

# Fault Diagnosis of Electric Submersible Pump System Based on Motor Current Signal Analysis and Deep Learning Method

Xin Lu<sup>1\*</sup>, Guoqing Han<sup>1</sup> , Luting Wang<sup>2</sup>, Zhuangzhuang Zhang<sup>3</sup>, and Xingyuan Liang<sup>1</sup> 

<sup>1</sup>Institute of Petroleum Engineering, China University of Petroleum-Beijing

<sup>2</sup>Chinese Academy of Sciences Institute of Mechanics

<sup>3</sup>Exploration & Development Research Institute of Changqing Oilfield Company

## Summary

The petroleum industry relies heavily on electric submersible pump (ESP) systems, which can cause substantial economic losses if malfunctioning. However, over half of these malfunctions are due to submersible electric motor (SEM) failures. To address this issue, motor current signature analysis (MCSA) is used in this work to monitor the condition of the SEM. This allows for further fault diagnosis of the ESP system, ensuring that malfunctions are detected and addressed promptly. In this study, a portable high-speed electrical signal acquisition tool was used to capture the current signal at a 10-KHz rate to monitor the SEM condition. Afterward, the fast Fourier transform (FFT) method was used to transform time-domain data into the frequency domain for spectrum analysis. Next, the motor rotational speed, which is critical for identifying fault characteristics, was obtained from the current spectrum based on the deep learning method, and the fault components were extracted from the spectrum to indicate various faults. Finally, maintenance decisions were made based on the detected failure characteristics to prevent further damage and improve system reliability. Fusing MCSA with deep learning techniques significantly improved fault diagnosis accuracy compared with MCSA alone, demonstrating the robustness and effectiveness of the proposed method.

## Introduction

Artificial lift techniques are critical for sustaining global oilfield operations, playing a pivotal role in prolonging well life and optimizing oil recovery. Among these techniques, the ESP is particularly prominent, with its adoption in approximately 15–20% of the world's one million wells (Pham et al. 2021). This widespread use is attributed to the ESP's capacity to maintain high production levels in deep well environments. Consequently, ESP-operated wells account for nearly 60% of the global annual crude oil output, while the associated costs represent 43% of the total global expenditure on artificial lift technologies (Yang et al. 2022).

The ESP system is a complex, multicomponent system, where the failure of any single component can lead to the malfunction of the entire system. When an ESP system fails, it can have a significant impact on oilfield production. In addition to the substantial costs associated with workover operations, there is also a loss of oil production. Failures in ESP systems can be broadly categorized into three primary types: electrical, mechanical, and operational. Electrical failures encompass issues such as cable defects, motor malfunctions, and overload conditions. Mechanical failures involve component damage, corrosion, and leaks, while operational failures pertain to challenges such as multiphase flow and scale deposition (Fakher et al. 2021). Among these, electrical failures are the most prevalent and have the most severe impact, as they can result in the total failure of the ESP system due to their potential to disrupt the entire operational integrity of the system. Motor malfunctions are of particular concern in electrical failures. When a SEM operates under abnormal conditions, such as eccentricity or broken rotor bar, it can lead to overloading of the motor. This overload condition increases the current load on both the motor and the cables, which may result in the degradation or damage of cable insulation, ultimately leading to cable failures. Therefore, to extend the service life of ESP systems and enhance the management level of oil fields, it is crucial and necessary to perform timely and accurate fault diagnostics on SEM.

Data analysis forms the cornerstone of fault diagnosis in ESP systems. With the rapid advancement of ESP data acquisition systems, there is increasing emphasis on the real-time monitoring and analysis of ESP production systems (AlBallam et al. 2022). Sensors and instruments installed in ESP wells gather extensive information from both downhole and surface environments, allowing for the periodic, real-time recording of ESP production parameters at a defined frequency. The data captured by sensors within the ESP system predominantly include dynamic, static, and historical data (Alamu et al. 2020). Dynamic data encompass surface temperature and pressure, pump inlet pressure and temperature, pump discharge pressure and temperature, pumphead, motor current, voltage, leakage current, and pump vibration. Static data involve the design parameters of the ESP, reservoir physical properties, and the completion methods of the ESP well, while historical data primarily include records of past well failures, workover times, and fault reports.

Among these data sets, vibration data provide the most direct indication of motor fault conditions. Rotor eccentricity or misalignment results in uneven forces during rotation, leading to significant vibration. Bearing wear or damage causes rotor imbalance, which increases the vibration amplitude. Furthermore, loosened internal mechanical components induce additional vibrations during operation. Electrical issues, such as winding short circuits or stator faults, also contribute to abnormal vibrations (Popaleny et al. 2018; Rodrigues et al. 2023). Even though ESP vibrations are tracked, they are typically treated as a static parameter, which only signals the potential onset of a problem when vibration levels surpass the alarm threshold. Determining the root cause, however, requires a more advanced vibration diagnostic system capable of displaying dynamic data. Implementing vibration monitoring is challenging due to the extreme depths, temperatures,

\*Corresponding author; email: luxincup@outlook.com

Copyright © 2025 Society of Petroleum Engineers

Original SPE manuscript received for review 14 November 2024. Revised manuscript received for review 10 January 2025. Paper (SPE 225426) peer approved 6 February 2025.

and pressures involved. Additionally, selecting the ideal position for mounting a vibration sensor, such as an accelerometer, is problematic. The recorded vibration is highly dependent on the placement of the accelerometer and the mechanical stiffness at different points along the ESP casing, which can vary significantly (Takacs and Takacs 2014; Thomson and Culbert 2016). Therefore, MCSA technology is a suitable alternative.

MCSA is a method used to detect faults in ESPs and SEMs by examining the dynamic currents of ESPs during operation. Compared with vibration analysis, current analysis is significantly simpler to perform, as it does not require the installation of sensors downhole, and all measurements can be made at the surface in the electric drive and pump-control equipment. The MCSA method leverages the FFT to convert time-domain current data into the frequency domain. Various faults can be identified by analyzing abnormal line frequencies in the frequency domain. Many studies have demonstrated the practical application of MCSA in diagnosing ESP faults and optimizing operations. For instance, field trials have shown that MCSA successfully identified bending loads and vibrations in high dogleg severity wells, where ESPs are prone to early failures (Walters et al. 2023). Additionally, controlled experiments comparing MCSA with traditional vibration analysis indicated that MCSA provided more accurate early fault detection, particularly in identifying subtle mechanical and electrical anomalies (Badkoubeh and Trevisan 2021). Extended field trials in a SAGD field further validated the use of MCSA to monitor and diagnose issues related to power quality, flow regime, and mechanical vibrations, demonstrating its ability to improve ESP reliability and optimize production performance (Badkoubeh et al. 2023).

MCSA has been widely applied in motor fault diagnostics; however, it still presents certain limitations. Primarily, MCSA relies on frequency-domain analysis, detecting faults by identifying characteristic frequencies within the current signal. This approach exhibits reduced sensitivity to complex faults, such as multiple simultaneous faults or weak fault signals, making accurate identification challenging. Moreover, environmental noise and nonfault-related signal disturbances can interfere with analysis results, leading to false positives or missed detections. Additionally, MCSA requires significant expertise, demanding that analysts possess deep theoretical knowledge of motors and substantial practical experience (Villalobos-Pina et al. 2024). Additionally, the MCSA technique requires motor speed data for accurate diagnostics (Thomson and Gilmore 2003). However, in downhole environments, measuring the speed of SEMs poses significant challenges.

To address these limitations, integrating deep learning methods can enhance fault detection and diagnosis. Deep learning models can learn complex patterns in the current data that might be difficult to identify with traditional methods. By training these models on large data sets, they can differentiate between subtle variations in the current signature that correspond to different types of faults, improving accuracy. Furthermore, deep learning can help in denoising the data, allowing for more reliable fault detection even in the presence of significant environmental noise. More importantly, the deep learning approach makes it possible to calculate the motor speed without installing a speed sensor. This study employs deep learning techniques to estimate motor speed without the need for speed sensors, thereby enabling the calculation of fault characteristics in SEM. By integrating MCSA with deep learning methods, the approach enhances the diagnosis of SEM faults.

## Models and Methods

**Motor Current Signal Analysis.** MCSA is a noninvasive method for detecting and diagnosing motor faults. It is based on analyzing the spectral characteristics of stator currents during motor operation to identify potential mechanical or electrical faults. The MCSA technique is widely used for motor condition monitoring and fault diagnosis due to its high efficiency, reliability, and cost-effectiveness. It is particularly important for the maintenance of rotating equipment in industry.

As an integral component of the ESP system, any malfunction of the SEM can result in ESP system failure, production stoppage, and potential safety hazards. Therefore, it is crucial to continuously monitor the motor's status and conduct proactive maintenance at the initial signs of a malfunction. While traditional motor fault detection methods, such as vibration analysis, require costly equipment and complex procedures, MCSA technology can detect early faults by analyzing the frequency spectrum of the motor stator current without disrupting the motor's regular operation. This approach effectively reduces downtime and maintenance expenses. The core of MCSA technology lies in the acquisition of motor stator current signals and the use of spectral analysis methods to identify the characteristic frequency components in the current signals. These frequencies are closely related to the internal faults of the motor, such as rotor broken bars, bearing faults, air gap eccentricity, and so on. The implementation process of MCSA is shown in Fig. 1.

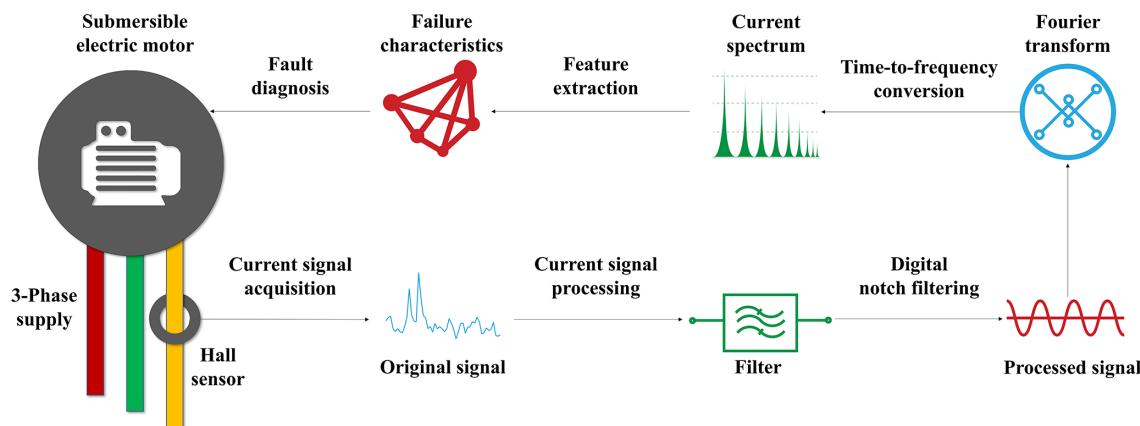


Fig. 1—Implementation process of MCSA technology.

**Step 1: Current Signal Acquisition.** First, the stator current signal of the motor in operation is collected in real time utilizing current sensors (e.g., Hall sensors or current transformers) installed in the electric drive and pump-control equipment. According to the Nyquist sampling theorem (Nyquist 1928), to accurately reconstruct the signal, the sampling frequency should be at least twice the highest frequency component of the signal. Therefore, the sampling frequency needs to satisfy the following conditions:

$$f_s \geq 2 \times f_{\max}, \quad (1)$$

where  $f_s$  is the sampling frequency and  $f_{\max}$  is the highest frequency component in the signal. In practice, the sampling frequency is generally much higher than the theoretical minimum requirement. This accounts for the high-frequency components that may be introduced by SEM faults and the possible aliasing effects during sampling. The power supply frequency of the SEM is usually 50 or 60 Hz, but the sampling frequency needs to be at least 1 kHz. For low-frequency faults, such as rotor bar breaks, the sampling frequency is typically at least 1 kHz to capture the low-frequency components associated with rotational speed. For high-frequency faults, such as bearing failures, the sampling frequency should be at least 10 kHz or higher to ensure that high-frequency vibration or resonance frequencies can be captured. At the same time, a higher sampling frequency generates more data, which places higher demands on data storage and subsequent processing capabilities. In oilfield applications, a balance needs to be found between sampling frequency and data processing capability.

**Step 2: Signal Processing.** Current signals acquired during analysis often contain significant amounts of noise and other uncorrelated frequency components, which can obscure the detection of fault characteristics. Therefore, preprocessing is essential to enhance the signal's clarity and relevance. This typically involves techniques such as denoising and filtering, which are applied to isolate the relevant information from the noise and improve the accuracy of subsequent fault analysis. In this study, a second-order digital notch filter was used to process the acquired current signals, thereby mitigating interference from power supply frequency. A second-order notch filter is a type of digital filter commonly used for the attenuation of specific frequency components. It exhibits good frequency selectivity by creating a narrow band of attenuation at a designated frequency, effectively suppressing unwanted spectral content. The transfer function of the second-order digital notch filter can be expressed as follows:

$$H(z) = \frac{1 - 2r\cos(2\pi f_0 T)Z^{-1} + r^2 Z^2}{1 - 2\cos(2\pi f_0 T)z^{-1} + Z^{-2}}. \quad (2)$$

The variable  $Z$  in the formula represents the  $z$ -transform variable, which serves as the complex frequency-domain representation of a discrete-time signal. Here,  $f_0$  denotes the frequency targeted for suppression. The sampling period  $T$  is defined as the inverse of the sampling frequency  $f_s$ , that is,  $T = 1/f_s$ . The radius parameter  $r$  of the filter controls the depth and bandwidth of the notch, typically taking values close to 1 (e.g., 0.99) to create a deep notch at the specified frequency. After the original signal  $x(n)$  is filtered,  $x_{\text{filtered}}(n)$  is obtained, and further windowing of  $x_{\text{filtered}}(n)$  is required. Windowing is primarily applied to signals to mitigate spectral leakage during Fourier analysis. Spectral leakage arises when a signal is either truncated or not perfectly periodic within the observation window, leading to the dispersion of its energy across multiple frequencies in the frequency domain. This dispersion can introduce inaccuracies in the identification of the true frequency components of the signal. By applying a window function, the signal's edges are tapered to zero, thereby reducing the abrupt discontinuities that result from truncation. This smoothing process minimizes the artificial spreading of spectral energy, yielding a more accurate representation of the signal's frequency content. In this study, the signal is processed using the Hanning window function to reduce the spectral leakage in signal processing. The Hanning window function can be expressed as follows:

$$w(n) = 0.51 \left( 1 - \cos\left(\frac{2\pi n}{N-1}\right) \right), \quad (3)$$

where  $N$  is the length of the window (i.e., the number of sampling points of the signal) and  $n$  is the discrete time index,  $n = 0, 1, 2, \dots, N-1$ . Applying the Hanning window function to  $x_{\text{filtered}}(n)$  gives the windowed signal  $x_w(n)$ :

$$x_w(n) = w(n)x_{\text{filtered}}. \quad (4)$$

**Step 3: FFT.** In SEM fault diagnosis, the FFT plays a crucial role. By converting time-domain signals into frequency-domain signals, the FFT rapidly reveals the frequency components contained within the signals. This is particularly important for identifying fault characteristics related to frequencies, such as imbalance in rotating machinery, bearing defects, and short circuits in motor stator windings. The FFT is an efficient algorithm for computing the discrete Fourier transform (DFT). The computational formula for the discrete Fourier transform is given by

$$x(k) = \sum_{n=0}^{N-1} x_{\text{filtered}}(n) e^{-j \frac{2\pi}{N} kn}, \quad (5)$$

where  $x(k)$  is the  $k$ th frequency component of the FFT,  $x_{\text{filtered}}(n)$  is the  $n$ th sample of the preprocessed current signal,  $j$  is the imaginary unit,  $N$  is the number of sampling points of the current signal, and  $K$  ranges from 0 to  $N-1$ .

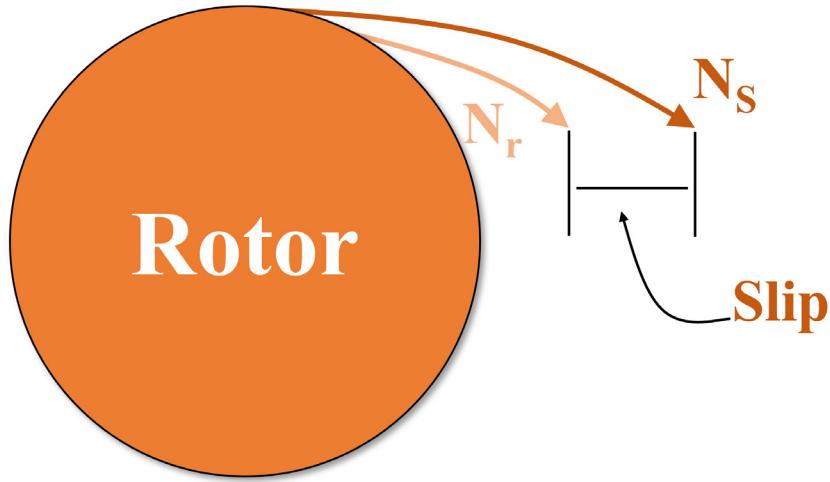
**Step 4: Feature Extraction.** Following the application of the FFT, the resulting current spectrum reveals key fault features. These features are then meticulously analyzed to diagnose motor malfunctions and to categorize the specific types of faults present. The cornerstone of extracting fault features lies in computing the slip, a critical parameter that describes the operational state of an induction motor. The slip indicates the extent of the difference between the rotational speed of the motor rotor and the synchronous speed, and a graphical representation of the slip is shown in Fig. 2.

Slip can be defined as follows:

$$s = \frac{N_s - N_r}{N_s}, \quad (6)$$

where  $N_s$  is the synchronous speed of SEM,  $N_r$  is the rotational speed, and they are usually expressed in revolutions per minute (RPM). The synchronous speed  $N_s$  is determined by the frequency of the power supply and the number of pole pairs of the motor, which can be calculated by the following equation:

$$N_s = \frac{120 \times f}{p}, \quad (7)$$

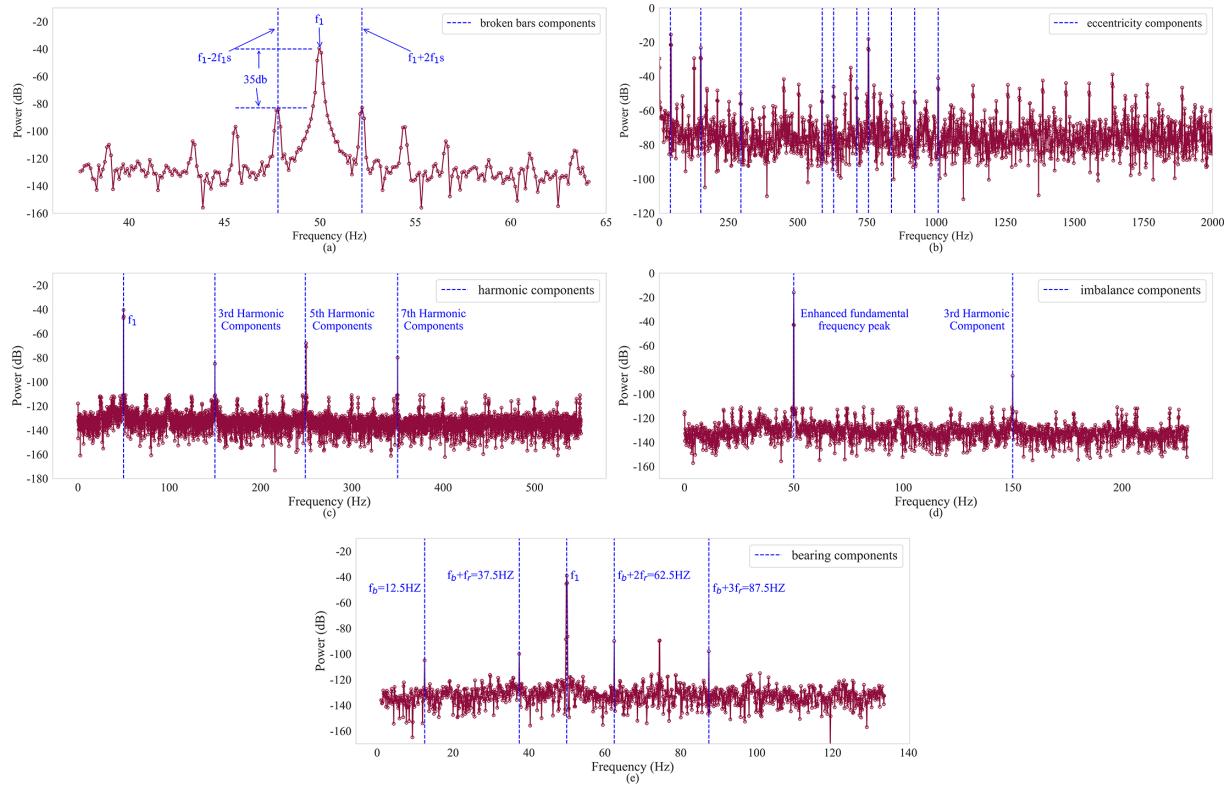


**Fig. 2—Illustration of slip.**

where  $f$  is the power supply frequency in Hz and  $p$  is the pole pair number of the motor. For the most common SEM fault, rotor broken bars, the fault characteristics can be expressed as follows:

$$f_{bb} = f_1 \pm 2f_1 s. \quad (8)$$

When a rotor develops broken bars, it results in an uneven distribution of the magnetic field inside the motor, which in turn generates sideband frequencies around the fundamental frequency in the spectrum. By detecting these frequencies and calculating their amplitude, it is possible to determine the presence and severity of a rotor broken bar fault. In the above equation,  $f_{bb}$  is the sideband frequency,  $f_1$  is the fundamental frequency, which is the power supply frequency, and  $s$  is the slip mentioned above. Observe the amplitude changes of these frequency sidebands, as a significant increase may indicate a broken bar fault. The characteristics of the rotor broken bars in the current spectrum are shown in **Fig. 3a**.



**Fig. 3—Spectral characteristics of SEM faults: (a) broken bars, (b) eccentricity, (c) harmonics, (d) imbalance, and (e) bearing.**

Another common fault that can fail the SEM is air gap eccentricity, and the fault characteristics of air gap eccentricity can be represented as follows:

$$f_{\text{ecc}} = f_1 \pm n \cdot N_r. \quad (9)$$

The uneven distribution of the air gap between the stator and rotor of a motor leads to changes in the magnetic flux density, which generates periodic electromagnetic forces during the operation of the motor. In the current spectrum, this manifests itself as an increase in the amplitude of a specific frequency component. The equation above shows how to calculate these specific frequency components, where  $f_{\text{ecc}}$  is the frequency component associated with air gap eccentricity,  $f_1$  is the supply frequency,  $n$  is a positive integer indicating the order of the fault frequency components, and  $N_r$  is the rotational speed of SEM in RPM. The characteristics of the air gap eccentricity in the current spectrum are shown in **Fig. 3b**.

After air gap eccentricity, another common fault in the SEM is harmonic distortion. Harmonics are unwanted frequency components that appear in the motor current spectrum, often caused by nonlinearities in the motor or external factors such as power supply issues. These harmonic components are multiples of the supply frequency and can indicate problems such as mechanical imbalances or electrical faults in the system. The equation for calculating harmonic components is as follows:

$$f_{\text{harm}} = m \cdot f_1, \quad (10)$$

where  $f_{\text{harm}}$  is the harmonic frequency,  $m$  is the harmonic order such as 2, 3, 4, 5, and so on. While harmonics of all orders can cause issues, odd-order harmonics are often more harmful.  $f_1$  represents the supply frequency. The characteristics of the harmonic distortion in the current spectrum are shown in **Fig. 3c**.

Three-phase imbalance is another critical issue in the SEM, often caused by asymmetrical loads, power supply issues, or internal faults. It introduces distinctive features in the motor current spectrum, such as zero-sequence components and third harmonic components. Zero-sequence currents occur when the sum of three-phase currents is nonzero, leading to increased amplitude at the fundamental supply frequency. The third harmonic, calculated using Eq. 10, appears at three times the supply frequency and is amplified under severe imbalance. These spectral characteristics, including enhanced fundamental and third harmonic amplitudes, are key indicators for diagnosing imbalance-related faults and are illustrated in **Fig. 3d**.

Bearing faults are one of the most common issues in rotating machinery and can be identified by analyzing specific frequency components in the spectrum. The characteristic fault frequencies depend on the geometry of the bearing and the rotational speed of the motor. The fundamental defect frequency can be calculated using the following equation:

$$f_b = \frac{N_r}{60} \cdot \frac{z}{d} \cdot \frac{N}{2}, \quad (11)$$

where  $f_b$  is the fundamental defect frequency,  $N_r$  is the rotational speed of the motor in RPM,  $z$  is the number of rolling elements,  $d$  is the diameter of the rolling elements, and  $N$  is the number of outer rings in the bearing. In addition, bearing faults cause sidebands near the fundamental frequency. Sidebands are caused by vibrations linked to variations in rotational speed or dynamic interactions between bearing elements. These frequencies are calculated as follows:

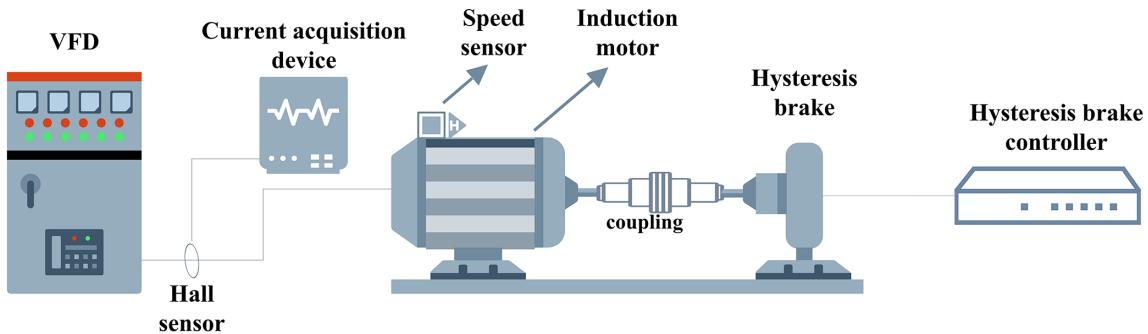
$$f_s = f_b \pm n \cdot f_r, \quad (12)$$

where  $f_s$  represents the sideband frequency,  $f_b$  is the fundamental defect frequency,  $f_r$  is the rotational frequency of the motor shaft, and  $n$  is the sideband order (e.g., 1, 2, and 3). The presence and amplitude of these sidebands serve as reliable indicators for diagnosing and monitoring the health of the bearing. The spectral characteristics of bearing faults, including fundamental frequencies and sidebands, are shown in **Fig. 3e**.

**Step 5: Fault Diagnosis.** The fault frequency components calculated in Step 4 are compared with the actual measured frequency components in the current spectrum to perform fault diagnosis. The signal amplitude of fault frequency components is an important basis for determining the severity of the fault. If the amplitude of the sidebands or the eigenfrequency is large, it means that the fault is more serious. Also, by continuously monitoring the current spectrum, it can be observed whether the amplitude of the eigenfrequency increases over time. If the amplitude continues to increase, the fault is worsening. In addition to spectrum analysis, a comprehensive diagnosis can be made in conjunction with other operating parameters of the motor such as temperature, vibration, and speed. For example, rotor broken bars and air gap eccentricity can trigger abnormal mechanical vibration, and the fault signal in the current spectrum can be further verified through vibration analysis.

**Rotational Speed Calculation Model.** In the diagnosis of the SEM, accurate computation of fault characteristics is paramount. A critical parameter for such computation, regardless of the fault type, is the motor speed. However, obtaining the motor speed is particularly challenging, as submersible motors operate in downhole environments, where direct measurement is difficult. This presents a significant obstacle to effective fault diagnosis. In this study, we established a laboratory-based speed measurement platform, enabling simultaneous acquisition of motor current and real-time speed data. By applying deep learning techniques, a nonlinear mapping between the motor current and real-time speed was developed. This model allows for accurate speed estimation using only current data, eliminating the need for downhole speed sensors.

The speed measurement platform consists of a three-phase squirrel-cage induction motor, a variable-frequency drive (VFD), a hysteresis brake, a coupling, a speed sensor, and a high-speed current acquisition system. The motor, with a rated power of 1.5 kW, a rated voltage of 380 V, and a frequency of 50 Hz, operates at a nominal speed of 2,880 rev/min. The insulation class of the motor is F, and its protection level is IP55, ensuring reliable operation under standard laboratory conditions. The motor and the hysteresis brake are connected via the coupling and mounted together on a 3-cm-thick steel plate. The hysteresis brake, which operates based on the hysteresis effect, provides a stable load for the motor. The motor speed is controlled using the variable-frequency drive, and once the motor reaches a steady speed, both speed and current data are collected for 60 seconds. The speed range of the experiment is from 20 to 60 Hz, with 500 sets of experimental data collected in total. The schematic and physical structure of the measuring platform are shown in **Figs. 4 and 5**, respectively.



**Fig. 4—Schematic of the measurement platform.**



**Fig. 5—Physical structure of the measurement platform.**

In oilfield operations, there are numerous models of SEMs, each with varying specifications. This variation leads to significant differences in operating current between different motors. To enhance the generalization capability of the rotational speed calculation model across different models of SEMs, normalization of the high-frequency current data is required. In this study, the min-max normalization method was used to map the current data into the range of  $[-1, 1]$ . Given the positive and negative variations in AC current data, normalizing the data to the  $[-1, 1]$  range preserves these differences, thereby retaining more of the original data information. Moreover, balanced positive and negative input data during the training of deep learning models facilitates better gradient updates, which can help reduce the vanishing gradient problem in deep networks and accelerate convergence. Importantly, mapping the current signals to an asymmetric interval such as  $[0, 1]$ , which lacks distinction between positive and negative values, would cause severe distortions in subsequent Fourier transforms. The Fourier transform relies on the periodicity and symmetry of the signal; normalizing to the  $[0, 1]$  range would disrupt this symmetry, leading to spectral distortion and introducing extraneous frequency components, particularly increasing the DC offset (the zero-frequency component) significantly because all signals would become positive values. The calculation formula for min-max normalization is as follows:

$$x_{\text{norm}} = 2 \times \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (13)$$

Where  $x$  is the original data,  $x_{\min}$  is the minimum value in the data,  $x_{\max}$  is the maximum value in the data, and  $x_{\text{norm}}$  is the normalized data in the range  $[-1, 1]$ . After completing the normalization of the data, a Fourier transform is applied to the data to obtain the current spectrum. The current spectrum data will be used as input to the deep learning model to map the speed of the motor. The current spectrum data will be used as input for the deep learning model, mapping to the rotational speed of the motor. The use of current frequency spectrum data as input for the rotational speed calculation model is justified by the fact that the motor's rotational speed directly influences its operating conditions and current variations. Eccentricity in the air gap, caused by manufacturing tolerances, introduces characteristic frequency components into the current signal, which are related to the rotational speed. Variations in rotational speed result in shifts of specific frequency components within the current spectrum. Consequently, the current frequency spectrum can carry rich information related to rotational speed. The relationship between the current frequency spectrum and rotational speed is likely to be highly nonlinear, and traditional analytical models may struggle to describe it accurately. In contrast, deep learning models can automatically learn this

relationship through training data. However, using current frequency spectrum data as input for deep learning models may result in issues related to input complexity.

The current data were collected at a frequency of 1 kHz over a period of 60 seconds, resulting in each data sample consisting of 60,000 data points. After performing the Fourier transform, the current spectrum also contains 60,000 frequency components. As the spectrum is symmetric after Fourier transformation, typically only the positive frequency part, known as the half-spectrum, is considered. Thus, the number of frequency components in the half-spectrum is 30,000. Even utilizing only the half-spectrum as input for the model remains overly complex. When deep learning models are fed with excessively complex inputs, several challenges can emerge. First, the model may face overfitting, leading to memorization of training data instead of learning general patterns and resulting in poor performance on new data. This issue is exacerbated by increased computational costs due to processing high-dimensional or intricate inputs, requiring more resources and longer training times. Training instability is also a concern, as intricate input features can cause problems like vanishing or exploding gradients. Additionally, including excessive or irrelevant information in the input can reduce the model's performance by introducing noise, ultimately affecting the accuracy and reliability of predictions. Therefore, effectively managing input complexity is crucial for ensuring the effectiveness and efficiency of deep learning models.

To tackle the complexity of the input data, this study has selected principal component analysis (PCA) as the dimensionality reduction technique for preprocessing before feeding it into the deep learning model. PCA is a widely-used statistical method that converts high-dimensional data into a lower-dimensional space by identifying the principal components—linear combinations of the original variables that capture the most variance within the data. This approach preserves the essential information in the data while decreasing the number of input features, thereby simplifying the model's input space without compromising on significant details. PCA is particularly advantageous because it reduces the risk of overfitting by decreasing the input dimensionality and focusing on the most informative aspects of the data. By compressing the data, PCA significantly decreases the computational burden on the model, thereby reducing training time and enhancing the efficiency of resource usage. This is especially crucial for high-dimensional data sets. Moreover, PCA provides a robust way to handle collinearity in the data, which can otherwise degrade the performance of the model. By projecting the data onto orthogonal principal components, the method ensures that the inputs are uncorrelated, further improving the stability and interpretability of the model.

PCA involves a series of mathematical steps to reduce the dimensionality of a data set while preserving as much variance as possible. The process begins with centering the data by subtracting the mean of each feature from the corresponding data points, ensuring that the data set has a mean of zero. This step is crucial because PCA is sensitive to the scale of the data, and centering ensures that the principal components are not skewed by differences in magnitude across features. Next, the covariance matrix of the centered data is calculated. The covariance matrix captures the relationships between features and measures how changes in one feature correlate with changes in another. For a data set with  $n$  features, the covariance matrix is an  $n \times n$  matrix, where each element represents the covariance between a pair of features. The formula for the covariance between two features  $x_i$  and  $x_j$ :

$$\text{Cov}(x_i, x_j) = \frac{1}{N} \sum_{k=1}^N (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j), \quad (14)$$

where  $N$  is the number of data points,  $x_{ki}$  and  $x_{kj}$  are the values of the  $i$ th and  $j$ th features for the  $k$ th data point, and  $\bar{x}_i$  and  $\bar{x}_j$  are the means of the respective features. Following the computation of the covariance matrix, the subsequent step involves the determination of its eigenvalues and corresponding eigenvectors. The eigenvectors identify the directions of the principal components, whereas the eigenvalues quantify the proportion of total variance explained by each respective principal component. This process is achieved through the solution of the eigenvalue equation:

$$Cv = \lambda v, \quad (15)$$

where  $C$  denotes the covariance matrix,  $\lambda$  represents the eigenvalue, and  $v$  is the corresponding eigenvector. The eigenvectors are subsequently ordered based on their corresponding eigenvalues, with the eigenvector associated with the largest eigenvalue indicating the direction of the greatest variance in the data. This ordering ensures that the first principal component captures the most significant amount of variance, followed by the second principal component, and so on. Finally, the data are projected onto the selected principal components by multiplying the original data set with the eigenvectors corresponding to the largest eigenvalues. This projection reduces the dimensionality of the data while retaining the most important variance. The new data set in the reduced space is represented as follows:

$$Z = Xv, \quad (16)$$

where  $Z$  is the transformed data,  $X$  is the original data matrix, and  $V$  is the matrix of eigenvectors. **Fig. 6** represents the dimensionality reduction process of PCA in two dimensions.

The spectrum data were downsampled to 50 dimensions, as the first 50 principal components accounted for over 80% of the total variance. This indicates that they significantly contribute to the representation of the data and the capture of the fundamental features of the original data set, thereby justifying the reduction in the number of principal components. The contribution of the principal components is illustrated in **Fig. 7**.

To extract the variation information of rotational speed, the original 60-second current signal is segmented into multiple overlapping time windows. Each window is set to a length of 50 seconds with a step size of 1 second, allowing for the extraction of 11 data segments. For each 50-second time window, FFT is applied to convert the signal from the time domain to the frequency domain, and PCA is utilized to reduce the dimensionality to 50 dimensions. Following this processing, 11 time windows are obtained, each corresponding to a 50D feature vector. Ultimately, these data are organized into a 3D array with a shape of  $11 \times 50 \times 1$ , where the first dimension represents the number of time windows, the second dimension indicates the feature dimensions of each time window, and the third dimension corresponds to the single-channel data. The data processing procedure is illustrated in **Fig. 8**.

After the dimensionality of the spectrum is reduced by PCA, the low-dimensional principal components are fed into a deep learning model for rotational speed calculation. Convolutional neural networks (CNNs) were chosen to compute the rotational speed in this study. CNNs were originally designed for image recognition tasks, but they perform equally well in processing time series data and frequency domain signals, especially in capturing local features and patterns. CNNs are able to automatically learn and extract useful features from raw input data without the need to manually design feature extraction algorithms, which is especially beneficial for complex spectral data. The design of the convolutional layers makes the model robust to shifts in the input data. This means that the CNN can effectively recognize the same frequency features even if they appear at different locations in the spectrum, which is useful for analyzing the current

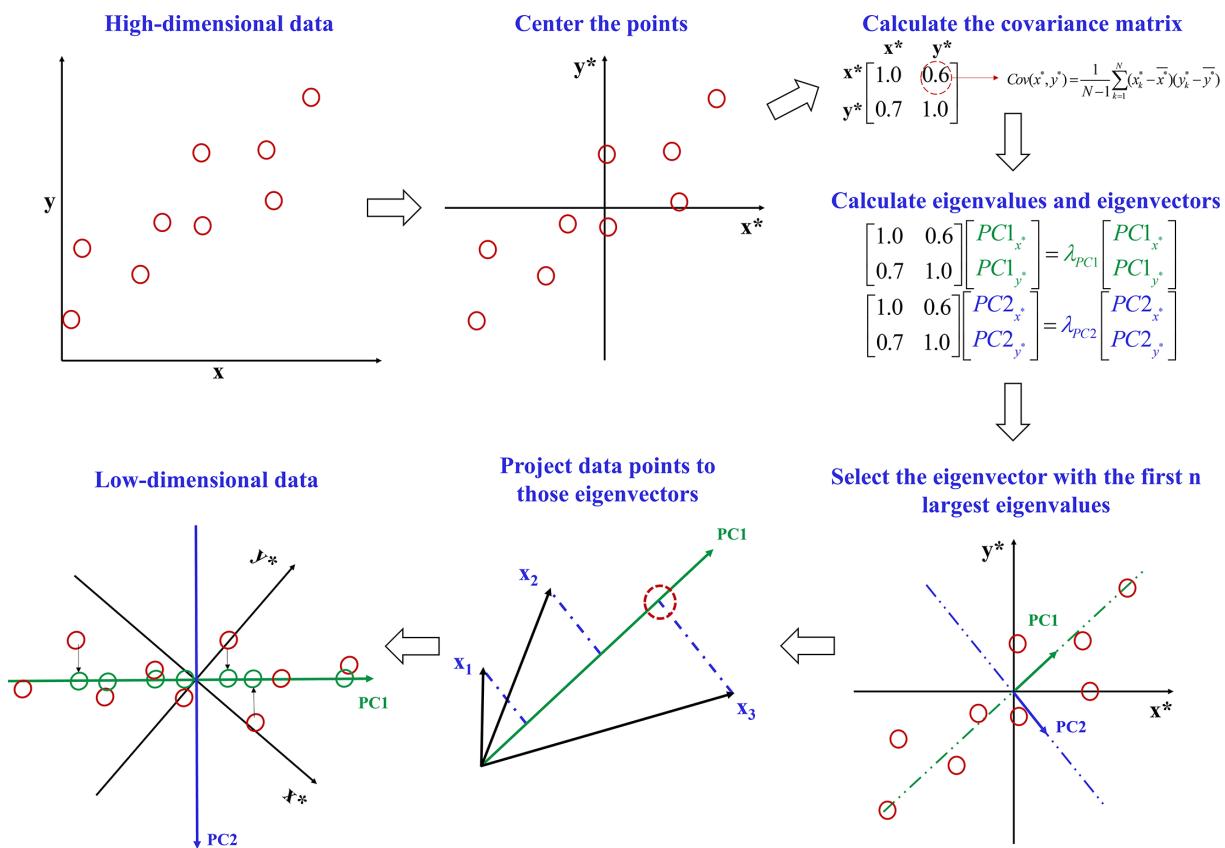


Fig. 6—The dimensionality reduction process of PCA.

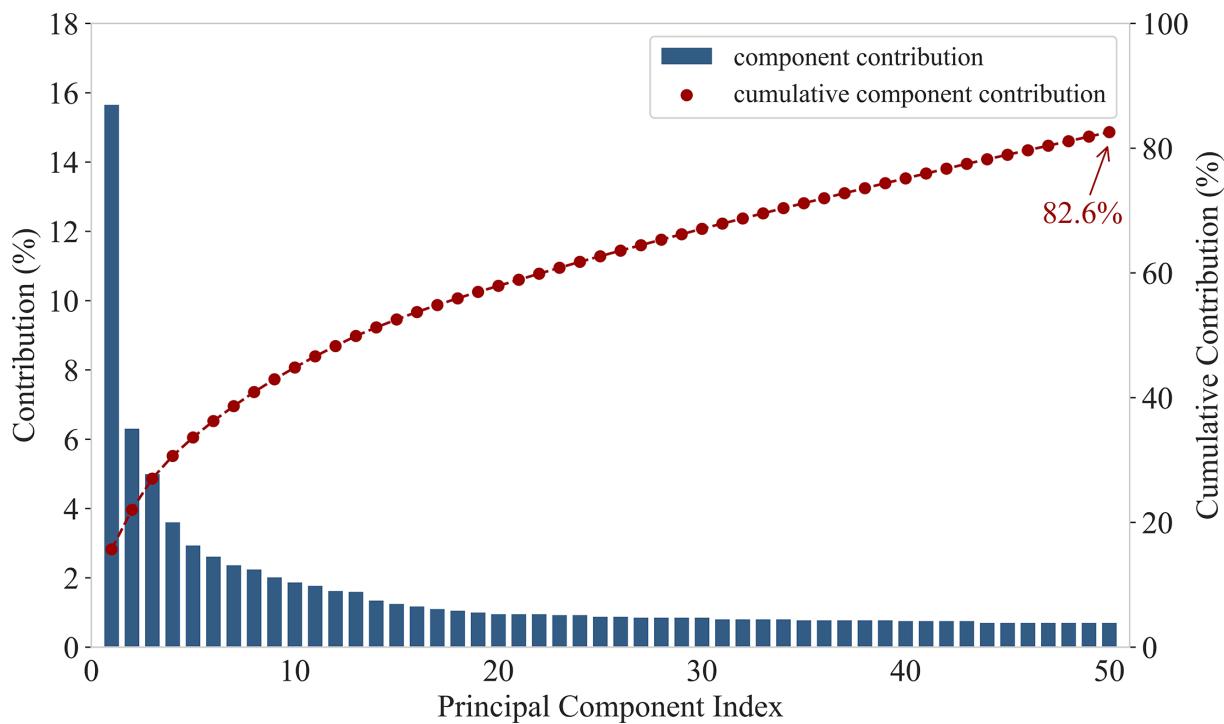
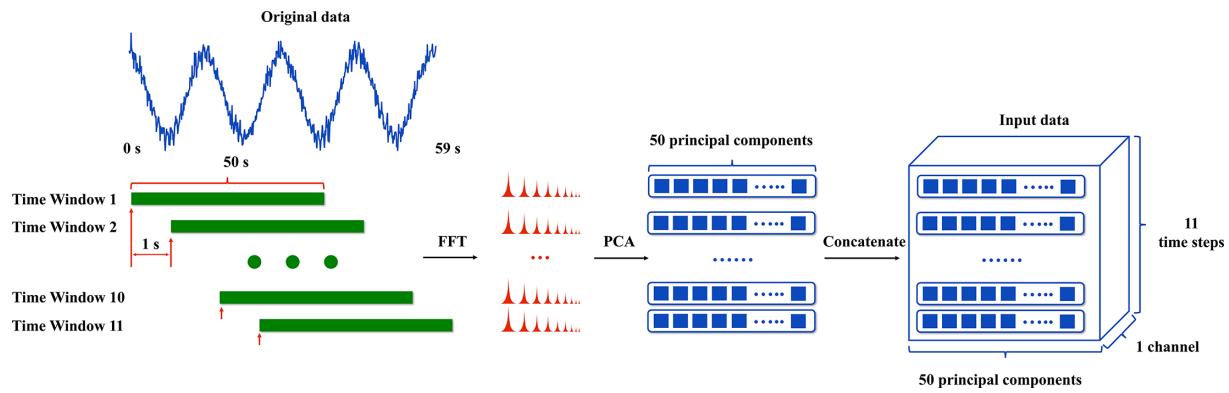


Fig. 7—Principal component contribution and cumulative contribution.

spectrum collected under different states. The CNN architecture consists of convolutional, pooling, flattening, and fully connected layers. Convolutional layers are responsible for extracting features from the input data by applying filters to generate feature maps. Pooling layers, typically using operations like max-pooling, reduce the dimensions of these feature maps, which helps to lower computational complexity and control overfitting. The flattening layer converts the multidimensional feature maps into a 1D vector, making it

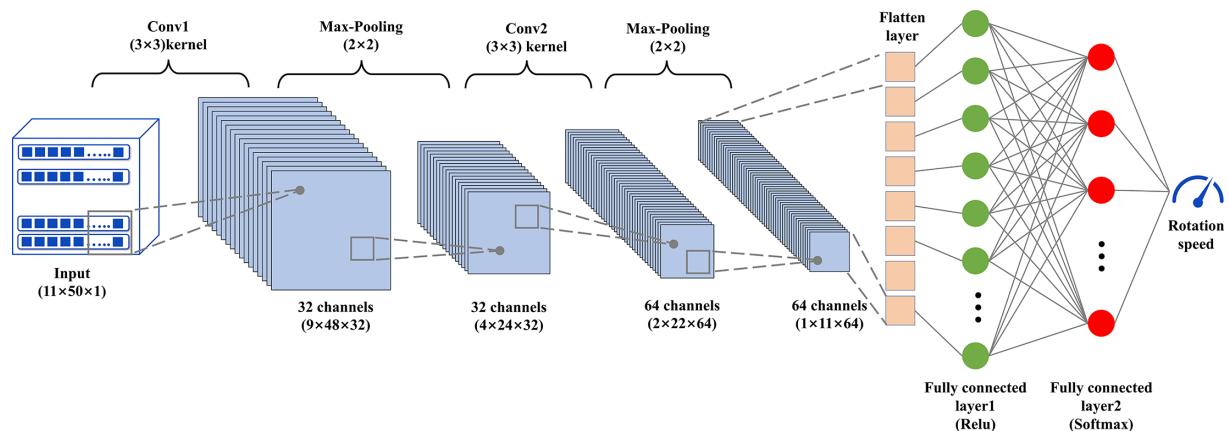


**Fig. 8—The data processing procedure.**

compatible for input to the fully connected layers. Fully connected layers then combine the extracted local features into a more global representation and ultimately map these features into the final output space, typically for tasks like classification or regression. The core of the CNN is the convolutional layer, and the mathematical expression for the convolutional operation can be simplified as follows:

$$Z = X * W + b, \quad (17)$$

where  $X$  is the input matrix,  $W$  is the convolution kernel, essentially a weight matrix,  $b$  is the bias term, and  $*$  represents the convolution operation. The convolution kernel slides over the input matrix, applying element-wise matrix multiplication and summing the results to produce the feature map  $Z$ . The CNN is utilized to construct the rotational speed calculation network, with its architecture detailed in Fig. 9.



**Fig. 9—The structure of the rotational speed calculation model.**

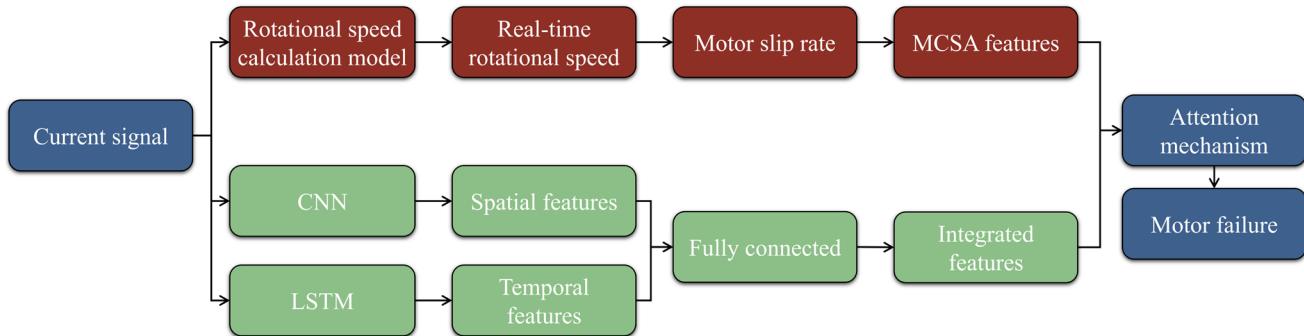
The network consists of nine layers in total. Input represents the input layer, where the preprocessed input data has dimensions of  $11 \times 50 \times 1$ . The first convolutional layer, Conv1, applies 32 convolutional filters, each with a size of  $3 \times 3$ . After passing through Conv1, the input data are convolved, resulting in a feature map with dimensions of  $9 \times 48 \times 32$ . This feature map is then passed through a  $2 \times 2$  max-pooling layer, which reduces the dimensions by half, resulting in  $4 \times 24 \times 32$ . As the size of the data in the first dimension is 9, the last row is discarded to ensure an even pooling operation, resulting in the first dimension being reduced to 4 after pooling. The second convolutional layer, Conv2, has 64 filters, each with a size of  $3 \times 3$ . After convolution through Conv2, the feature map dimensions become  $2 \times 22 \times 64$ . A subsequent max-pooling operation reduces the dimensions to  $1 \times 11 \times 64$ . The resulting feature map is then flattened into a 1D vector with a size of 704, which is passed to the fully connected layers for further computation. Fully Connected Layer 1 has 128 neurons with a rectified linear unit activation function. The second fully connected layer, FC2, has 64 neurons and uses the softmax activation function to map the neurons to the output layer, which predicts the rotational speed.

Applying this model to downhole ESP systems introduces challenges due to long-distance signal transmission. Signal attenuation, noise, and high-frequency loss from cable impedance can distort the current signal, affecting the accuracy of rotational speed calculations. These issues can be mitigated through surface-based processing methods, such as filtering or signal reconstruction, to correct distortions. Increasing the sampling rate can also enhance the resolution of high-frequency components, ensuring reliable feature extraction. For high-speed motors like permanent magnet motors operating above 5,000 rev/min, the higher characteristic frequencies in the spectrum require a sampling rate of 20–50 kHz to maintain sufficient resolution. These adjustments make the model adaptable to future motor technologies and field conditions.

**Fault Diagnosis Model.** Despite the effectiveness of individual methods such as CNNs, recurrent neural networks (RNNs), and analytical approaches like MCSA in fault diagnosis, each approach has inherent limitations. CNNs are adept at capturing spatial features but may overlook important temporal patterns in the data. In contrast, RNNs excel at modeling temporal dependencies but lack the ability to effectively extract spatial features. Similarly, MCSA is powerful in identifying fault signatures from motor current spectra but struggles to

fully integrate spatial and temporal interactions within the data. These limitations highlight the need for a more comprehensive approach that combines the strengths of each method while mitigating their weaknesses. To address this, the fault diagnosis model proposed in this study integrates CNNs, LSTM, MCSA, and an attention mechanism to provide a more robust solution for fault diagnosis, enhancing accuracy and reliability in machinery monitoring.

The fault diagnosis model integrates MCSA, CNN, long short-term memory (LSTM), a type of RNN, and an attention mechanism to perform accurate fault diagnosis. The model takes the current signal as an input signal, these signals are processed through parallel branches, one branch employs CNNs to capture spatial features by detecting local patterns within the signals, while the other branch utilizes LSTMs to extract temporal features that represent the sequence dependencies over time. The spatial and temporal features extracted from the CNN and LSTM branches are then processed through the fully connected layer to form integrated features, which enhances the model's ability to recognize complex relationships within the data. Simultaneously, the input signals undergo FFT to convert the time-domain signals into frequency spectra. These spectra are crucial as they serve as inputs for a rotational speed calculation model, which computes the rotational speed of the SEM. Further analysis of the calculated rotational speed is conducted to identify the key parameter used in MCSA for determining fault components, which is slip. Spatiotemporal fusion features and fault components identified by MCSA are subsequently merged via an attention mechanism, which allows the model to dynamically weight and integrate these diverse feature sets based on their diagnostic importance. Finally, the combined features are passed through fully connected layers to perform the final task of fault diagnosis, leveraging the enriched feature representation to accurately classify the operational state of the monitored equipment. The structure of the fault diagnosis model is illustrated in **Fig. 10**.



**Fig. 10—The structure of the fault diagnosis model.**

In this model, the LSTM plays a crucial role in capturing the temporal dependencies in the current signal. The LSTM is specifically designed to handle data sequences and can effectively learn patterns over time through a gating mechanism that regulates the flow of information. This mechanism consists of input, output, and forget gates, which allow the model to retain important features while discarding irrelevant information. By processing the input signal through its recurrent layers, the LSTM can recognize and model long-term dependencies, making it particularly well-suited for time-series data, such as the current signals in this application. The LSTM algorithm executes a defined series of operations at every discrete timestep, denoted as  $t$ . Here  $x_t$  signifies the input vector at time  $t$ ,  $h_t$  denotes the hidden state, and  $c_t$  represents the cell state. Additionally,  $f_t$  corresponds to the output of the forget gate,  $i_t$  to the output of the input gate, and  $o_t$  to the output of the output gate. The LSTM model's operations are regulated by the subsequent equations:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (18)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (19)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (20)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t, \quad (21)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (22)$$

$$h_t = o_t \cdot \tanh(c_t). \quad (23)$$

The structure of LSTM is shown in **Fig. 11**.

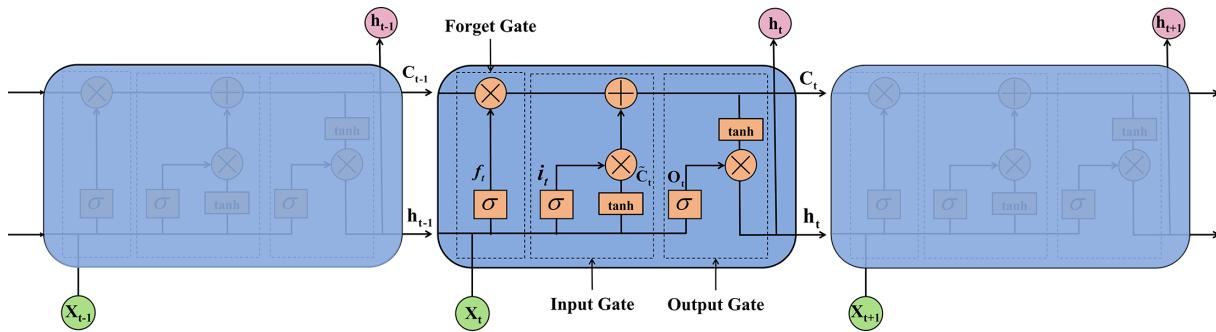
In the proposed model, an attention mechanism is used to enhance feature representation by assigning varying importance to the fused features derived from the CNN and LSTM outputs, as well as the MCSA features. The attention scores are computed using a feed-forward neural network, which takes the concatenated feature vectors  $F = [F_{\text{intergrade}}, F_{\text{MCSA}}]$  as input, the attention weights  $a$  are then obtained through a softmax function, expressed mathematically as

$$a = \text{Softmax}(W \cdot F + b), \quad (24)$$

where  $W$  represents the learned weight matrix and  $b$  is the bias term. The resulting attention weights are used to compute a weighted feature representation  $F_{\text{att}}$ :

$$F_{\text{att}} = \sum_{i=1}^n a_i \cdot F_i, \quad (25)$$

where  $a_i$  denotes the attention weight corresponding to feature  $F_i$ . This weighted representation is subsequently fed into a fully connected layer for fault diagnosis, allowing the model to focus on the most relevant features while diminishing the influence of less informative ones. The proposed fault diagnosis model is designed with specific architectural parameters to enhance its performance. The CNN branch includes two convolutional layers with 32 and 64 filters, each using a kernel size of  $5 \times 5$  and rectified linear unit activation. Each



**Fig. 11—The structure of the LSTM model.**

convolutional layer is followed by a max-pooling layer with a pool size of  $2 \times 2$ , effectively capturing spatial features from the input current signal. The LSTM branch contains a single layer with 128 units, designed to model long-term temporal dependencies in the time-series data. A dropout rate of 0.2 is applied to prevent overfitting. The spatial features extracted by the CNN and the temporal features from the LSTM are concatenated with fault signature features derived from MCSA. These combined features pass through an attention mechanism, consisting of a fully connected layer with 256 neurons and a softmax activation function to compute attention weights.

The proposed model shows strong scalability across different ESP models and operating conditions. This is due to three main factors: the use of universal input features, the robustness of the architecture, and the diversity of the training data set. The model processes raw current signals, which are available for all ESP models. Using FFT, it extracts frequency-domain features such as harmonics, slip, and fault components, which depend on motor dynamics and are not tied to specific motor designs. The architecture combines CNNs for spatial feature extraction, LSTMs for temporal patterns, and an attention mechanism for feature weighting. This design allows the model to handle variations in motor sizes, power ratings, and load conditions. The training data set includes motors with different specifications and covers various operating conditions like speed changes and load shifts. This ensures the model performs well with different ESP types and under diverse field conditions.

## Results and Discussion

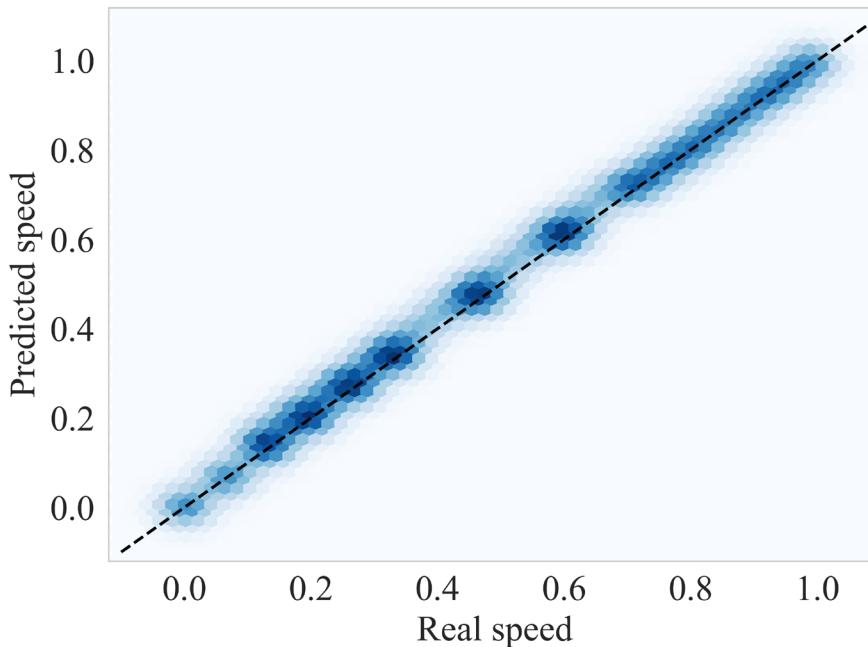
**Calculation of Rotational Speed.** The network was trained using 80% of the data and tested using 20% of the data. The error of the network was evaluated using the root mean square error (RMSE) and the mean relative error (MRE), demonstrated in **Table 1**.

Evaluation Metrics	Calculation Formula
RMSE	$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_t - y_{\text{pre}})^2}$
MAE	$\text{MAE} = \frac{\sum_{t=1}^n  y_t - y_{\text{pre}} }{n}$

**Table 1—Rotational speed calculation model evaluation metrics.**

The adaptive moment estimation (Adam) optimizer is utilized to facilitate the training process. Adam is a popular optimization algorithm that combines the advantages of both AdaGrad and RMSProp. It dynamically adjusts the learning rate for each parameter based on the first and second moments of the gradients, which helps in accelerating convergence and stabilizing the training process. The model was trained with a batch size of 32 and an initial learning rate of 0.0005. Training was conducted for a maximum of 200 epochs, with early stopping applied if the validation loss did not improve for 10 consecutive epochs. Following the pretraining of the rotational speed calculation model, the model exhibited an MRE of 0.07865 and an RMSE of 0.0459 on the training set. On the test set, the MRE was recorded at 0.07182 with an RMSE of 0.0476. These metrics indicate that the model achieved a high level of accuracy. The prediction results of the rotational speed calculation model on the entire data set are shown in **Fig. 12**. Due to the large amount of data, a data density distribution graph was used to represent the computational effect of the model. The horizontal axis represents the normalized real rotational speed, while the vertical axis corresponds to the normalized predicted rotational speed. The gradation of colors indicates the density of the data distribution. It is evident that the data predominantly cluster along the diagonal, signifying that the rotational speed predictions generated by the model are highly reliable. To validate the robustness of the model, the predictive performance of the model was evaluated under various operating conditions.

**Fig. 13** demonstrates that the model consistently achieves low MRE and RMSE across various operating speeds—high, medium, and low—highlighting minimal variability in prediction error levels among different speed settings. This consistent performance across a range of conditions underscores the model's robustness and confirms its capability to maintain accuracy and reliability under diverse operational scenarios. On the other hand, during the ESP production process, factors such as temperature fluctuations and mechanical vibrations can introduce noise into the current signal. Therefore, it is essential for the model to maintain robust performance even under noise interference. To assess the model's performance when input data are disturbed, Gaussian noise was deliberately introduced into the current data. **Fig. 14** illustrates the scenario in which noise is added to the original current signal. The blue curve represents the original current signal, while the red semitransparent curve depicts the current data with noise introduced.

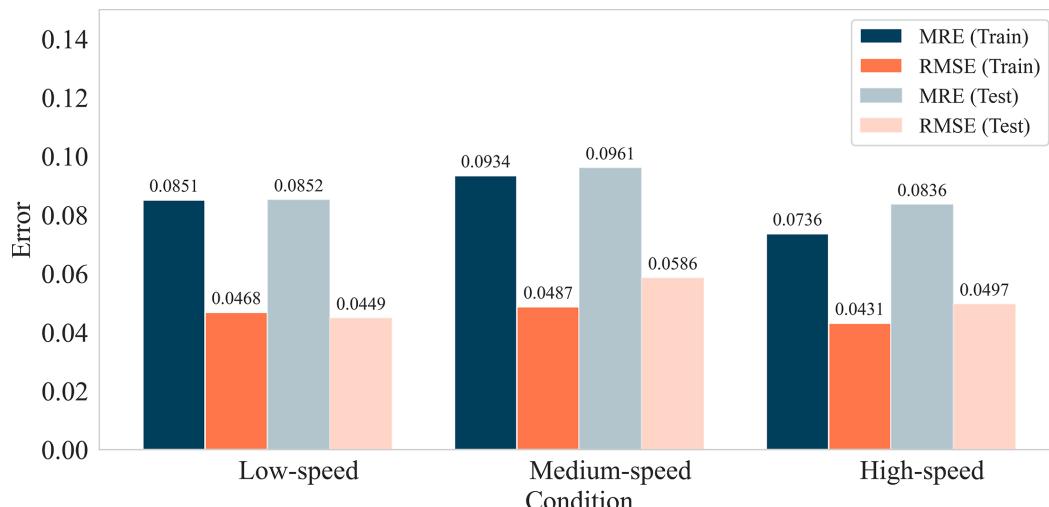


**Fig. 12—Rotational calculation model predictions for full sample.**

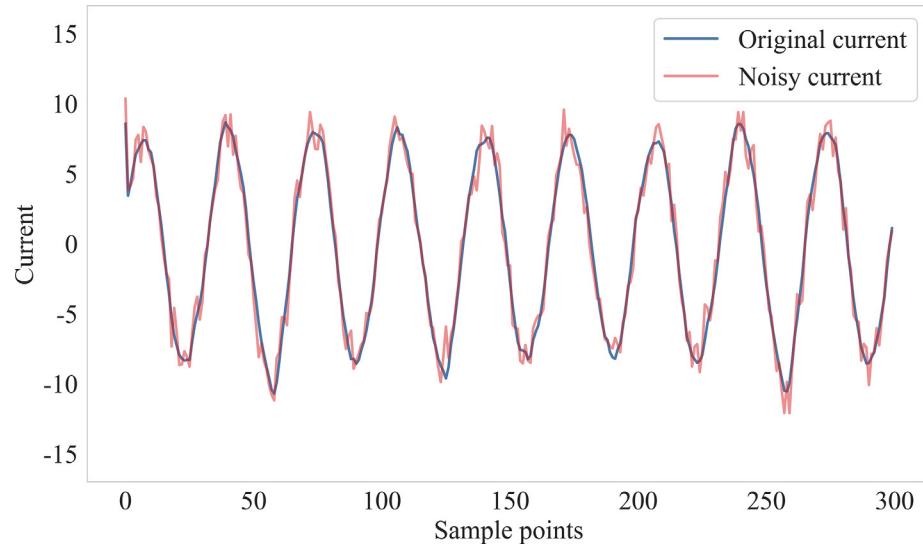
The calculations of rotational speed using the noisy current data yield an MRE of 0.08137 and an RMSE of 0.0461 for the training set, while the test set shows an MRE of 0.07213 and an RMSE of 0.0501. Compared with the results obtained using the original data, there is no significant discrepancy. Furthermore, as shown in Fig. 15, the rotational speed calculation model demonstrates considerable robustness under the influence of Gaussian noise, effectively simulating random noise such as sensor errors or background interference typically encountered in production environments.

To further evaluate the model's robustness, additional types of noise were introduced to replicate more realistic operating conditions. Sinusoidal noise was applied to simulate the effects of mechanical vibrations, while impulse noise was introduced to mimic electromagnetic interference. Finally, a combination of Gaussian, sinusoidal, and impulse noise was added to the current data to recreate complex noise conditions often encountered in oilfield environments. These extended experiments provide a more comprehensive understanding of the model's performance under various noise conditions, further confirming its capability to maintain accuracy and reliability in challenging real-world scenarios. The model's errors under different noise conditions are shown in Fig. 16. It can be seen that the model performs robustly under various noise conditions, with only slight increases in error for Gaussian, sinusoidal, impulse, and mixed noise, highlighting its reliability in diverse noise scenarios.

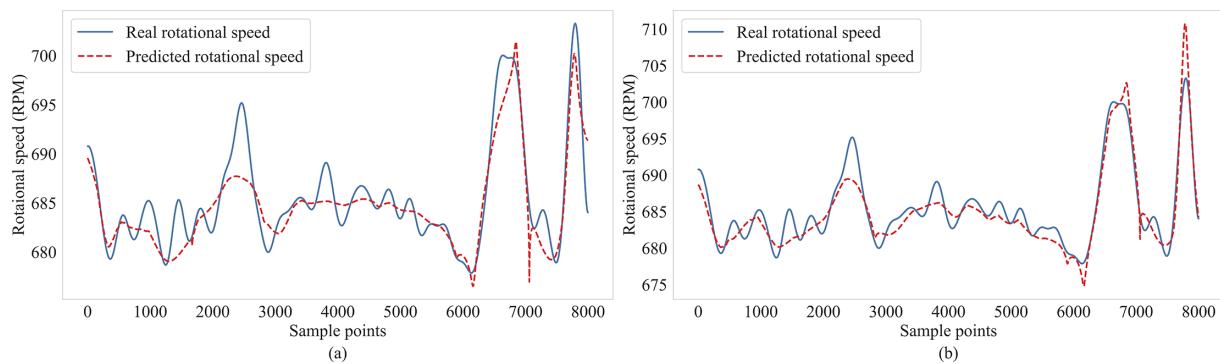
The model's strong robustness is attributable to the implementation of the PCA mechanism. By utilizing PCA for feature selection and dimensionality reduction, the influence of noise on the model is mitigated, resulting in enhanced accuracy and reduced training time. Fig. 17 presents a comparison between models with and without the PCA mechanism, focusing on three aspects: model accuracy, robustness, and training time measured in seconds. Here, robustness refers to the RMSE error in rotational speed calculations using noisy current



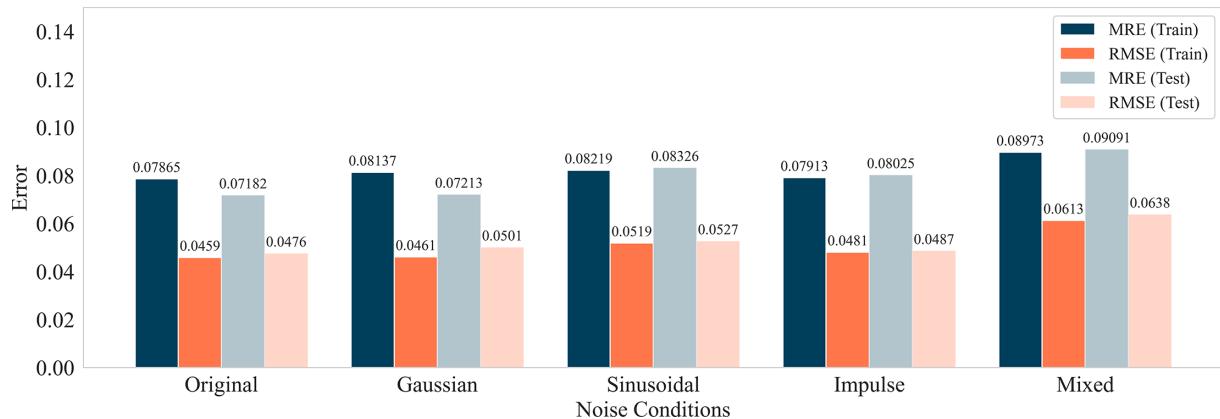
**Fig. 13—Modeling effects under different operating conditions.**



**Fig. 14—Current signal data with added noise.**



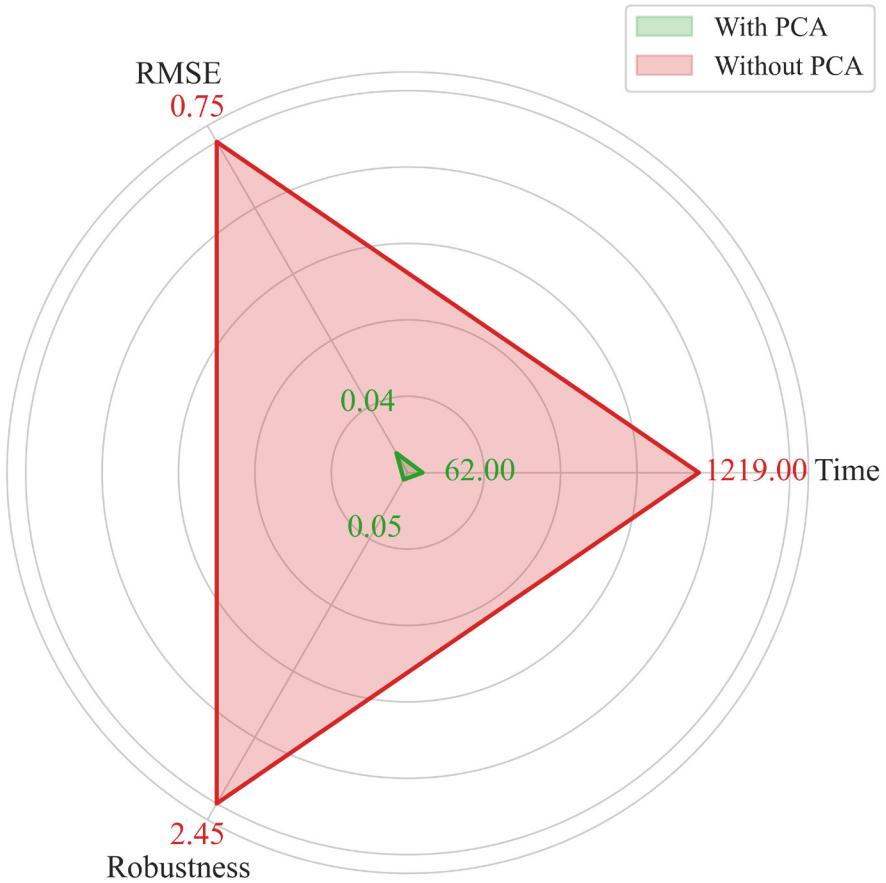
**Fig. 15—Comparison of calculation results for different current signals. (a) Calculation results of original current data. (b) Calculation results of noise current data.**



**Fig. 16—Model performance under different noise conditions.**

data, where lower values indicate greater robustness. As shown in **Fig. 17**, PCA significantly contributes to the improvement of the model's performance. From various test results, it can be concluded that the rotational speed calculation model can accurately compute rotational speed data using only current data, without the need for a speed sensor. This model exhibits strong robustness, thereby providing essential support for the subsequent calculation of MCSA features.

**Fault Diagnosis of SEM.** To train the integrated model effectively, the Adam optimizer was utilized due to its adaptive learning rate and computational efficiency, which are well-suited for handling the complex architecture comprising integrated features, MCSA features,



**Fig. 17—Comparison of model performance with and without PCA mechanism.**

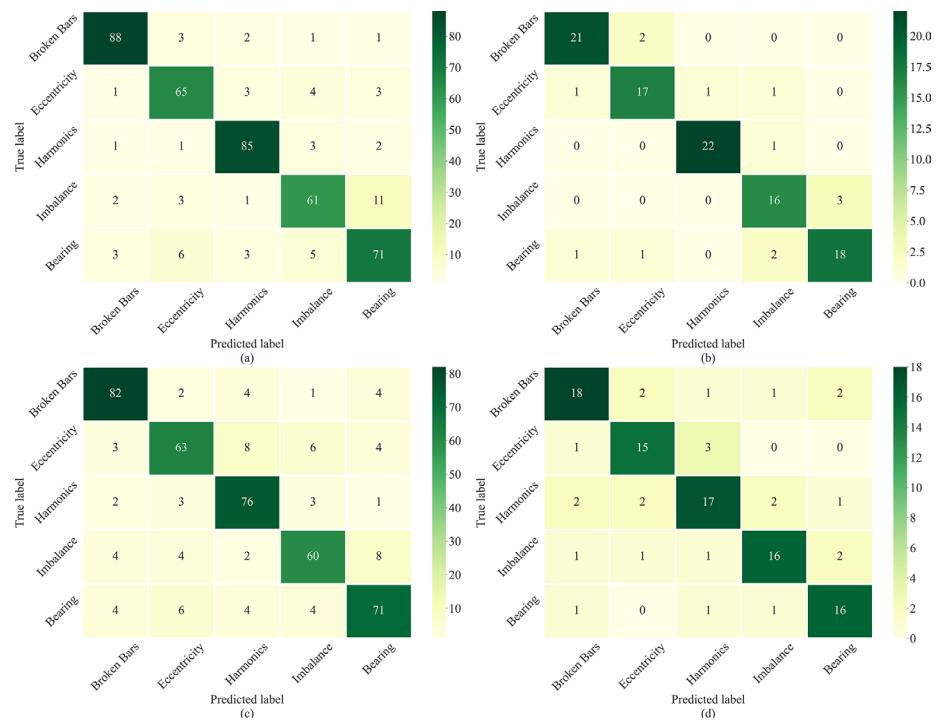
and attention mechanisms. The optimization objective was to minimize the categorical cross-entropy loss function, which is appropriate for multiclass fault diagnosis tasks. For the multiclass fault diagnosis tasks, the cross-entropy loss can be expressed as follows:

$$L(y, \hat{y}) = - \sum_{i=1}^C y_i \log(\hat{y}_i), \quad (26)$$

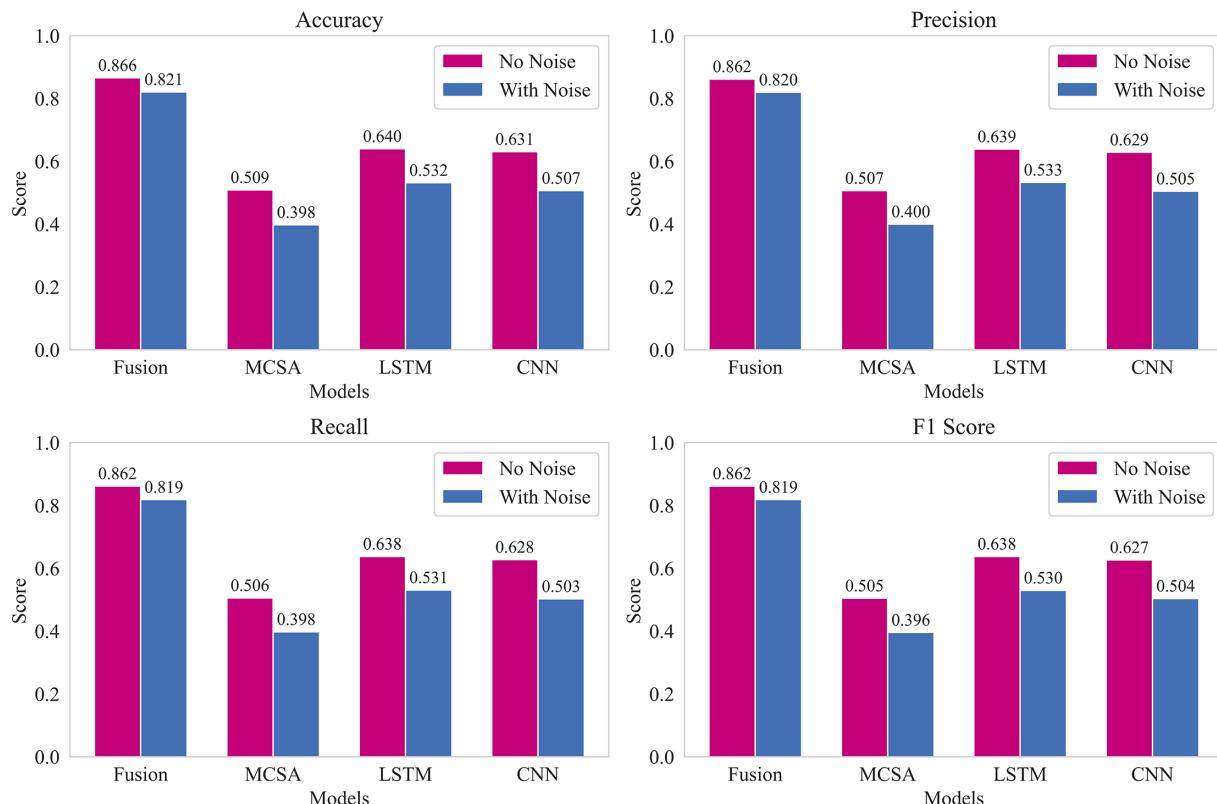
where  $C$  represents the number of fault classes,  $y_i$  is the one-hot encoding of the true label, and  $\hat{y}_i$  is the predicted probability for class  $i$ . The data set used for this analysis comes from fault records of the Bohai PL oil field in China and maintenance records from repair facilities, providing real-world fault data for training and evaluation. The training process was conducted with a batch size of 64 and an initial learning rate of 0.001. The model was trained for a maximum of 500 epochs, with early stopping applied if the validation loss failed to improve for 20 consecutive epochs. The data set was divided into 80% for training and 20% for testing, ensuring sufficient data for model evaluation and fine-tuning. In this study, a confusion matrix was utilized to evaluate the performance of the fault classification model. The confusion matrix provides a comparative analysis between the true labels and the predicted labels from the model, facilitating a deeper understanding of the classification effectiveness across different categories. To achieve intuitive visualization, the confusion matrix was presented as a heatmap, clearly illustrating the classification results for various categories, as shown in **Figs. 18a and 18b**. The heatmap indicates that the model performs exceptionally well on both the training and testing data sets. Notably, in the broken bars and harmonics categories, the model's prediction accuracy exceeds 90%, demonstrating its strong classification capability. Additionally, the model also performs well in the eccentricity, imbalance, and bearing categories, effectively identifying the vast majority of samples. In addition to its strong classification performance, the proposed model exhibits remarkable robustness in the presence of noise. Similarly, Gaussian noise was introduced into the current data to assess the robustness of the fault diagnosis model in the presence of interference. The results of fault diagnosis using the added noise data are shown in **Figs. 18c and 18d**.

The confusion matrix reveals that the introduction of noise data has somewhat affected the model's diagnostic capabilities; however, the overall classification performance remains robust. In both the training and testing sets, the accuracy for all fault categories consistently exceeds 80%. This demonstrates the model's strong robustness, affirming its effectiveness and reliability under noisy conditions. The favorable performance of the fault diagnosis model can be attributed to the integration of multisource features; this fusion approach leverages the strengths of each feature type, enabling a comprehensive understanding of the fault characteristics from multiple dimensions. In contrast, models utilizing a single feature type, such as only MCSA or spatial features, typically exhibit reduced diagnostic capabilities.

**Fig. 19** illustrates the diagnostic performance of four models—Fusion, MCSA, LSTM, and CNN—evaluated across four metrics: Accuracy, Precision, Recall, and F1 Score, under both noise-free and noisy conditions. The Fusion model consistently outperforms the others across all metrics, demonstrating robust performance with and without noise. In contrast, the MCSA model shows the most significant decrease in all metrics when noise is introduced, with scores dropping notably, indicating higher sensitivity to noise. The LSTM and CNN models display moderate noise resistance, with the LSTM model achieving slightly higher scores than CNN in noisy conditions. Notably, the scores for Accuracy, Precision, Recall, and F1 Score are almost identical across all models and conditions. This phenomenon



**Fig. 18—Heatmap of confusion matrix for raw and noisy data. (a) Heat map of confusion matrix for the training set. (b) Heat map of the confusion matrix for the test set. (c) Heatmap of the confusion matrix of the training set with added noise. (d) Heatmap of the confusion matrix of the test set with added noise.**



**Fig. 19—Comparison of diagnostic performance across different models.**

occurs because the data set used in this study is well-balanced, ensuring that true positives, false positives, and false negatives are distributed proportionally across fault classes. As a result, precision and recall closely align, leading to consistent F1 scores and accuracy values. Overall, the Fusion model's superior performance across all metrics, even under noisy conditions, highlights its effectiveness and robustness for fault diagnosis in challenging environments. To further improve the performance of the fault diagnosis model, a grid search

Learning Rate	LSTM Units	Kernel Size	Batch Size	Training Set Accuracy	Test Set Accuracy
0.001	512	8	128	0.896	0.874
0.001	448	7	128	0.871	0.863
0.003	256	10	64	0.866	0.833
0.002	256	8	32	0.847	0.805
0.004	256	9	64	0.803	0.791

Table 2—Performance comparison of different parameters of the fault diagnosis model.

method was used to optimize the model's hyperparameters. The hyperparameters to be optimized include the number of LSTM units, the kernel size, the learning rate, and the batch size. The learning rate is set to range from 0.001 to 0.01, with a tuning step of 0.001. The number of LSTM units ranges from 64 to 512, with a tuning step of 64. The kernel size ranges from 5 to 10, with a tuning step 1. The batch size ranges from 16 to 128, with a tuning step of 16. Some of the results are shown in **Table 2**. The final model parameters are a learning rate of 0.001, 512 LSTM units, a kernel size of 8, and a batch size of 128.

The fusion model was applied to monitor all ESP systems on a production platform in the Bohai PL oil field in China. During the monitoring, the model detected abnormal operations in seven wells. The detection occurred before the actual failures of the ESP systems. After the abnormalities were identified, the ESP systems in these wells failed over time. This confirmed the accuracy and effectiveness of the model's diagnostic results. The diagnostic records and failure details of the seven ESP wells are shown in **Table 3**. The data in the table demonstrate that the model provided timely warnings before the failures occurred. This gave the maintenance team enough time to prepare and respond. The results show that the model performed well in the field. It improved the efficiency of ESP system maintenance and reduced unplanned downtime. This makes the model a reliable tool for smart oilfield management.

Various results indicate that, by integrating multiple sources of features, the diagnostic accuracy of the fault diagnosis model can be significantly improved, enabling timely and accurate diagnosis of SEM faults. However, accurately diagnosing faults alone is not sufficient to meet the requirements of sustainable production. After identifying the corresponding fault conditions, it is essential to implement preventive measures tailored to each specific fault. The five fault conditions discussed in this study, along with their causes and possible preventive measures, are as follows:

1. Broken Bars. If the SEM operates under an overload condition for an extended period, the rotor bars may weaken due to overheating, eventually leading to fracture. Additionally, the presence of sand particles in the well fluid or metal debris from the pump unit caused by corrosion can further damage the rotor bars. Once rotor bars are fractured, the motor's performance declines significantly, which may result in a complete system shutdown. To prevent broken rotor bars, the motor load current should be maintained below 85% of its rated current. If significant current fluctuations or increasing load trends are detected, the production rate should be gradually reduced to avoid prolonged stress on the rotor bars. Sand control measures should be implemented to minimize the impact of particles or debris.
2. Eccentricity. During SEM operation, changes in temperature and mechanical stress can cause deformation of the pump casing and internal components, leading to eccentricity between the rotor and stator. This misalignment increases the risk of motor overheating.

Well No.	Date of Model-Detected Anomaly	Fault Type Diagnosed by Model	Date of ESP System Failure	Fault Records of the Oil Field
1	21 January 2023	Eccentricity	1 March 2023	Decline in efficiency on 25 January 2023. Current fluctuations detected on 19 February 2023. Motor failed due to overload. Inspections showed air gap deviation.
2	9 February 2023	Bearing fault	15 February 2023	Motor temperature exceeded 102°C on 11 February 2023, followed by severe vibration. Motor failed due to burnout. Inspections found bearing fractures and localized burning.
3	31 March 2023	Broken bars	10 Apr 2023	Pressure drop and flow decline detected on 2 April 2023. Motor failed on 10 April 2023. Inspections revealed two broken rotor bars.
4	26 September 2023	Broken bars	30 September 2023	Abnormal vibrations on 26 September 2023 and current fluctuations on 28 September 2023. Motor failed on 30 September 2023. Inspections revealed three broken rotor bars.
5	9 November 2023	Imbalance	29 November 2023	Increased power consumption and voltage fluctuations on 11 November 2023. Motor failed due to cable failure. Inspections found phase voltage imbalance.
6	20 November 2023	Harmonics	25 Nov 2023	Current exceeded normal range on 23 November 2023. Motor shutdown due to overload. Inspections found harmonics caused by frequent adjustments.
7	16 December 2023	Eccentricity	8 January 2024	Increased vibration on 17 December 2023. Irregular current fluctuations detected on 4 January 2024. Motor shutdown due to overload. Inspections found uneven air gaps.

Table 3—Field monitoring records of ESP wells and model diagnostics results.

ing and further deformation, which can accelerate wear and reduce motor lifespan. Severe eccentricity may eventually result in system failure due to mechanical friction or rotor damage. To prevent eccentricity, the motor's current fluctuations, power factor, and vibration levels should be closely monitored. If vibration levels exceed 2 mm/s, the motor load should be reduced by adjusting the production rate or operating frequency. It is important to ensure the motor operates within 70–85% of its rated current to limit overheating. Regular evaluation of power factor data can help identify early signs of misalignment.

3. **Harmonics.** Harmonics in the power supply system can result from instability or pollution caused by inverters or other nonlinear loads. Frequent start-stop cycles or fluctuating loads in the ESP system can also generate current harmonics, which may lead to overheating, increased energy consumption, and motor stress. Severe harmonics can eventually trigger system shutdowns due to overload protection. To mitigate harmonics, filters should be used to keep total harmonic distortion below 5%. If distortions are detected, adjustments should be made to stabilize power delivery and avoid rapid load fluctuations. Frequent start-stop cycles should also be avoided, and maintaining a stable load schedule is essential to reduce harmonic-related risks.
4. **Imbalance.** Inconsistent voltage losses during long-distance power transmission can lead to phase voltage imbalances. Additionally, aging cables, poor connections, or localized damage to transformers may further exacerbate the imbalance. Phase voltage imbalance can cause uneven current distribution in the motor, leading to overheating, increased vibrations, and eventual motor or cable failure. To address imbalance, the voltage symmetry should be carefully monitored, and any signs of imbalance should trigger maintenance actions. Regular inspections of transformers, cables, and connection points should be conducted to identify and resolve potential issues. Adjustments to the production schedule can help reduce operational stress during periods of voltage instability. Installing a three-phase imbalance protection device can further safeguard the system by preventing severe damage through automatic shutdowns when necessary.
5. **Bearing.** Excessive motor load increases the pressure on bearings, leading to surface damage and reduced service life. Electrical issues, such as current fluctuations or leakage currents, may cause electrical erosion, resulting in pits on the bearing surface. Over time, these issues can cause severe vibrations and overheating, ultimately leading to motor failure. To prevent bearing faults, the motor's temperature and vibration levels should be carefully monitored. A warning should be triggered if the motor temperature exceeds 100°C or if vibration velocity amplitude surpasses 2.5 mm/s. When such warnings are issued, the production rate or motor operating frequency should be adjusted to reduce stress on the bearings. Large frequency adjustments should be avoided to minimize the risk of sudden current fluctuations that could exacerbate bearing damage.

The preventive measures proposed in this study provide a general direction for reducing the occurrence of SEM faults. However, specific measures should be tailored to the conditions of individual oilfields and ESP wells to ensure their effectiveness in practical applications. The successful implementation of these preventive measures relies not only on their adaptability to specific field conditions but also on the computational efficiency and hardware capabilities required to support the proposed fault diagnosis model. To ensure the model's practicality in real-world applications, it is crucial to evaluate its hardware demands and processing requirements under field conditions. The fault diagnosis model can run in real time for ESP systems because it is trained off-site and then loaded onto edge computing devices for inference. This reduces the need for powerful hardware in the field since training is not required on-site. The edge computing devices perform tasks such as signal processing, feature extraction, and fault classification. A quad-core CPU (e.g., Intel Core i5) and a lightweight GPU (e.g., NVIDIA Jetson Nano or Xavier NX) are enough for inference. The GPU speeds up CNN feature extraction and attention mechanisms, while the CPU handles FFT and LSTM inference. Devices with 4–8 GB of RAM and enough storage for the pretrained model can process high-frequency signals in real time, with a latency of 20–50 ms per signal. This setup supports the requirements for ESP systems and allows for accurate fault diagnosis in the field.

The hardware configuration described above makes it possible to evaluate the cost-effectiveness of the system. A cost-benefit analysis was done to compare the system with traditional monitoring methods. For a single ESP well, the proposed system needs an initial investment of USD 6,000–8,000, which includes the edge computing device and a share of model development costs. Annual maintenance costs are USD 2,000–3,000 per well. Traditional methods rely on manual inspections and vibration analysis, with annual costs of USD 600–2,400 for labor and USD 8,000–10,000 for equipment. These methods often fail to prevent about 5 days of downtime per year. For a well producing 500 barrels per day at USD 75 per barrel, this results in losses of USD 187,500 annually. The proposed system can reduce downtime by 50–70%, saving USD 93,750–131,250 each year, and lower repair costs by 30–50%, saving USD 3,000–10,000 per well. The cost savings depend on when faults occur and will take time to appear. For a single well, annual savings range from USD 96,750 to 141,250, with a return on investment achieved in 6–12 months. For a system with 10 wells, the total investment is USD 60,000–80,000, with annual savings of USD 967,500–1,412,500, achieving a return on investment in 12–18 months. The system provides long-term benefits by preventing faults, reducing downtime, and improving reliability. To further enhance the practicality of the system, it is also important to consider how the proposed system can be seamlessly integrated into existing oilfield monitoring infrastructure. The proposed system integrates with existing oilfield monitoring infrastructure through edge computing devices that process high-frequency current signals locally, minimizing the need for major modifications. It supports standard protocols such as Modbus and OPC UA (open platform communications unified architecture) to receive input signals and transmit diagnostic results, allowing seamless communication with existing sensors, programmable logic controllers, and supervisory control and data acquisition systems. These diagnostic results can be visualized on central platforms and incorporated into current monitoring workflows. Its modular design allows incremental deployment, starting with critical locations and scaling up as needed. This approach ensures minimal disruption to existing workflows, enabling operators to use familiar interfaces while benefiting from automated diagnostics, early warnings, and improved operational efficiency.

## Conclusion

A rotational speed calculation model for the SEM was built using CNNs, allowing the calculation of rotational speed of SEM solely based on current data, without the need for a speed sensor. This provides a basis for accurately calculating MCSA fault features. Furthermore, the introduction of the PCA mechanism reduced the training time of the speed calculation model and enhanced its stability in the presence of noisy data. Additionally, by utilizing LSTM networks and CNNs to extract spatiotemporal features, and incorporating an attention mechanism to fuse these spatiotemporal features with MCSA fault features, a fault diagnosis model was developed. This model is capable of accurately diagnosing SEM faults and demonstrates strong stability even under the influence of noise. Future work will focus on improving the model's robustness under more complex noise conditions, such as non-Gaussian noise and electromagnetic interference, to ensure accuracy in challenging environments. Additionally, optimizing the model's real-time performance will enable faster processing of high-frequency data streams, meeting the operational needs of oilfield monitoring systems. By accurately diagnosing SEM faults and implementing corresponding preventive measures, the management level of oilfields can be improved, making oil field production more sustainable.

## Funding

This research was supported by the Science Foundation of China University of Petroleum-Beijing (Grant No. 2462023YJRC019) and the National Key Research and Development Project of China (Grant No. 2023YFA1011700).

## References

- Alamu, O. A., Pandya, D. A., Warner, O. et al. 2020. ESP Data Analytics: Use of Deep Autoencoders for Intelligent Surveillance of Electric Submersible Pumps. Paper presented at the Offshore Technology Conference, Houston, Texas, USA, 4–7 May. <https://doi.org/10.4043/30468-MS>.
- AlBallam, S., Karami, H., and Devegowda, D. 2022. A Data-Based Reliability Analysis of ESP Failures in Oil Production Wells. *J Energ Power Technol* **04** (4): 1–29. <https://doi.org/10.21926/jept.2204036>.
- Badkoubeh, A. and Trevisan, E. F. 2021. Motor Intelligence Versus Vibration Analysis for Early-Stage Diagnosis of Electric Submersible Pumps. Paper presented at the SPE Gulf Coast Section Electric Submersible Pumps Symposium, Virtual and The Woodlands, Texas, USA, 4–8 October. <https://doi.org/10.2118/204522-MS>.
- Badkoubeh, A., Waldner, L. B., and Imanfarid, M. 2023. Field Validation: ESP Reliability Monitoring and Production Optimization by Trending Condition and Performance Data Extracted from ESP Surface Electrical Signals. Paper presented at the SPE Gulf Coast Section - Electric Submersible Pumps Symposium, The Woodlands, Texas, USA, 2–6 October. <https://doi.org/10.2118/214728-MS>.
- Fakher, S., Khlaifat, A., Hossain, M. E. et al. 2021. Rigorous Review of Electrical Submersible Pump Failure Mechanisms and Their Mitigation Measures. *J Petrol Explor Prod Technol* **11** (10): 3799–3814. <https://doi.org/10.1007/s13202-021-01271-6>.
- Nyquist, H. 1928. Certain Topics in Telegraph Transmission Theory. *Trans Am Inst Electr Eng* **47** (2): 617–644. <https://doi.org/10.1109/T-AIEE.1928.5055024>.
- Pham, S. T., Vo, P. S., and Nguyen, D. N. 2021. Effective Electrical Submersible Pump Management Using Machine Learning. *OJCE* **11** (1): 70–80. <https://doi.org/10.4236/ojce.2021.111005>.
- Popaleny, P., Duyar, A., Ozel, C. et al. 2018. Electrical Submersible Pumps Condition Monitoring Using Motor Current Signature Analysis. Paper presented at the Abu Dhabi International Petroleum Exhibition & Conference, 12–15, Abu Dhabi, UAE. <https://doi.org/10.2118/193081-MS>.
- Rodrigues, D. A., Martins, G. S. O., David, E. R. et al. 2023. Fault Diagnosis of Electric Submersible Pumps Using Vibration Signals. *J Braz Soc Mech Sci Eng* **45** (9): 445. <https://doi.org/10.1007/s40430-023-04370-z>.
- Takacs, G. and Takacs, G. 2014. *Electrical Submersible Pumps Manual*. Houston, Texas, USA: Gulf Professional Publishing.
- Thomson, W. T. and Culbert, I. 2016. *Current Signature Analysis for Condition Monitoring of Cage Induction Motors: Industrial Application and Case Histories*, 1st ed. Hoboken, New Jersey, USA: Wiley-IEEE Press. <https://doi.org/10.1002/9781119175476>.
- Thomson, W. T. and Gilmore, R. J. 2003. *Motor Current Signature Analysis To Detect Faults In Induction Motor Drives - Fundamentals, Data Interpretation, And Industrial Case Histories*. College Station, Texas, USA: Texas A&M University. Turbomachinery Laboratories. <https://doi.org/10.21423/R1CS91>.
- Villalobos-Pina, F. Javier., Reyes-Malanche., J. A., Cabal-Yepez., E. et al. 2024. Electric Fault Diagnosis in Induction Machines Using Motor Current Signature Analysis (MCSA). In *Time Series Analysis - Recent Advances, New Perspectives and Applications*. IntechOpen. <https://doi.org/10.5772/intechopen.1004002>.
- Walters, R., Ehman, K., Olson, K. et al. 2023. Trials of VSD Electrical Signature Analysis Technology for ESP Vibration in High Dog Leg Severity Wells at Surmont. Paper presented at the SPE Gulf Coast Section - Electric Submersible Pumps Symposium, The Woodlands, Texas, USA, 2–6 October. <https://doi.org/10.2118/214729-MS>.
- Yang, J., Wang, S., Zheng, C. et al. 2022. Fault Diagnosis Method and Application of ESP Well Based on SPC Rules and Real-Time Data Fusion. *Math Probl Eng* **2022**: 1–15. <https://doi.org/10.1155/2022/8497299>.