
1	6.9973×10^{-161}	6.5023×10^{-146}	279.1487	1.1222	0.0033	0.1477	7
8	9.7225×10^{-162}	3.0918×10^{-125}	281.2508	1.1306	0.0031	0.0696	8
10	2.9642×10^{-56}	4.2975×10^{-79}	75.9913	0.3055	0.0022	−1.3343	9
9	8.8779×10^{-39}	8.9785×10^{-35}	50.5643	0.2032	0.0021	−2.0106	10

Abbreviations: ANOVA: Analysis of variance; K–W: Kruskal–Wallis; VMF: Variational mode functions.

4.2. Prediction effect of ALO-optimized SOM model on multiple anomalies

To verify the effectiveness of the ALO–SOM neural network model for identifying fixed-point deformation anomaly signals, 100 repeated experiments were carried out to automatically classify five anomaly types: lightning interference, meteorological anomaly, collapse-induced earthquake, artificial blasting, and power-supply failure. The training-to-test sample ratio was set to 8:2. Detailed experimental results are presented in Figures 6–8. Figure 6 illustrates the hyperparameter-searching results of the ALO-optimized SOM across the 100-run fixed-point deformation anomaly prediction experiments. Figure 7 reports the prediction accuracy obtained by the ALO–SOM model under 100 rounds of random-sampling trials, and Figure 8 provides the boxplot of prediction results from these 100 random-sampling experiments. For the ALO algorithm, the LB and UB of the search domain are [1, 1] and [10, 10], respectively. The number of search agents is 20, the maximum number of iterations is 50, and the dimension of independent variables is 2.

Since 80% of the total samples are randomly selected for the construction of the training set, the optimal competitive layer dimension and the optimal training times obtained from each round of the ALO-optimized SOM model are different. This is because the "knowledge base" learned by the SOM model is a different subset of samples when each round of ALO starts the search, which leads to differences in the optimal mapping (such as the grid that can best retain the subset topology) of the SOM when learning the topology, which is specifically reflected in the change in the dimension (grid shape) of the optimal competitive layer. At the same time, some subsets may be more "easy to fit," and the model converges quickly, so the optimal training times are shorter; some subsets may have more noise or fuzzy search boundaries and need more iterations to be stable, so the optimal training times increase. Based on each pair of optimal hyperparameters obtained by ALO, substitute them into the SOM model to replace the original default values, and then perform the anomaly identification step to get 100 prediction results (Figure 7).

Based on the statistical accuracy curves derived from 100 independent repeated test-set experiments, the proposed ALO–SOM model maintains a stable and high overall classification level, with all repeated test-set accuracy values exceeding 97%. Benefiting from the random partitioning of training and test subsets in each trial, slight numerical volatility exists across the multi-round experimental results, which reflects the statistical robustness of the model rather than performance instability. In terms of single-category classification performance on the test set, the model presents the strongest robustness for power-supply failure signals, with only one individual trial obtaining a relatively lower accuracy of 97.1428%.

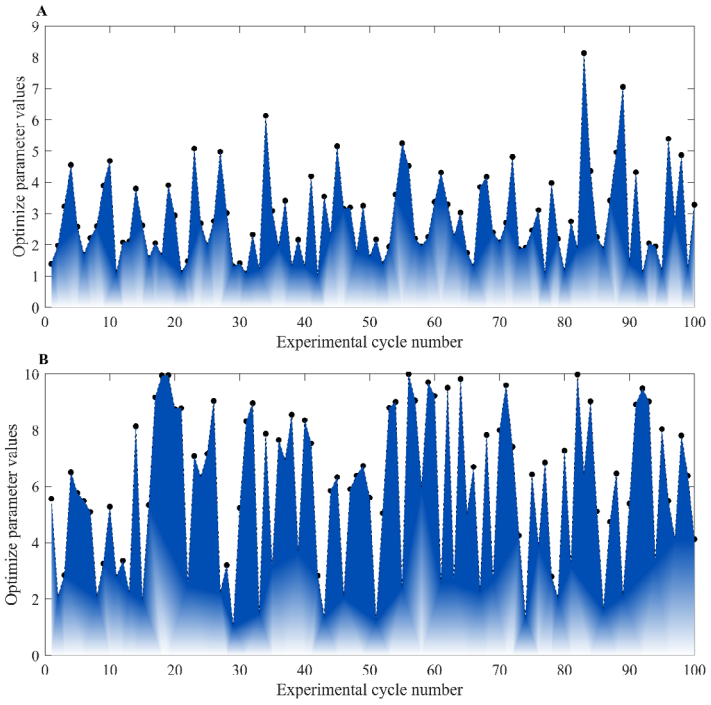


Figure 6. Hyperparameter search results for ant lion optimizer-optimized self-organizing map following 100-cycle fixed-point deformation anomaly prediction experiments. (A) Competitive network layer dimensions. (B) Number of training epochs.

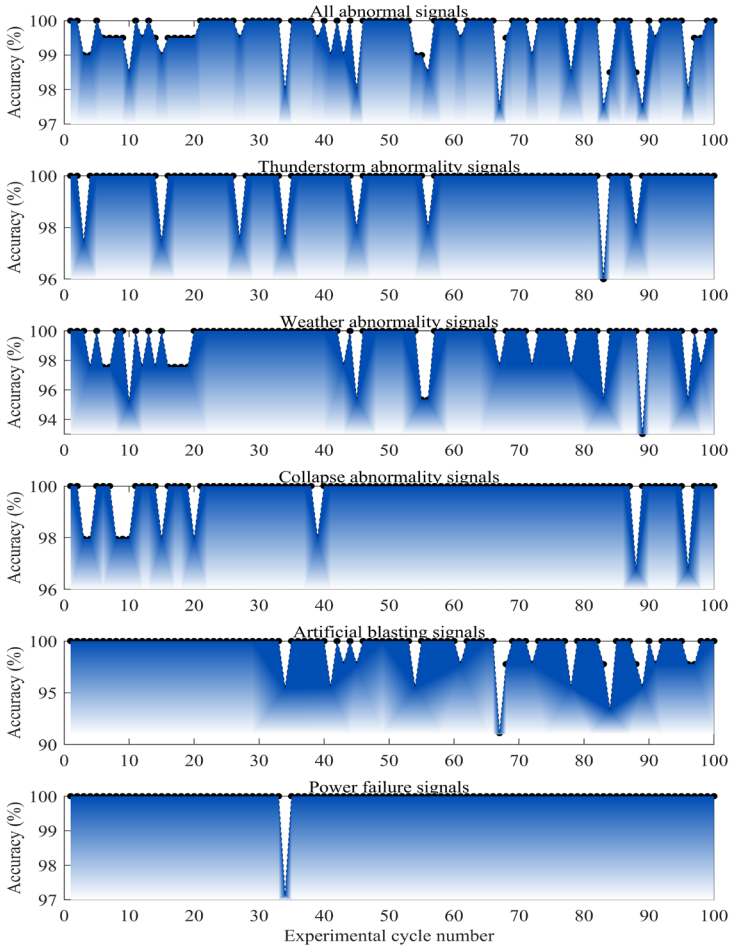


Figure 7. Prediction accuracy results of the ant lion optimizer–self-organizing map model under 100 rounds of random sampling

By comparison, a small number of repeated trials yield relatively low classification accuracy for other anomaly categories; in particular, the minimum test-set accuracy of artificial blasting signals drops to 91.1111% across the 100 trials. To intuitively visualize the statistical distribution and fluctuation characteristics of multi-round classification performance, a boxplot is plotted for the 100 repeated test-set classification results of fixed-point deformation anomalies. The boxplot quantitatively characterizes key statistical indicators, including the median, quartiles, and outliers, as illustrated in Figure 8.

Figure 8 reveals the statistical visualization differences between a single exception type and the overall sample. The total sample, "lightning interference," "weather anomaly," "collapse earthquake," "artificial blasting," "power failure," and other abnormal types are automatically detected with 3, 4, 5, 3, 5, and 1 outlier, respectively, in the box plot. The difference in the number of outliers between the total and a single abnormal type reflects the uneven distribution of the advantages and disadvantages of the classification results of each type, which is also consistent with the treatment method of random sampling of the training set. The statistical indicators such as the mean value, standard deviation, range value, median, and worst value of 100 rounds of fixed-point deformation anomaly classification results are shown in Table 3.

Statistical results derived from 100 independent repeated test-set experiments show that the overall classification accuracy achieves a mean value of 99.6200%, a standard deviation of 0.6361%, a median of 100%, a worst-case accuracy of 97.5000%, and a range of 2.5000%, demonstrating the stable and excellent overall performance of the proposed model. In terms of single-category classification stability on the test set, the power-supply failure anomaly presents the highest stability with a standard deviation of only 0.2857. In contrast, the artificial blasting category exhibits the poorest stability, with a standard deviation of 1.5452%. It also corresponds to the minimum accuracy (91.1111%) and the maximum accuracy range (8.8888%) across all anomaly types. This

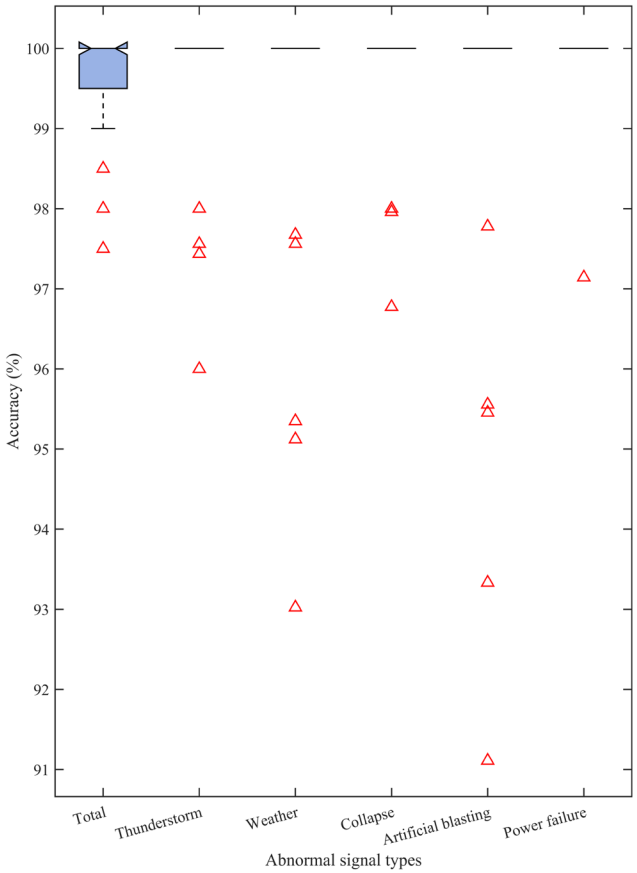


Figure 8. Box plot of ant lion optimizer–self-organizing map model prediction results under a 100-round random sampling experiment

Table 3. Statistical results of ant lion optimizer–self-organizing map model prediction for point-specific deformation anomalies

Statistical value (%)	Overall (%)	Lightning interference (%)	Abnormal weather (%)	Collapse earthquake (%)	Manual blasting (%)	Power failure (%)
Mean value	99.6200	99.7999	99.3131	99.7726	99.3989	99.9714
Standard deviation	0.6361	0.7038	1.4703	0.7021	1.5452	0.2857
Median	100	100	100	100	100	100
Worst value	97.5000	96.0000	93.0232	96.7742	91.1111	97.1428
Range value	2.5000	4	6.9767	3.2258	8.8888	2.8571

phenomenon is mainly attributed to the partial feature similarity between artificial blasting signals and other deformation anomaly signals. For instance, the waveform morphology of VMF_{10} for meteorological anomalies (Figure 4) is highly overlapped with that of blasting signals, posing a certain challenge for the accurate classification of the ALO–SOM model.

4.3. Comparison of prediction effects under various machine learning models

To verify the performance advantages of the ALO–SOM model for anomaly classification, eighteen competing baseline models were adopted for comparative evaluation, including standard BP neural network, GRNN, probabilistic neural network (PNN), radial-basis-function

(RBF) network, decision tree, SVM, AdaBoost, naive Bayes, linear discriminant analysis, quadratic discriminant analysis, random forest, k-nearest neighbors (KNN), logistic regression, multi-layer perceptron (MLP), LSTM, gated recurrent unit (GRU), voting ensemble, and stacking ensemble. Overall accuracy and per-class accuracy (%) of prediction outputs were utilized as evaluation metrics. Consistent with the preceding experiments, the dataset was partitioned via random sampling, with 80% of samples assigned to the training set and 20% to the test set.

To ensure fair horizontal comparison, all models adopt identical input features and preprocessing pipelines, where Min–Max normalization is uniformly applied to map features into the range [0,1]. Stratified random sampling is used for dataset partitioning: the original samples are split into the first-generation training set and the first-generation test set at a ratio of 8:2. The first-generation test set is never involved in any model parameter optimization and serves as a unified external test set. Only the proposed ALO–SOM model further divides the first-generation training set into the second-generation training set and the second-generation validation set at an 8:2 ratio for hyperparameter optimization using the ALO. The ALO is configured with a population size of 30 and a maximum iteration number of 50. The optimized variables include the row and column numbers of the SOM competitive layer grid (search range: 4–12, integers), the initial neighborhood radius of SOM [1.0, 5.0], and the learning rate attenuation coefficient [0.6, 0.95]. The classification error rate on the second-generation validation set is taken as the objective function for minimization. The maximum number of training iterations for the SOM network is set to 200, and Euclidean distance is adopted to measure sample similarity.

The 18 comparison models do not utilize the second-generation subsets, and no hyperparameter search is performed; all hyperparameters are fixed. For SVM with RBF kernel, the penalty coefficient C = 10 and kernel bandwidth $\gamma = 0.1$. The random forest consists of 100 base decision trees with a maximum tree depth of 10. AdaBoost uses decision trees of depth 1, with 80 base learners and a learning rate of 0.1. The number of neighbors for KNN is set to k = 5. Neural networks, including MLP, LSTM, GRU, deep neural network, and convolutional neural network, share a unified maximum number of iterations of 200 with the Adam optimizer. The smoothing parameters of GRNN and PNN are fixed at 0.2 and 0.1, respectively. The standalone SOM adopts a fixed 8×8 grid. Voting and Stacking ensemble models use the preset combination of base classifiers. It should be noted that the core purpose of this group of comparisons is to evaluate the performance improvement contributed by the ALO adaptive hyperparameter-optimization module. The proposed ALO–SOM represents a complete pipeline with meta-heuristic automatic hyperparameter searching, while all eighteen baseline models use empirically preset fixed hyperparameters without adaptive tuning. This experiment compares model schemes with and without adaptive hyperparameter optimization, rather than comparing bare model structures under identical hyperparameter-search conditions.

A total of 100 repeated experiments are carried out to evaluate model stability. In each trial, stratified sampling is re-performed on the original dataset to generate an independent first-generation training set and first-generation test set. Within each trial, ALO–SOM re-executes ALO-based parameter optimization, while the comparison models are trained with fixed hyperparameters. Evaluation metrics for all models are calculated on the corresponding first-generation test set in each run. The specific comparison results are shown in Figure 9 and Table 4.

The mean test-set overall accuracy and per-class accuracy of the 18 machine-learning models exhibit substantial discrepancies, which indicates that models built on different theoretical principles show markedly different matching performance for these nonlinear noise-contaminated deformation features. For overall classification performance, only seven models, namely SVM, AdaBoost, random forest, KNN, MLP, LSTM, voting ensemble, and stacking ensemble, yield mean test-set accuracy values inferior to the proposed ALO–SOM framework (99.62%, averaged across 100 independent repeated experiments). Among the compared methods, the PNN achieves the lowest mean test-set accuracy of 20%; it can only identify lightning interference signals and is therefore unsuitable for multi-class classification of fixed-point deformation anomalies. Models with mean test-set accuracy below 80% (including GRNN, RBF, and linear discriminant analysis) present large room for performance improvement, and dedicated optimization would be required before practical deployment.

In terms of per-class classification performance, apart from the aforementioned low-performance models, Naive Bayes, quadratic discriminant analysis, and GRU show unsatisfactory classification results for the power-supply failure category, with mean test-set accuracies of 45%, 62.5%, and 75%, respectively. This poor per-class performance directly pulls down their overall classification accuracy below 90%. In addition, BP and GRU also achieve relatively low mean test-set accuracies for the lightning interference class (75% and 77.5%). The detailed quantitative classification results of all 18 comparative models are summarized in Table 4.

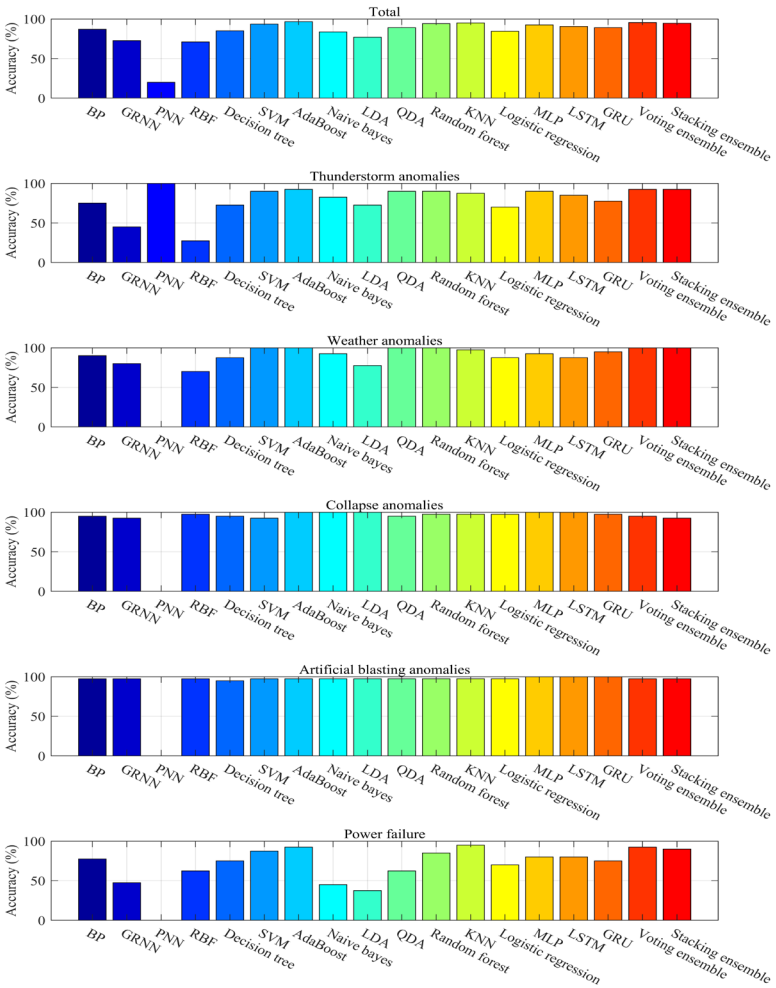


Figure 9. Histogram comparing the performance of 18 machine learning models in identifying pointwise deformation anomalies
Abbreviations: BP: Backpropagation; GRNN: General regression neural network; GRU: Gated recurrent unit; KNN: k-nearest neighbors; LDA: Linear discriminant analysis; LSTM: Long short-term memory; MLP: Multi-layer perceptron; PNN: Probabilistic neural network; QDA: Quadratic discriminant analysis; RBF: Radial basis function; SVM: Support vector machine.

As listed in Table 4, among all comparative models, AdaBoost achieves the highest mean test-set accuracy of 96.5% across 100 independent repeated experiments, followed by the Voting ensemble model (95.5%), KNN (95.0%), Stacking ensemble (94.5%), and SVM (93.5%). These several models exhibit competitive comprehensive classification performance on the five-class deformation anomaly classification task and can serve as alternative baseline methods for the proposed ALO–SOM framework. In contrast, the remaining comparative models yield mean test-set accuracies below 87% in either overall or single-class classification. Such limited performance renders them incompetent for solving the nonlinear multi-class classification problem of complex deformation anomaly signals, and these methods fail to maintain a favorable balance between single-class classification stability and overall classification accuracy in prediction results.

The BP neural network exhibits highly unbalanced per-class classification performance. It achieves high accuracy for manual-blasting samples reaching 97.5%, whereas its accuracy drops to 75.0% for lightning-interference samples and 77.5% for power-failure samples, yielding an overall accuracy of only 87%. This phenomenon is mainly caused by two aspects. First, lightning-interference and power-failure anomalies contain diverse transient impulse waveforms, leading to large intra-class dispersion of extracted features. Second, the BP neural network has limited capability for nonlinear clustering of multi-source mixed deformation signals. It can well capture the obvious oscillatory characteristics of blasting-related waveforms, but struggles to separate complex transient features from lightning interference and power-failure events. Such imbalanced classification performance across different anomaly categories degrades its overall classification effect.

Table 4. Comparison of fixed-point deformation anomaly signal classification results across 18 common machine learning models

Machine learning model	Overall (%)	Lightning interference (%)	Abnormal weather (%)	Collapse earthquake (%)	Manual blasting (%)	Power failure (%)
BP	87.0	75.0	90.0	95.0	97.5	77.5
GRNN	72.5	45.0	80.0	92.5	97.5	47.5
PNN	20.0	100.0	0	0	0	0
RBF	71.0	27.5	70.0	97.5	97.5	62.5
Decision tree	85.0	72.5	87.5	95.0	95.0	75.0
SVM	93.5	90.0	100.0	92.5	97.5	87.5
AdaBoost	96.5	92.5	100.0	100.0	97.5	92.5
Naive bayes	83.5	82.5	92.5	100.0	97.5	45.0
LDA	77.0	72.5	77.5	100.0	97.5	37.5
QDA	89.0	90.0	100.0	95.0	97.5	62.5
Random forest	94.0	90.0	100.0	97.5	97.5	85.0
KNN	95.0	87.5	97.5	97.5	97.5	95.0
Logistic regression	84.5	70.0	87.5	97.5	97.5	70.0
MLP	92.5	90.0	92.5	100.0	100.0	80.0
LSTM	90.5	85.0	87.5	100.0	100.0	80.0
GRU	89.0	77.5	95.0	97.5	100.0	75.0
Voting ensemble	95.5	92.5	100.0	95.0	97.5	92.5
Stacking ensemble	94.5	92.5	100.0	92.5	97.5	90.0

Abbreviations: BP: Backpropagation; GRNN: General regression neural network; GRU: Gated recurrent unit; KNN: k-nearest neighbors; LDA: Linear discriminant analysis; LSTM: Long short-term memory; MLP: Multi-layer perceptron; PNN: Probabilistic neural network; QDA: Quadratic discriminant analysis; RBF: Radial basis function; SVM: Support vector machine.

The 95% confidence interval of the overall accuracy of the ALO–SOM model from 100 repeated runs is calculated as [99.49%, 99.75%]. The narrow confidence interval demonstrates that the fluctuation of model prediction performance is limited and the generalization result is stable. One-way ANOVA was performed on the 100-round overall accuracy of ALO–SOM and all 18 baseline models; the result yields $F = 247.36$ and $p < 0.001$, which proves that significant differences exist in classification accuracy among these models. The subsequent Tukey's honestly significant difference post-hoc multiple comparison test further confirms that the ALO–SOM model significantly outperforms all 18 comparative models ($p < 0.001$). Combined with the previous feature distinguishability verification based on ANOVA, the Kruskal–Wallis nonparametric test, and the composite score evaluation for each VMF component, these statistical results demonstrate that the high 99.62% average classification accuracy is not an accidental outcome caused by random data splitting. Instead, the performance advantage originates from the high discriminability of selected VMF adaptive sample entropy features and the hyperparameter optimization advantage of the ALO–SOM framework. The confidence interval and hypothesis test jointly validate the statistical credibility of the proposed fixed-point deformation anomaly identification method.

5. Discussion

5.1. Technical mechanism analysis of method effectiveness

The classification framework proposed in this paper, which integrates PSO–VMD, adaptive sample entropy, and ALO–SOM, achieves significantly better classification performance than 18 comparison models on five types of fixed-point deformation anomaly data. The effectiveness of the framework depends on collaborative optimization at the following three levels:

(i) Firstly, at the signal decomposition level, PSO–VMD realizes high-fidelity and low-aliasing preprocessing. In traditional VMD, if the preset mode number k is too small, the signals of different physical processes will be accidentally integrated into the same mode, resulting in mode aliasing; If the penalty factor α is not appropriate, it will affect the center frequency and bandwidth of the mode, resulting in incomplete or over-smoothed decomposition. PSO automatically finds the best {K, α} parameter pair by using the sum of the entropies of the VMF envelopes as the fitness function. This ensures that the original mixed signal is adaptively decomposed into a series of eigenmode functions with discrete central frequencies and relatively clear physical meaning, which fundamentally lays a high-quality foundation for subsequent feature extraction and is the primary premise for the success of the overall method.

(ii) Secondly, at the feature extraction level, multi-scale adaptive sample entropy constructs robust and rich feature expression. The traditional sample entropy uses a fixed similarity tolerance, which is sensitive to noise and amplitude changes. The adaptive sample entropy determines the optimal tolerance at each scale through a data-driven approach, which significantly improves the robustness of features. More importantly, applying this method to each IMF component decomposed by PSO–VMD is equivalent to analyzing the complexity of the signal from different frequency bands (scales). Experiments show that the low-frequency IMF entropy of natural slowly changing anomalies is significantly different from the high-frequency IMF entropy mode of transient interference such as lightning and explosion. This multi-scale and adaptive entropy feature vector can reveal the inherent dynamic complexity differences of the five types of anomalies more profoundly than the single-scale time-domain or frequency-domain Statistics (such as mean and variance), and provide high-discriminant input for a machine learning classifier.

(iii) Finally, at the level of classification decision, ALO–SOM realizes the accurate mapping of network structure optimization. The strength of SOM lies in its topology-preserving property, which can nonlinearly project the high-dimensional feature space to the two-dimensional competitive layer, so that similar samples are gathered. However, the grid dimension of its competitive layer (such as 5×5 or 10×10) and the number of training times greatly affect the clustering effect. If the dimension is too small, different categories will be forced to squeeze together, resulting in confusion; If it is too large, it may be overfitted and reduce the prediction efficiency. ALO can efficiently find the optimal structural parameters of the SOM network by simulating the intelligent random search mechanism of ant lion hunting and guided by the classification accuracy. This avoids the blindness of manual parameter adjustment and makes the structure of the SOM network best match the complexity of the problem to be classified, so as to give full play to its clustering and classification ability.

5.2. Comparative analysis with existing methods

Compared to the 18 comparative models listed, the method presented in this paper shows systematic advantages:

(i) Compared to the traditional machine learning models: BP, SVM, decision tree, and other models, which directly process the original signal or simple features, there are "inborn deficiencies" in the ability of feature expression, and it is difficult to deal with the non-stationary and low signal-to-noise ratio deformation data, which is the bottleneck of their performance.

(ii) Compared with complex deep learning models, LSTM, GRU, and other sequential models are theoretically suitable for processing time-series data, and their performance depends heavily on massive annotated data. Through the strategy of "optimal decomposition + physical meaning feature extraction," this method achieves better results with fewer but higher-quality features.

(iii) Compared to the integration class model, the integration of voting and stacking improves the generalization ability by combining multiple base learners, which has a certain effect on the overall classification rate and the classification effect of a single type, which proves that the single model may have limitations. The ALO–SOM in this paper can be regarded as a deeply optimized "supercell model," and its performance is better than that of simple integration, indicating that the targeted optimization of a single model for a specific problem (feature space) may be more effective than the collective voting mechanism.

5.3. Limitations of the method

Although the method is effective, it still has the following limitations and needs to be improved in future work:

(i) Calculation cost of parameter optimization: The PSO–VMD feature extraction stage took an average of 38.30 s, with a standard deviation of 0.241 s; the ALO–SOM iterative training and optimization took an average of 4.30 s, with a standard deviation of 0.047 s; once the model parameters had fully converged

and stabilized, ALO–SOM required an average of only 0.210 s to make a test prediction for a single new sample. It can be seen that the time overhead of the entire method is primarily concentrated in the PSO–VMD parameter optimization and feature computation stages. The nested PSO iterative search, which repeatedly solves for modal parameters for each segment of the deformation signal, accounts for the vast majority of the computation time; ALO–SOM training accounts for a relatively smaller proportion of the time, while the computational cost of the model inference stage is very low. The average total time per round for the entire training and optimization process of the original model is 42.60 s. The training stage is computationally intensive, and high-precision classification results are achieved at the cost of high training computational overhead; the native approach is more suitable for offline post-analysis of historical deformation waveforms from fixed-point stations. This paper proposes two lightweight acceleration schemes for subsequent engineering implementation: Scheme 1 employs a pre-trained parameter library grouped by station combined with a local fine-tuning strategy, eliminating the need for global PSO searches on each new waveform; Scheme 2 combines SOM competitive layer neuron pruning with GPU parallel acceleration to reduce the time required for iteration and feature extraction.

(ii) Interpretability and physical relevance of features: Although the adaptive sample entropy feature is effective, its physical meaning is relatively abstract (representation complexity). How to establish a more intuitive and quantitative correlation between these high-dimensional statistical characteristics and specific geophysical processes (such as blasting vibration energy, step characteristics of power failure) is the key to improving the physical interpretability of the method.

(iii) Limited generalization verification: All experiments are based on single-station tilt data from Yixian station (2016–2021). Differences in geological background, cave environment, and instrument noise among different fixed-point deformation stations will affect model performance. Multi-station and multi-instrument cross-validation is required in further research.

(iv) Poor geophysical interpretability of statistical features: Adaptive sample entropy only reflects signal complexity and lacks direct physical correspondence with deformation mechanisms. It can only serve as classification input. Further fusion of physically-driven features (tidal residual, amplitude-frequency indicators) is required for geophysical interpretation.

(v) The optimal modal number K and penalty factor α yielded by PSO–VMD are not static fixed parameters, but are adaptively optimized according to the noise level and spectral features of the input dataset. When applying this framework to observational data from other fixed-point deformation stations, re-running the PSO optimization on local station samples can automatically obtain site-adapted K and α. This adaptive searching mechanism enhances the robustness and adaptability of signal decomposition and subsequent feature-extraction workflows for different station-specific signals. However, the specific optimal parameters K = 10 and α = 2,000 acquired from the Yixian seismic station cannot be directly transplanted to other sites; direct reuse may lead to modal aliasing or over-decomposition for waveforms under distinct noise backgrounds.

(vi) Asymmetric hyperparameter-tuning configuration for comparative experiments. The proposed ALO–SOM framework incorporates automated hyperparameter searching based on an internal validation subset, while all eighteen comparative baseline models adopt empirically fixed hyperparameters without adaptive optimization. This experimental design targets the performance gain from the introduced optimization module. Nevertheless, this workflow difference may potentially benefit the performance of ALO–SOM. Full-scale grid-search tuning for all baseline models is not implemented due to the heavy computational burden of 100 repeated experimental runs.

6. Conclusion

To enable automatic high-precision identification of multiple predefined abnormal signal categories in fixed-point deformation monitoring, this paper develops and validates a technical framework implementing an end-to-end collaborative optimization pipeline covering signal preprocessing, feature extraction, and classification decision-making. PSO optimizes key parameters for VMD to achieve adaptive, high-fidelity decomposition of mixed observational signals. Based on the decomposed modes, multi-scale adaptive sample entropy is adopted to extract robust, discriminative complexity-oriented features. The ALO further optimizes the competitive-layer topology and training procedure of the SOM network, yielding an efficient classifier adapted to the characteristics of the input feature space.

Future research should focus on three areas: model lightweighting and acceleration, optimization of the feature system, and expansion and validation of the dataset.

In terms of model engineering optimization, the current nested iterative methods of PSO–VMD and ALO–SOM have two main drawbacks. These drawbacks are high training computational overhead and limited direct applicability for real-time station operations. To address these limitations, we should comprehensively employ multiple acceleration techniques to improve algorithmic efficiency. First, we will introduce GPU parallel computing to accelerate the VMD decomposition and the iterative optimization of envelope entropy, thereby reducing the time required for batch waveform processing. Second, for PSO–VMD, we will conduct offline pre-optimization for fixed stations. During the operational inference phase, we will directly reuse the previously converged optimal K and parameters, bypassing the PSO iterative search process. Third, we will simplify the ALO–SOM model by pruning the structure of the SOM competitive layer network. This simplification will be implemented under the condition that the loss in classification accuracy remains controllable, so as to reduce the network grid dimension. Concurrently, we will optimize the ALO iterative termination criterion by setting an error threshold to prematurely terminate unnecessary iterative loops, thereby reducing the computational load of hyperparameter search. Fourth, we explore knowledge distillation techniques to transfer the discriminative knowledge from the optimized complex SOM model to a lightweight network, further reducing online inference overhead and facilitating the deployment of the algorithm in operational seismic station scenarios.

From the perspective of feature optimization and physical interpretability enhancement, the current adaptive sample entropy-based feature system still has certain limitations. Although multi-scale entropy statistics can effectively quantify waveform complexity differences and support high-precision intelligent classification, these purely statistical indicators lack explicit geophysical constraint information and cannot directly reflect the intrinsic physical mechanisms of different deformation anomalies. To further improve the generalization ability and practical applicability of the identification model in future engineering applications, subsequent research will abandon the single statistical feature dependence mode. We will fuse multi-dimensional physically constrained features, including tidal residual variation characteristics, waveform amplitude attenuation features, dominant frequency spectral energy distribution, and time-domain mutation parameters, to construct a coupled statistical-physical collaborative feature system. This hybrid feature framework can retain the high separability advantage of entropy statistical features while introducing geophysical a priori constraints, effectively solving the problem of ambiguous physical interpretation of pure entropy classification results. It can further improve the discrimination stability under complex noise interference and enhance the physical credibility and mechanism rationality of intelligent identification results, providing more reliable technical support for real-time monitoring, anomaly tracing, and geodynamic analysis of fixed-point deformation observation stations.

With regard to the generalization validation of the dataset and algorithm, the experimental data in this study were obtained exclusively from a single seismic station; the model's generalization performance across stations and instruments has not yet been fully tested. In the future, we will continue to collect anomalous samples from fixed-point deformation observations at multiple stations, involving various instrument types and different geological backgrounds, to enhance the multi-source annotated dataset. We will also conduct model transfer and cross-validation across different geological environments to further evaluate the practical applicability of this algorithm in real-world operational settings.

This study lays a preliminary foundation for advancing fixed-point deformation anomaly identification from laboratory-level academic prototypes toward practical application. It offers valuable technical references for earthquake monitoring analysis and geoscience-related exploration research.

Acknowledgments

None.

Funding

This research was financially supported by Mahasarakham University; Hubei Provincial Natural Science Foundation of China (No.2026AFB752;2025AFB834); Open Fund of Wuhan Gravitation and Solid Earth Tides, National Observation and Research Station, No. WHYWZ202406; The Spark Program of Earthquake Technology of CEA, No.XH25019C; Earthquake Monitoring and Early Warning Tasks, No.CEA-JCYJ-202601025; Hubei Provincial Program for Science and Technology Innovation Platforms, No.2026CFC002; Research grants from the National Institute of Natural Hazards, Ministry of Emergency Management of China (Grant Number: ZDJZDJ2025-58); 2025 Doctoral Special Support Program Project of Chengdu Jincheng College (NO.2025JCKY(B)0018).

Conflict of interest

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

Author contributions

Conceptualization: Cong Pang, Wugang Ma, Piyapatr Busababodhin
Formal analysis: Tianwen Zhao, Guoqing Chen, Wugang Ma, Chawei Li
Investigation: Cong Pang, Wugang Ma, Chawei Li, Yong Jiang
Methodology: Cong Pang, Tianwen Zhao, Chawei Li, Yong Jiang, Piyapatr Busababodhin
Writing–original draft: Cong Pang, Tianwen Zhao, Guoqing Chen
Writing–review & editing: Cong Pang, Tianwen Zhao, Piyapatr Busababodhin

Availability of data

The data that support the findings of this study are not publicly available due to restrictions imposed by the relevant seismic monitoring authorities. These data were collected from the Yixian seismic station in Hebei Province, China, and contain sensitive geophysical monitoring records. Access may be granted to qualified researchers upon reasonable request to the corresponding author, subject to institutional approval.

Further disclosure

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.

References

1. Kabel M, Böhlke T, Schneider M. Efficient fixed point and Newton–Krylov solvers for FFT-based homogenization of elasticity at large deformations. *Comput Mech*. 2014;54(6):1497-1514.

   doi: 10.1007/s00466-014-1071-8

2. Schirmer H. Fixed point sets of deformations of pairs of spaces. *Topol Appl*. 1986;23(2):193-205.

   doi: 10.1016/0168-8641(86)90041-6

3. Shimada S, Bock Y. Crustal deformation measurements in central Japan determined by a Global Positioning System fixed-point network. *J Geophys Res: Solid Earth*. 1992;97(B9):12437-12455.

   doi: 10.1029/92JB00809

4. Sun W, Okubo S, Fu G, Araya A. General formulations of global co-seismic deformations caused by an arbitrary dislocation in a spherically symmetric earth model-applicable to deformed earth surface and space-fixed point. *Geophys J Int*. 2009;177(3):817-833.

   doi: 10.1111/j.1365-246X.2009.04113.x

5. Zhang Y, Wang D, Zhao Y, Shi H. Present and progress of fixed-point deformation observation. *Geomatics and Information Science of Wuhan University*. 2022;47(6):830-838.

   doi: 10.13203/j.whugis20220104

6. Liu S, Pang C, Wang X, Song M. Comparison and Analysis of Multiple Entropy Feature Tests for Abnormal Signals in Fixed-Point Deformation Monitoring. *International Scientific Technical and Economic Research*. 2026;4(2):206-231.

   doi: 10.71451/istaer2622

7. Peng Z, Shao Y. 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). 2024:598-604.

   doi: 10.1109/ICAICE63571.2024.10864303

8. Sheng D, Yang J, Wang H, Zhao CS, Sheng C. Interference and precursor signal analysis based on random forest and improved logistic regression model. *International Scientific Technical and Economic Research*. 2025;3(1):8-24.

   doi: 10.71451/m417wa34

9. Chen F, Diao F, Haghighi MH, et al. Mechanism and implications of the post-seismic deformation following the 2021 Mw 7.4 Maduo (Tibet) earthquake. *Geophys J Int*. 2024;237(1):203-216.

   doi: 10.1093/gji/ggae034

10. Doan NP, Nguyen BP, Park SS. Seismic deformation analysis of earth dams subject to liquefaction using UBCSAND2 model. *Soil Dyn Earthq Eng*. 2023;172:108003.

    doi: 10.1016/j.soildyn.2023.108003

11. Guo ZM, Wang ZY, Lei Y, Li YM. High-speed railway signal processing based on a combination of VMD and KSVD. *Appl Geophys*. 2025:1-16.

    doi: 10.1007/s11770-025-1223-y

12. Kohler M, Puzrin AM. Mechanism of co-seismic deformation of the slow-moving La Sorbella landslide in Italy revealed by MPM analysis. *J Geophys Res Earth Surf*. 2022;127(7).

    doi: 10.1029/2022JF006618

13. Zhang S, Jia P, Hou B, Hao M. Vertical deformation before and after the 2022 Menyuan Ms6.9 earthquake and analysis of earthquake precursors. *Geod Geodyn*. 2023;14(6):582-588.

    doi: 10.1016/j.geog.2023.03.005

14. Zhao S, McClusky S, Cummins PR, Miller MS. Co-seismic and post-seismic deformation associated with the 2018

Lombok, Indonesia, earthquake sequence, inferred from InSAR and seismic data analysis. *Remote Sens Environ*. 2024;304:114063.

doi: 10.1016/j.rse.2024.114063

15. Kulshrestha A, Chang L, Stein A. Use of LSTM for sinkhole-related anomaly detection and classification of InSAR deformation time series. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2022;15:4559-4570.

doi: 10.1109/JSTARS.2022.3180994

16. Maes K, Salens W, Feremans G, Segher K, Francois S. Anomaly detection in long-term tunnel deformation monitoring. *Eng Struct*. 2022;250:113383.

doi: 10.1016/j.engstruct.2021.113383

17. Zhang J, Tang H, Zhou B, Wen T, Zhang S. A new early warning criterion for landslides movement assessment: Deformation Standardized Anomaly Index. *Bull Eng Geol Environ*. 2024;83(5):205.

doi: 10.1007/s10064-024-03672-3

18. Milillo P, Sacco G, Di Martire D, Hua H. Neural network pattern recognition experiments toward a fully automatic detection of anomalies in InSAR time series of surface deformation. *Front Earth Sci*. 2022;9:728643.

doi: 10.3389/feart.2021.728643

19. Zhang T, Zhang W, Cao D, Yi Y, Wu X. A new deep learning neural network model for the identification of InSAR anomalous deformation areas. *Remote Sens*. 2022;14(11):2690.

doi: 10.3390/rs14112690

20. You H, Hua L, Su T. Power Transformer Fault Diagnosis and Prediction Based on Deep Learning. *Journal of Digital Frontier*. 2026;1(1):144-184.

doi: 10.67541/jdf2606

21. Wang W. Design Scheme Style Feature Extraction and Machine Learning Classification Algorithm Construction. *International Scientific Technical and Economic Research*. 2026;4(3):1-21.

doi: 10.67541/istaer2623

22. Castillo-Villamor L, Hardy A, Bunting P, et al. The Earth Observation-based Anomaly Detection (EOAD) system: a simple, scalable approach to mapping in-field and farm-scale anomalies using widely available satellite imagery. *Int J Appl Earth Obs Geoinf*. 2021;104:102535.

doi: 10.1016/j.jag.2021.102535

23. Liu C, Pan J, Wang J. An LSTM-based anomaly detection model for the deformation of concrete dams. *Struct Health Monit*. 2024;23(3):1914-1925.

doi: 10.1177/14759217231199569

24. Wibowo A, Purnama SR, Pratama C, Heliani LS, Sahara DP, Wibowo ST. Anomaly detection on displacement rates and deformation pattern features using tree-based algorithm in Japan and Indonesia. *Geod Geodyn*. 2023;14(2):150-162.

doi: 10.1016/j.geog.2022.07.003

25. Ma X, Lin Z, Jiao L, et al. Phase identification with VMD and HT combined method for an active seismic source experiment. *Measurement*. 2022;201:111689.

doi: 10.1016/j.measurement.2022.111689

26. Ou B, Zhang C, Xu B, et al. Innovative approach to dam deformation analysis: integration of VMD, fractal theory, and WOA-DELM. *Struct Control Health Monit*. 2024;(1):1710019.

doi: 10.1155/2024/1710019

27. Zhiyao L, Yong D, Denghua L. Coupling VMD and MSSA denoising for dam deformation prediction. *Structures*. 2023;58:105503.

doi: 10.1016/j.istruc.2023.105503

28. Zhou J, Zhang Y, Wang H, Yang J, Li P, Wang J. SABO-optimized VMD for seismic damage assessment of frame structures considering soil–structure interaction. *Buildings*. 2025;15(11):1822.

doi: 10.3390/buildings15111822

29. Bo Y, Guo X, Liu Q, Pan Y, Zhang L, Lu Y. Prediction of tunnel deformation using PSO variant integrated with XGBoost and its TBM jamming application. *Tunn Undergr Space Technol*. 2024;150:105842.

doi: 10.1016/j.tust.2024.105842

30. Li Q, Cheng F, Zhang X. Numerical simulation and deformation prediction of deep pit based on PSO-BP neural network inversion of soil parameters. *Sensors*. 2024;24(10):2959.

doi: 10.3390/s24102959

31. Li X, Jia C, Zhu X, Zhao H, Gao J. Investigation on the deformation mechanism of the full-section tunnel excavation in the complex geological environment based on the PSO-BP neural network. *Environ Earth Sci*. 2023;82(13):326.

doi: 10.1007/s12665-023-10963-7

32. Cui S, Jia J, Pang X, Wen J, Shi Y, Zeng J. A data-driven method with sample entropy and CEEMDAN for short-term performance degradation prediction of dynamic hydrogen fuel cells. *Int J Hydrogen Energy* 2024;83:916-932.

doi: 10.1016/j.ijhydene.2024.08.123

33. You H, Hua L, Su T. Economic Dispatch Optimization of Microgrid Based on an Improved Particle Swarm Optimization Algorithm. *International Scientific Technical and Economic Research*. 2026;4(3):90-124.

doi: 10.67541/istaer2627

34. Delgado-Bonal A, Marshak A. Approximate entropy

and sample entropy: a comprehensive tutorial. *Entropy*. 2019;21(6):541.

doi: 10.3390/e21060541

35. Bigdeli A, Maghsoudi A, Ghezelbash R. Application of self-organizing map (SOM) and K-means clustering algorithms for portraying geochemical anomaly patterns in Moalleman district, NE Iran. *J Geochem Explor*. 2022;233:106923.

doi: 10.1016/j.gexplo.2021.106923

36. Rahman ATMS, Kono Y, Hosono T. Self-organizing map improves understanding on the hydrochemical processes in aquifer systems. *Sci Total Environ* 2022;846:157281.

doi: 10.1016/j.scitotenv.2022.157281

37. Roeva O, Zoteva D, Roeva G, Lyubenova V. An efficient hybrid of an ant lion optimizer and genetic algorithm for a model parameter identification problem. *Mathematics*. 2023;11(6):1292.

doi: 10.3390/math11061292

38. Mirjalili S. The ant lion optimizer. *Adv Eng SoftwAdv Eng Softw*. 2015;83:80-98.

doi: 10.1016/j.advengsoft.2015.01.010

39. Vu Hong Son P, Soulisa FV. A Hybrid Ant Lion Optimizer (ALO) Algorithm for Construction Site Layout Optimization. *J Soft Comput Civ Eng*. 2023;7(4):50-71.

doi: 10.22115/scce.2023.365303.1540

40. Wang A, Pang C, Chen G, Li C, Zhao T. Optimization of the SOM neural network model using CEEMDAN distribution entropy and ALO for seismic and blasting identification. *J Seism Explor*. 2025;34(2):28-43.

doi: 10.36922/JSE025280033

41. Zosso D, Dragomiretskiy K, Bertozzi AL, Weiss PS. Two-Dimensional Compact Variational Mode Decomposition. *J Math Imaging Vis*. 2017;58(2):294-320.

doi: 10.1007/s10851-017-0710-z

42. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol - Heart Circ Physiol*. 2000;278(6):H2039-H2049.

doi: 10.1152/ajpheart.2000.278.6.H2039

43. Akbari MA, Zare M, Azizipanah-Abarghooee R, Mirjalili S, Deriche M. The cheetah optimizer: a nature-inspired metaheuristic algorithm for large-scale optimization problems. *Sci Rep*. 2022;12(1):10953.

doi: 10.1038/s41598-022-14338-z