

Original Paper

Downhole vibration recognition using a wavelet transform based convolutional neural network model

Xue-Ying Wang^{a,b}, Ming Liu^{a,b}, Jian-Min Wang^c, Wei Li^d, Fei-Fei Zhang^{a,b,*}

^a School of Petroleum Engineering, Yangtze University, Wuhan, 430100, Hubei, China

^b Hubei Key Laboratory of Oil and Gas Drilling and Production Engineering, Wuhan, 430100, Hubei, China

^c Drilling Division, CNPC Offshore Engineering Co., Ltd., Tianjin, 300451, China

^d Geosteering & Logging Research Institute, Sinopec Matrix Corporation, Dongying, 257000, Shandong, China

ARTICLE INFO

Article history:

Received 30 June 2025

Received in revised form

13 January 2026

Accepted 9 June 2026

Available online 15 June 2026

Edited by Xi Zhang

Keywords:

Downhole vibrations

Backward whirl

Stick-slip

Wavelet transform

Convolutional neural network

ABSTRACT

Severe downhole drill-string vibrations can accelerate fatigue damage, leading to substantial economic losses and increased non-productive time. Existing vibration monitoring approaches typically rely on multiple high-cost downhole sensors, which has hindered their widespread deployment in field operations. This work presents a real-time vibration monitoring approach based solely on acceleration measurements and machine-learning techniques. Radial and tangential acceleration signals are generated using a rotor dynamics model and transformed into time–frequency scalograms via continuous wavelet transform. Among various whirl modes, backward whirl is recognized as the most damaging to drilling tools and wellbore integrity. Accordingly, vibration events are classified into backward whirl and non-backward whirl categories, achieving classification accuracies exceeding 94% across different training–testing split ratios. For stick-slip identification, the proposed method distinguishes stick-slip from non-stick-slip conditions with 100% accuracy. Furthermore, based on the stick-slip index, stick-slip vibrations are subdivided into four severity levels—non-stick-slip, low, medium, and high—with classification accuracies exceeding 95% across all severity classes. The generalization capability of the proposed model is validated using independent acceleration data from ABAQUS simulations and field measurements, demonstrating robust performance beyond the training dataset. Backward whirl and stick-slip are accurately classified using single-point acceleration data alone, without requiring rotational speed measurements or symmetrically distributed accelerometers. This approach reduces sensor requirements, lowers operational costs, and facilitates practical real-time deployment.

© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

During drilling operations, severe drill string vibrations degrade bit and bottom-hole assembly (BHA) performance, resulting in reduced rate of penetration, increased mechanical specific energy, and elevated non-productive time (Bailey et al., 2008; Esmaeili et al., 2012). Such vibrations may also compromise wellbore stability and increase the risk of wellbore collapse (Zhu and Liu, 2013). Accurate detection and timely mitigation of

downhole vibrations are therefore essential for safe and efficient drilling.

Drill string vibrations are commonly categorized as axial, torsional, and lateral. Lateral vibrations of a rotating BHA can readily induce whirl (Besaisow et al., 1988), which is generally classified as forward, backward, or chaotic based on the relative rotational and orbital motion of the BHA. Stick-slip, a dominant form of torsional vibration, is characterized by alternating phases of sticking and slipping and may arise from either cutting-action mechanisms or frictional interactions (Chen et al., 2020).

Current approaches for whirl identification rely on measurements such as acceleration and bending moment. Early implementations were limited by sensor constraints and bandwidth restrictions, leading to measurement inaccuracies and delayed

* Corresponding author.

E-mail address: feifei-zhang@yangtzeu.edu.cn (F.-F. Zhang).

Peer review under the responsibility of China University of Petroleum (Beijing).

real-time detection (Shyu, 1989; Zannoni et al., 1993). Advances in sensor technology have improved measurement accuracy and temporal resolution. Minett-Smith et al. (2010) demonstrated real-time whirl detection using a two-point measurement scheme with radially symmetric accelerometers, while Bowler et al. (2016) employed recorded long-duration, high-frequency downhole measurements to identify whirl behavior and mode transitions. However, large data volumes and the high cost of precision sensors continue to hinder widespread field deployment.

Accurate monitoring of rotational speed is essential for calculating the stick-slip index (SSI). One approach estimates rotational speed by tracking variations in bit rotation using triaxial accelerometers (Townsend et al., 2021); however, reliable measurement at low rotational speeds remains challenging. Gyroscopes provide more accurate rotational speed measurements and are often integrated with other sensors. For example, Jones et al. (2017) developed a downhole sensing system combining a triaxial gyroscope with vibration sensors to detect stick-slip behavior. Despite these advances, gyroscopes are constrained by limited measurable speed ranges and reduced sensitivity to low-frequency signals. Moreover, combining multiple sensing technologies increases system complexity and cost. As a result, achieving accurate and cost-effective rotational speed monitoring in downhole environments remains a significant challenge.

Whirl, stick-slip, and other vibration modes exhibit distinct signatures in acceleration signals. However, traditional identification methods typically depend on rotational speed, bending moment, and other auxiliary measurements, increasing both system complexity and cost. This study adopts a single-point acceleration-based approach combined with machine learning (ML) to enable accurate vibration classification without rotational speed or bending moment measurements. ML has been successfully applied across the oil and gas industry, including drilling fluid property prediction (Gowida and Elkatatny, 2024), reservoir characterization (Moya et al., 2024), and drilling performance optimization (Olukoga and Feng, 2021). It has also been employed for drill-string vibration analysis using surface or downhole measurements (Baumgartner and van Oort, 2014; Millan et al., 2019; Saadeldin et al., 2023; Wang et al., 2024). However, existing approaches often depend on multi-sensor inputs or simplified dynamic models, limiting their robustness under complex downhole vibration conditions.

Continuous wavelet transform (CWT) is a powerful time–frequency analysis technique that combines multi-resolution analysis with strong localization capability and has been widely applied in petroleum engineering. Previous studies have demonstrated the effectiveness of CWT in a range of applications, including hydraulic fracturing diagnostics (Gabry et al., 2025a), fracture characterization (Gabry et al., 2025b), inter-well connectivity analysis (Gabry et al., 2025c), and microseismic signal interpretation (Sheng et al., 2015; Zhao et al., 2021). These applications highlight the ability of CWT to capture transient and non-stationary signal features under complex subsurface conditions. Building on these advances, this study employs CWT to transform acceleration signals into time–frequency scalograms, which are subsequently classified using a convolutional neural network based on AlexNet.

Section 2 describes the dynamic modeling of acceleration signals associated with backward whirl and stick-slip and analyzes their characteristic features. Section 3 outlines the fundamentals of the wavelet transform, the Morse wavelet, and the AlexNet architecture. Section 4 presents the classification results obtained from simulated data and validates the proposed method using ABAQUS simulations and field data. The overall workflow of the study is illustrated in Fig. 1.

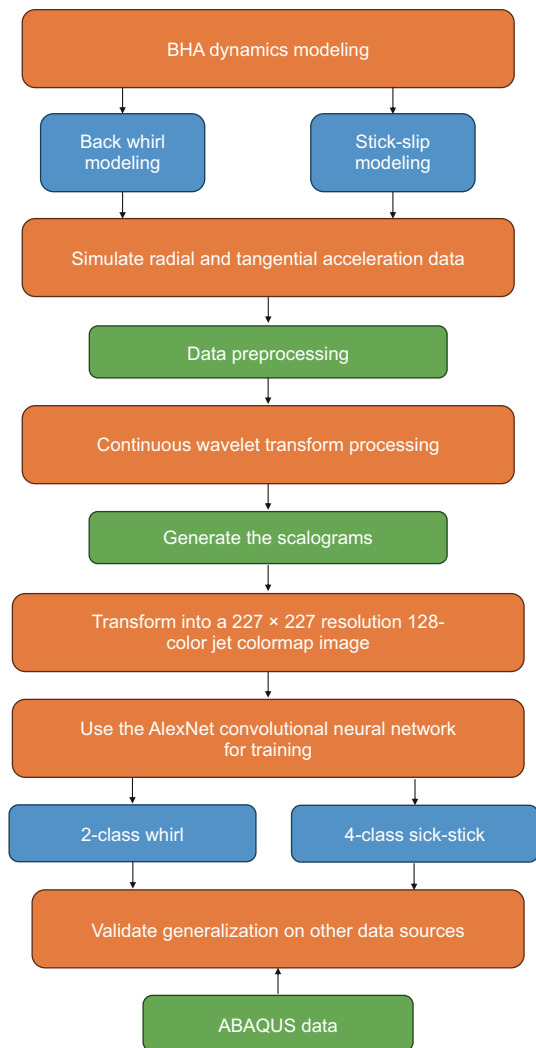


Fig. 1. Workflow of the proposed backward whirl and stick-slip classification method.

2. Dynamic modeling and acceleration signal generation

This section presents the dynamic modeling and acceleration characteristics of backward whirl and stick-slip. A single-disk rotor dynamics model is used to generate training data, from which radial and tangential acceleration signals are obtained under different vibration conditions. Spectral analysis is subsequently performed to examine the characteristic features of the simulated acceleration signals.

2.1. Dynamic modeling framework

The single-disk rotor dynamics model adopted in this study is a classical theoretical framework that represents the BHA as a rotating disk. By prescribing the disk rotational speed, the model enables effective simulation of backward whirl and stick-slip vibrations. This simplified representation captures the essential downhole vibration behavior in a tractable mathematical form, allowing efficient simulation of multiple vibration modes encountered during drilling operations. Owing to its flexibility and computational efficiency, the model is well suited for investigating downhole vibration mechanisms. The model configuration is illustrated in Fig. 2.

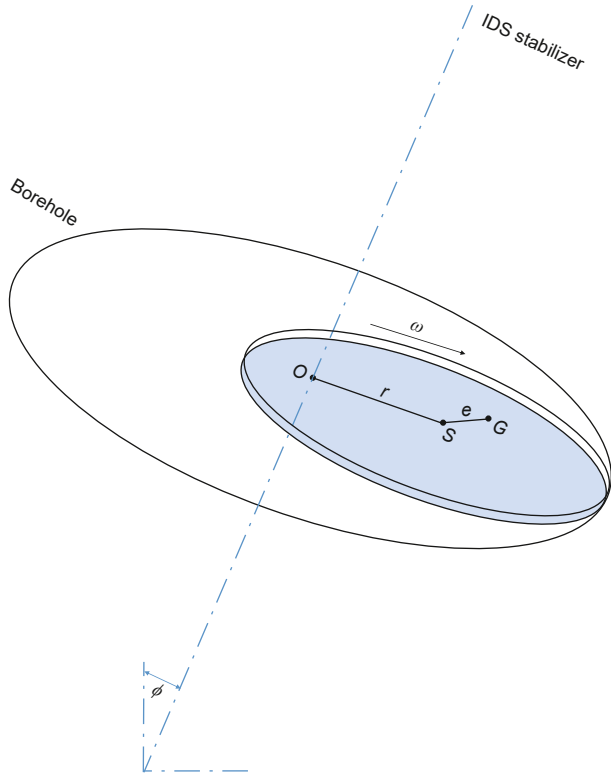


Fig. 2. Schematic of the single-disk rotor dynamics model (after Stroud et al., 2011).

As illustrated in Fig. 2, ϕ denotes the inclination angle, and O represents the geometric center of the wellbore. Points S and G correspond to the geometric and mass centers of the disk, respectively, with their separation defined as the eccentricity e , which plays a key role in inducing whirl. r is the distance from S to O. The disk rotates at an angular velocity ω and while simultaneously undergoing orbital motion.

2.1.1. Whirl modeling

Assuming a constant rotational speed ω , the disk undergoes whirl motion. The bending stiffness of the drill string is denoted by k . When contact with the wellbore wall occurs, the disk is subjected to a normal contact force F_{BN} and a tangential friction force F_T characterized by the friction coefficient μ . In addition, a linear damping force with coefficient C is applied at point S, as illustrated in Fig. 3.

When the radial displacement r exceeds the critical clearance R_c , a contact force F_{BN} is exerted on the disk by the wellbore wall. If $r \leq R_c$, no contact is assumed to occur, as shown below.

$$F_{BN} = \begin{cases} 0 & r \leq R_c \\ k_B(r - R_c) & r > R_c \end{cases} \quad (1)$$

where k_B is the wellbore wall contact stiffness.

The resulting equation of motion for the disk in the radial direction is given as follows (Stroud et al., 2011):

$$m[\ddot{r} - r\dot{\theta}^2 - g \sin \phi \cos \theta - e\omega^2 \cos(\omega t - \theta)] = -kr - C\dot{r} - F_{BN} \quad (2)$$

The equation of motion in the tangential direction is as follows (Stroud et al., 2011):

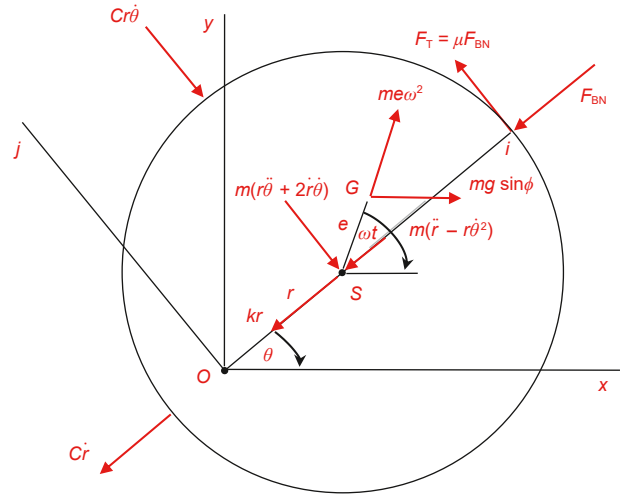


Fig. 3. Force analysis of the rotating disk (Stroud et al., 2011).

$$m[r\ddot{\theta} + 2\dot{r}\dot{\theta} + g \sin \phi \cos \theta - e\omega^2 \sin(\omega t - \theta)] = -C\dot{\theta} - F_T \quad (3)$$

where m is the mass of the disk; g is the gravitational acceleration; θ is the angular position around the wellbore center; F_T is the friction force between the disk and the wellbore wall, acting opposite to the direction of disk motion.

The tangential friction force F_T is calculated as follows:

$$F_T = \text{sign}(V_{\text{rel}})\mu F_{BN} \quad (4)$$

$$V_{\text{rel}} = \omega R_S + \dot{\theta}r \quad (5)$$

where V_{rel} is the relative velocity between the disk and the wellbore wall, R_S is the radius of the disk.

By substituting Eqs. (1) and (4) into Eqs. (2) and (3) and simplifying, the expressions for radial acceleration $\ddot{r}$ and tangential acceleration $\ddot{\theta}$ are obtained as follows (Stroud et al., 2011):

$$\ddot{r} = -\frac{C}{m}\dot{r} + \left(\dot{\theta}^2 - \frac{k}{m}\right)r + g \sin \phi \cos \theta + e\omega^2 \cos(\omega t - \theta) - \frac{k_B}{m}(r - R_c) \quad (6)$$

$$\ddot{\theta} = -\left(\frac{2\dot{r}}{r} + \frac{C}{mr}\right)\dot{\theta} - \frac{g}{r} \sin \phi \sin \theta + \frac{e\omega^2}{r} \sin(\omega t - \theta) - \frac{\mu k_B}{mr}(r - R_c) \quad (7)$$

2.1.2. Stick-slip modeling

Stick-slip simulation using the single-disk rotor dynamics model follows the same conditions as whirl simulation, with a revised assumption for the rotational speed ω (Wang et al., 2024):

$$\omega = \begin{cases} \frac{\omega_s T \pi}{2T_1} \sin\left(\frac{\pi}{T_1} t\right) & T \leq t < T_1 + nT, n = 0, 1, 2, \dots \\ 0 & T_1 + nT \leq t < T_2 + nT, n = 0, 1, 2, \dots \end{cases} \quad (8)$$

where T_1 is the sliding duration, T_2 is the sticking duration, T is the stick-slip cycle period, and ω_s is the surface rotational speed.

2.2. Acceleration signal characteristics

The acceleration signal of backward whirl has significant periodicity in the time domain, frequently accompanied by high-amplitude transient impacts resulting from intermittent contact between the drill string and the wellbore wall. Stick-slip exhibits a nonlinear pattern marked by alternating phases of steady motion and abrupt changes.

2.2.1. Backward whirl acceleration characteristics

Whirl modes are classified according to the relationship between disk rotational speed ω and the orbital speed about the borehole center Ω . Table 1 summarizes the resulting whirl regimes and defines d_h and d_c as the borehole and rotor diameters, respectively. Fig. 4 provides a schematic illustration of the different whirl modes.

In synchronous forward whirl, the rotational speed ω equals the whirl speed Ω , and both rotate in the same direction. As a result, the same point on the drill string remains in continuous contact with the borehole wall, which can lead to localized and severe unilateral wear of the BHA. When $0 < \Omega < \omega$, forward whirl occurs with the rotation center offset from the wellbore center, representing a transition between no whirl and synchronous forward whirl. At $\Omega = 0$, only pure drill string rotation exists without whirl. When Ω is opposite in direction to ω , backward whirl arises, which can damage both the drill string and the wellbore. If $\Omega = -\omega \left[\frac{d_c}{d_h - d_c} \right]$, synchronous backward whirl occurs, characterized by a higher whirl frequency and increased risk to downhole components (Macpherson et al., 2015).

The periodicity of backward whirl acceleration is examined through simulations conducted at a disk speed of 120 r/min. Fig. 5 presents the time-domain radial and tangential acceleration signals, where the raw data exhibit strong fluctuations and high-amplitude impacts due to intermittent contact between the drill string and the borehole wall. Because the periodicity is not readily discernible in the raw signal, fast Fourier transform (FFT) is applied to extract the dominant frequency components (Nussbaumer, 1982). The filtered signal (Fig. 5) and the enlarged view in Fig. 6 reveal peak-to-peak intervals ranging from 0.0625 s to 0.0775 s, corresponding to the backward whirl period.

To further confirm the periodicity of backward whirl, frequency spectra of the raw radial and tangential acceleration signals are analyzed. The dominant frequency components observed in Fig. 7 verify the periodic nature of the backward whirl motion.

As shown in Fig. 7, both radial and tangential acceleration spectra exhibit a dominant frequency at 12.1094 Hz, while high-frequency components can be neglected. This frequency agrees well with the backward whirl period identified in the time domain (Fig. 6). According to the kinematic relationship in Table 1, the

backward whirl frequency depends on the BHA rotational speed, diameter, and borehole clearance. Using the model parameters, the theoretical backward whirl frequency is calculated as 12 Hz, which closely matches the spectral result and corresponds to a whirl period of approximately 0.084 s.

In practical downhole conditions, slip commonly occurs due to insufficient friction, borehole irregularities, tool wear, and nonlinear drill-string dynamics (Popp et al., 2018). As illustrated in Figs. 8 and 9, backward whirl with slip at 120 r/min exhibits a dominant frequency of 9.96 Hz, which is about 80% of the theoretical value. This reduction reflects the weakened rolling constraint at the contact interface and indicates that backward whirl cannot be reliably identified using frequency-based criteria alone. Therefore, slip effects should be explicitly included in the recognition samples to ensure robustness and engineering relevance under realistic drilling conditions.

2.2.2. Stick-slip acceleration characteristics

Cutting-action-induced stick-slip is characterized by near-zero rotational speed during the sticking phase, with minimal axial and lateral vibration. In this regime, torsional and axial motions are coupled, and the axial response may also exhibit intermittent stick-slip behavior. Friction-induced stick-slip can be further classified into drill pipe, BHA, and bit stick-slip. In the case of drill pipe stick-slip, pseudo stick-slip may occur, in which the bit maintains a non-zero rotational speed during the sticking phase, typically corresponding to the mud motor speed.

Fig. 10 shows a representative acceleration signal associated with stick-slip vibration, characterized by periodic transitions between low-activity (sticking) and high-activity (sliding) phases. During the sticking phase, acceleration and rotational speed remain low, whereas the sliding phase is marked by abrupt increases in acceleration amplitude and impulsive responses. The rotational speed exhibits pronounced oscillations, reflecting rapid energy release and accelerated bit motion relative to the borehole.

Fig. 11 presents field-acquired stick-slip acceleration data sampled at 1000 Hz, where an eccentrically mounted sensor and gyroscope were used to measure tri-axial acceleration and rotational speed, respectively, during drilling in a siltstone formation with a 215.9 mm PDC bit under a WOB of 12 tones.

Compared with the simulated signal in Fig. 10, field-measured acceleration and rotational speed exhibit noticeable differences. Although periodic stick-slip behavior persists, acceleration remains non-zero during the nominal sticking phase, indicating pseudo stick-slip. To better represent field conditions, such pseudo stick-slip behavior—characterized by non-zero acceleration and rotational speed during sticking—is incorporated into the training dataset. This is achieved by introducing low-amplitude random disturbances into the single-disk rotor model, resulting in simulated signals that reproduce the observed field behavior, as shown in Fig. 12.

3. Method

3.1. Continuous wavelet transform

Vibration signals often display irregular waveforms and substantial fluctuations, downhole vibration signals are no exception, with amplitudes varying significantly over time. CWT is adept at analyzing transitory features. As one type of wavelet transform, CWT serves as a robust tool for time-frequency analysis and is essential in numerous scientific and engineering disciplines.

Wavelet transform, like Fourier transform, can identify dominant frequencies in a signal. Unlike Fourier transform, it uses a mother wavelet such as Haar, Daubechies, Morlet, or Symlet for analysis (Safavian et al., 2005). By translating and scaling the

Table 1
Whirl modes based on ω - Ω relationship.

Whirl modes	Relationship
Synchronous forward whirl	$\Omega = \omega$
Forward whirl	$0 < \Omega < \omega$
No whirl	$\Omega = 0$
Backward whirl	$-\omega \left[\frac{d_c}{d_h - d_c} \right] < \Omega < 0$
Synchronous backward whirl	$\Omega = -\omega \left[\frac{d_c}{d_h - d_c} \right]$

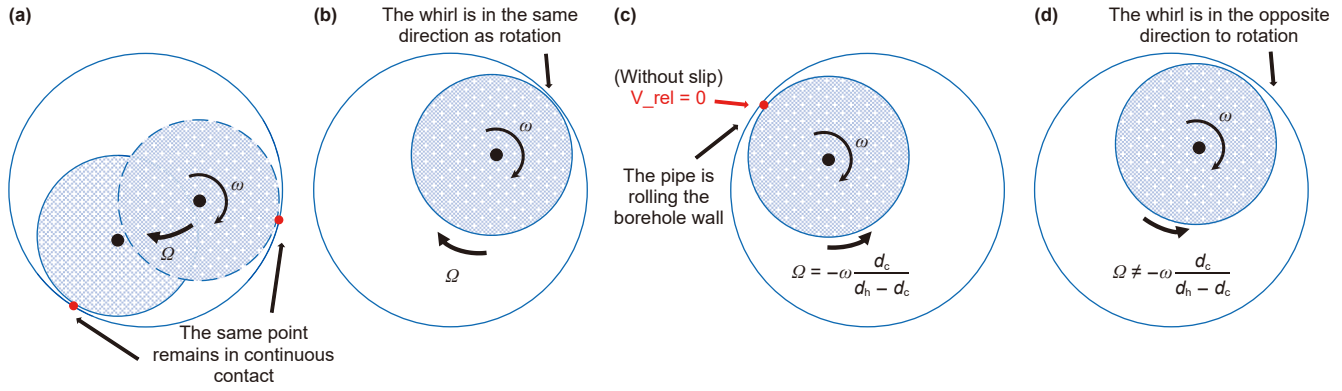


Fig. 4. Schematic illustration of four whirl modes in a borehole. (a) Synchronous forward whirl. (b) Forward whirl. (c) Synchronous backward whirl. (d) Backward whirl.

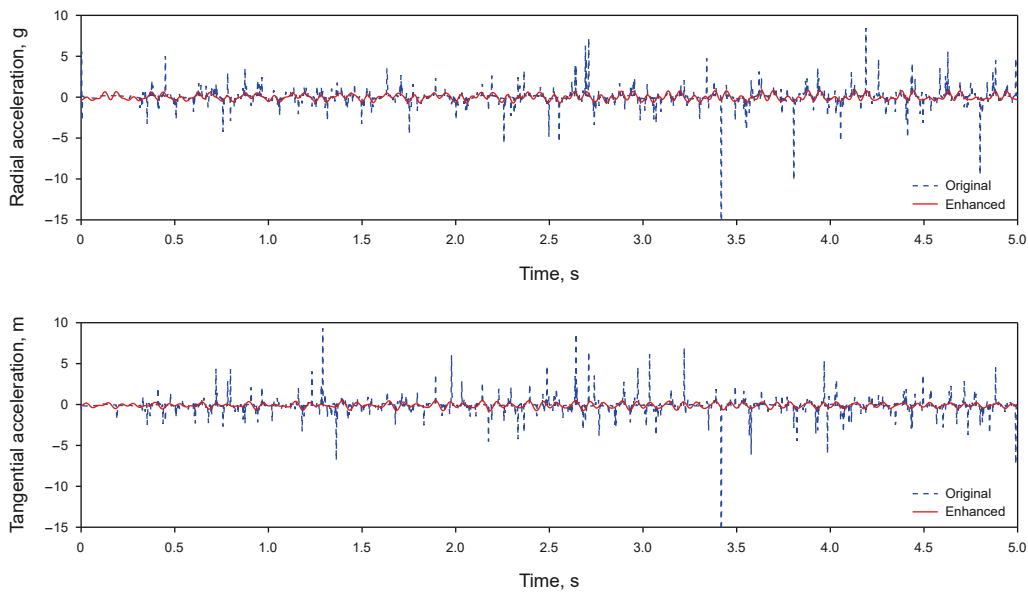


Fig. 5. Acceleration curves under backward whirl conditions.

mother wavelet, wavelet transform enables localized time–frequency analysis. It maps the signal onto a time–scale plane, where scale represents signal variation. Since scale and frequency are inversely related, small scales capture high-frequency components, while large scales correspond to low-frequency information (Sifuzzaman et al., 2009).

The CWT of a signal $f(t)$ involves translating and scaling the mother wavelet $\psi\left(\frac{t-b}{a}\right)$. The transform is defined as follows (Byeon et al., 2019):

$$\text{CWT}(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \times \psi\left(\frac{t-b}{a}\right) dt \quad (9)$$

where $f(t)$ is the original signal, $\psi\left(\frac{t-b}{a}\right)$ is the mother wavelet function, a is the scale parameter controlling the width of the wavelet—narrow at high frequencies and wide at low frequencies, and b is the translation parameter determining the position of the wavelet in time.

3.2. Morse wavelet

This study employs the Morse wavelet, a distinguished subgroup of analytic wavelets recognized for its superior time–frequency localization. It is extensively utilized in mechanical vibration analysis for the acquisition of transients and the management of non-stationary signals (Du et al., 2022; Toma et al., 2021). The frequency-domain representation is provided by the Fourier transform below (Lilly and Olhede, 2009; Byeon et al., 2019):

$$\Psi_{\beta, \gamma}(\omega) = U(\omega) a_{\beta, \gamma} \omega^{\beta} e^{-\omega^{\gamma}} \quad (10)$$

where $U(\omega)$ is the Heaviside step function, $a_{\beta, \gamma}$ is a normalization constant, defined as $a_{\beta, \gamma} \equiv 2(e\gamma/\beta)^{\beta/\gamma}$, γ controls the symmetry of the Morse wavelet, β is a parameter that governs decay and compactness;

The peak frequency of the wavelet is determined by $\omega_{\beta, \gamma}$, as given by the following equation (Lilly and Olhede, 2009):

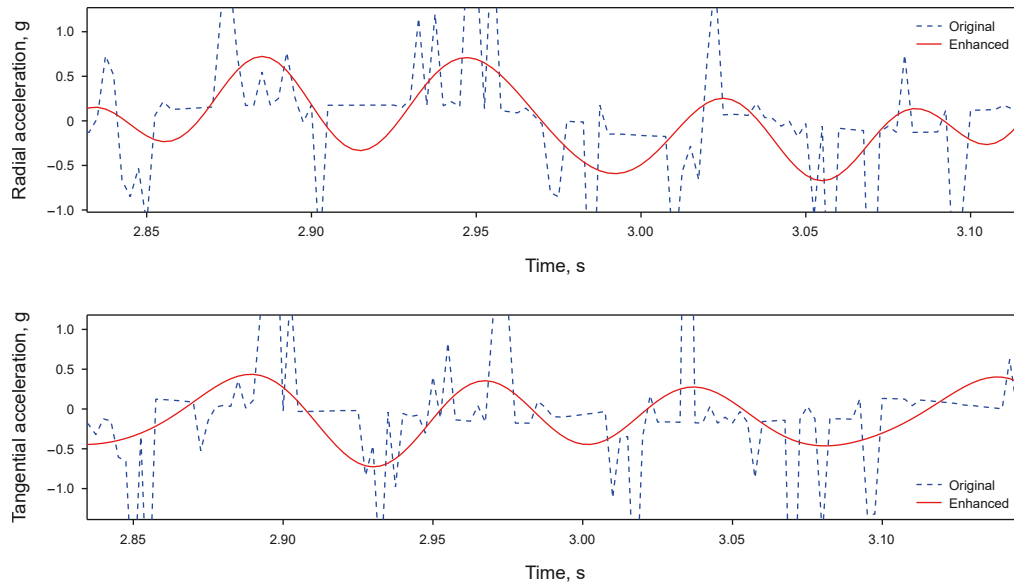


Fig. 6. Acceleration curve showing adjacent peak intervals.

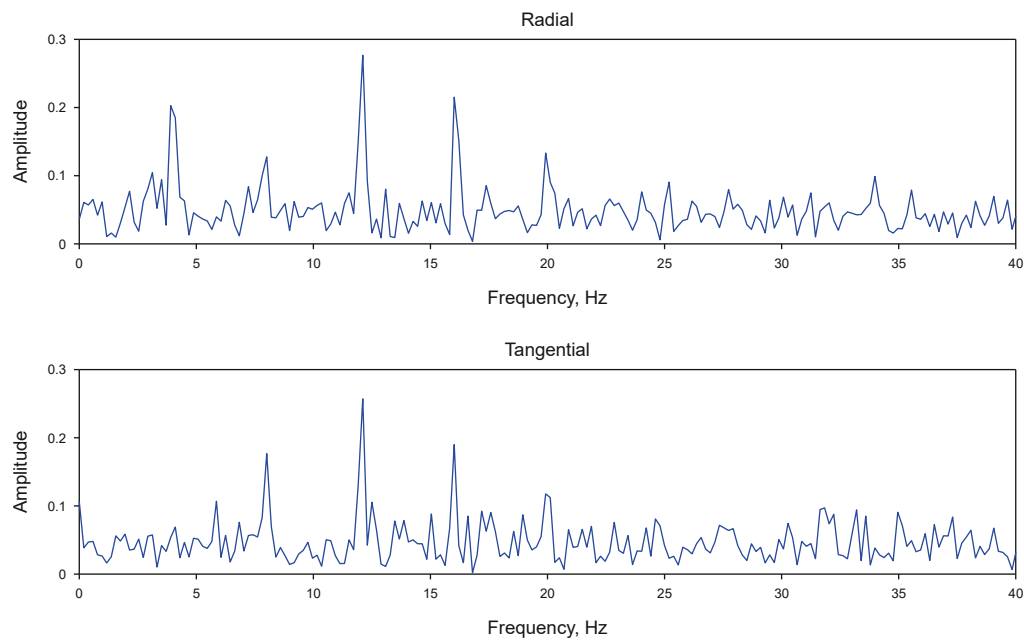


Fig. 7. Frequency spectra of radial and tangential acceleration.

$$\omega_{\beta,\gamma} = (\beta/\gamma)^{1/\gamma} \quad (11)$$

Fig. 13 illustrates the time-domain curves of the stick-slip acceleration signal alongside the associated wavelet coefficients produced by the Morse wavelet. The coefficients closely reflect the original signal, showing a high degree of resemblance. Consequently, the Morse wavelet is chosen as the mother wavelet for the wavelet transform in this research.

3.3. AlexNet convolutional neural network model

AlexNet is a deep convolutional neural network that attained significant success in the 2012 ImageNet competition. Fig. 14 illustrates that the design comprises five convolutional layers, three max-pooling layers, two fully connected layers, and one output layer. The dimensions of the input image are fixed at 227×227 pixels. The initial convolutional layer employs 96 filters measuring

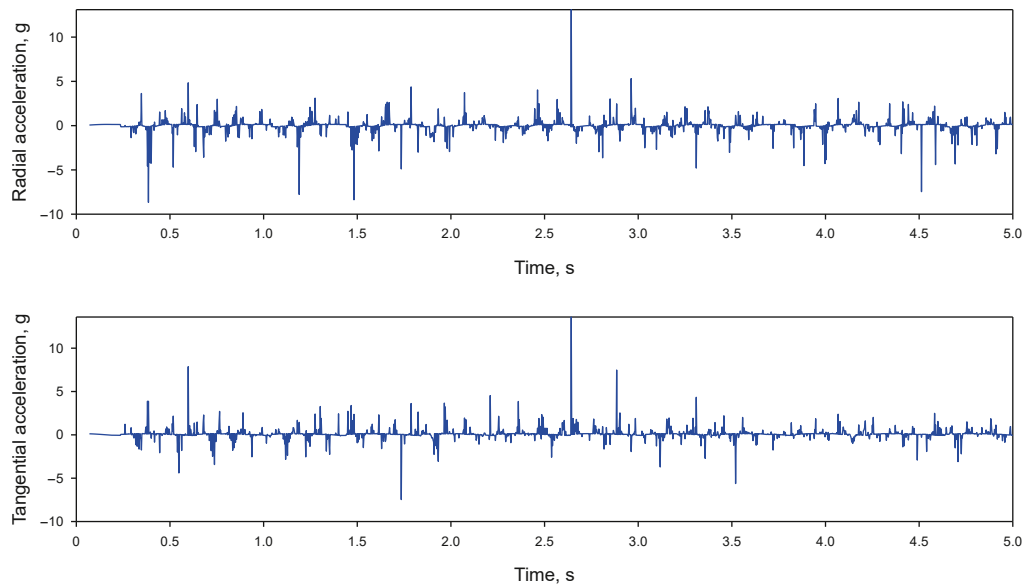


Fig. 8. Acceleration signal of backward whirl with slip.

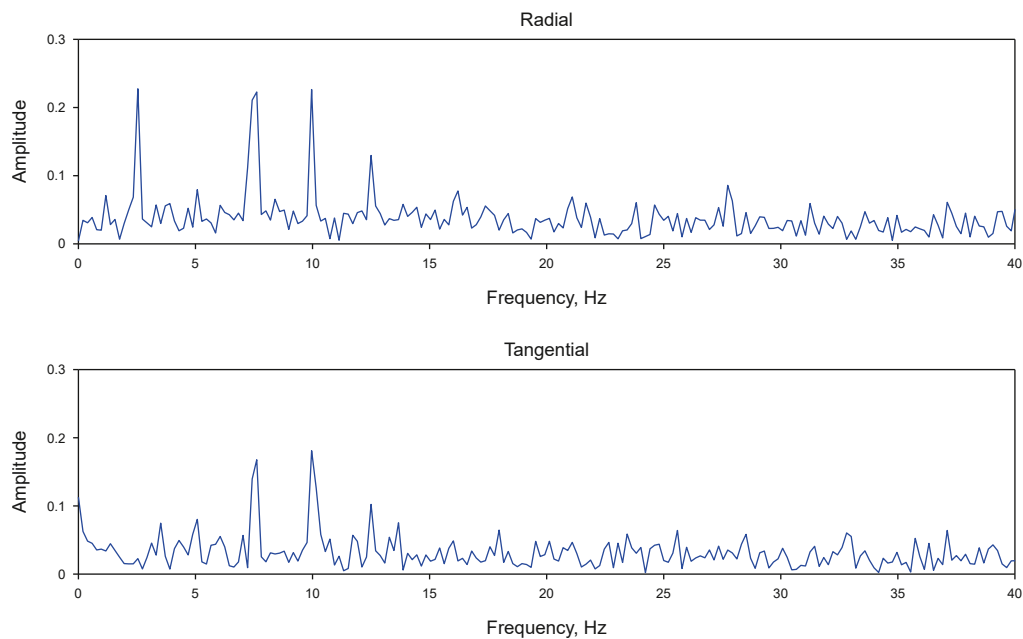


Fig. 9. Frequency spectrum of backward whirl acceleration with slip.

11×11 with a stride of 4, resulting in a feature volume of dimensions $55 \times 55 \times 96$. This is succeeded by a 3×3 max-pooling layer, which decreases the dimensions to $27 \times 27 \times 96$. The second convolutional layer employs 256 filters of dimensions 5×5 with a stride of 1, resulting in an output of $27 \times 27 \times 256$, subsequently down sampled to $13 \times 13 \times 256$ by the second pooling layer. The third and fourth convolutional layers utilize 384 filters of dimensions 3×3 with a stride of 1, preserving the spatial resolution at 13×13 . The fifth convolutional layer utilizes 256 filters of identical dimensions and is succeeded by a third max-pooling layer, which diminishes the output to $6 \times 6 \times 256$. The output is subsequently flattened into a one-dimensional vector of length 9216 and sent over two fully connected layers. The terminal output

layer executes categorization into 1000 categories utilizing the Softmax activation function (Krizhevsky et al., 2012).

3.4. Vibration recognition via CWT-AlexNet

This study proposes a method that integrates CWT with the AlexNet convolutional neural network for the classification of backward whirl and stick-slip vibrations. By directly using the scalograms generated from the multi-scale decomposition of CWT as input samples for model training, the approach eliminates the need for manual feature extraction. In contrast, traditional ML algorithms such as Decision Tree, Naïve Bayes, and Support Vector Machine (SVM) rely on a variety of statistical

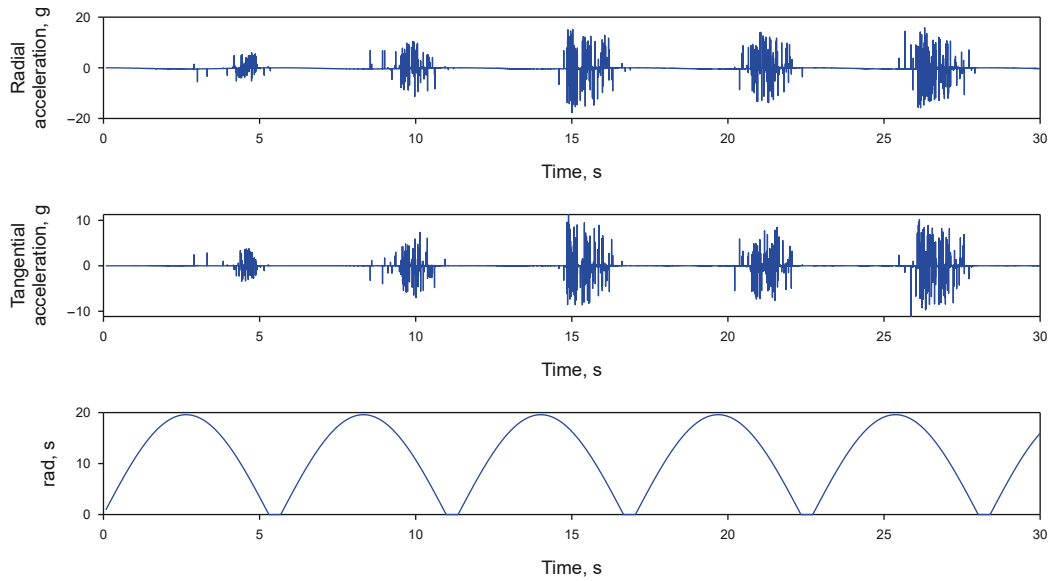


Fig. 10. Typical acceleration signal of stick-slip vibration.

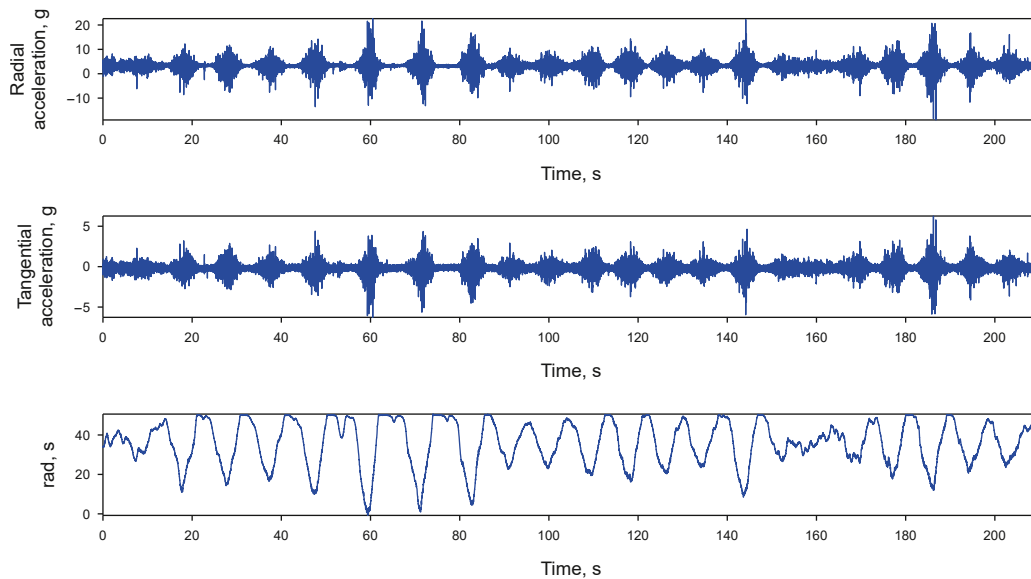


Fig. 11. Field-measured stick-slip acceleration signal sampled at 1000 Hz.

features—such as kurtosis, shape factor, skewness, and standard deviation. Experimental results demonstrate that the type and quantity of these features significantly affect classification accuracy, while often failing to fully capture the transient and dynamic characteristics of vibration signals. The proposed method eliminates the need for extracting numerous time and frequency-domain features, significantly simplifying the data preprocessing process. It removes the subjectivity of manual feature engineering, thereby enhancing classification accuracy and model stability.

In memory-constrained MWD/LWD systems, the suggested method, following INT8 quantization, attains a substantial reduction in model size to roughly 20–30 MB, with an inference delay of 30–50 ms and runtime memory utilization below 100 MB. These measures satisfy the essential criteria for downhole real-time deployment regarding latency, memory, and power usage.

4. Backward whirl and stick-slip classification

This section outlines a three-step workflow. Acceleration signals associated with whirl and stick-slip are first preprocessed and transformed into two-dimensional scalograms using continuous wavelet transform with the Morse wavelet. These scalograms are then used as inputs to the AlexNet model for training and classification, and the trained model is subsequently validated using simulation data generated by ABAQUS or field measurement data.

4.1. Signal preprocessing

To enhance recognition accuracy and model convergence, the raw acceleration signals are partitioned using fixed-length sliding windows to capture non-stationary vibration characteristics,

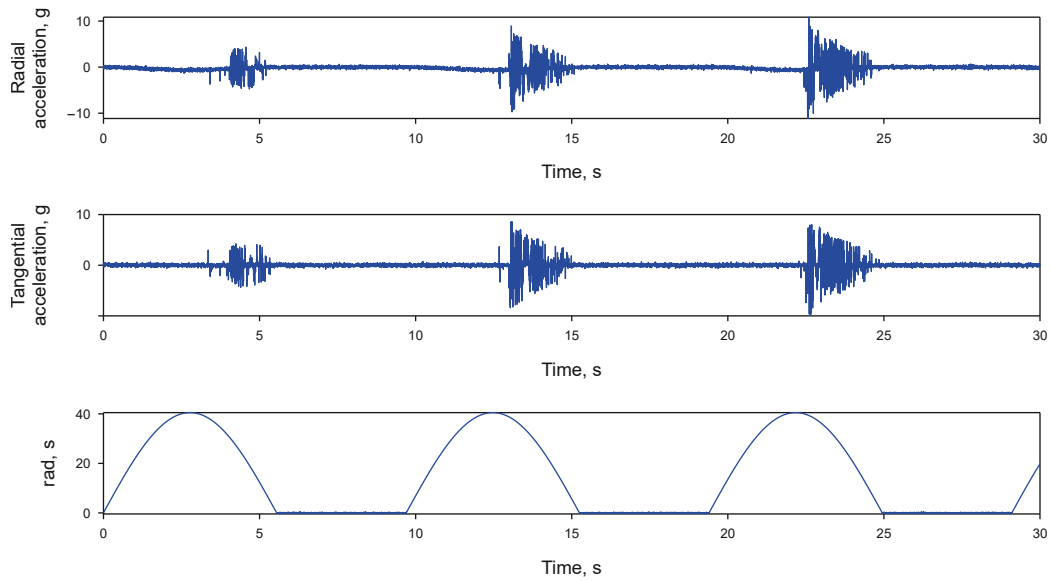


Fig. 12. Acceleration signal after adding low-amplitude random vibrations.

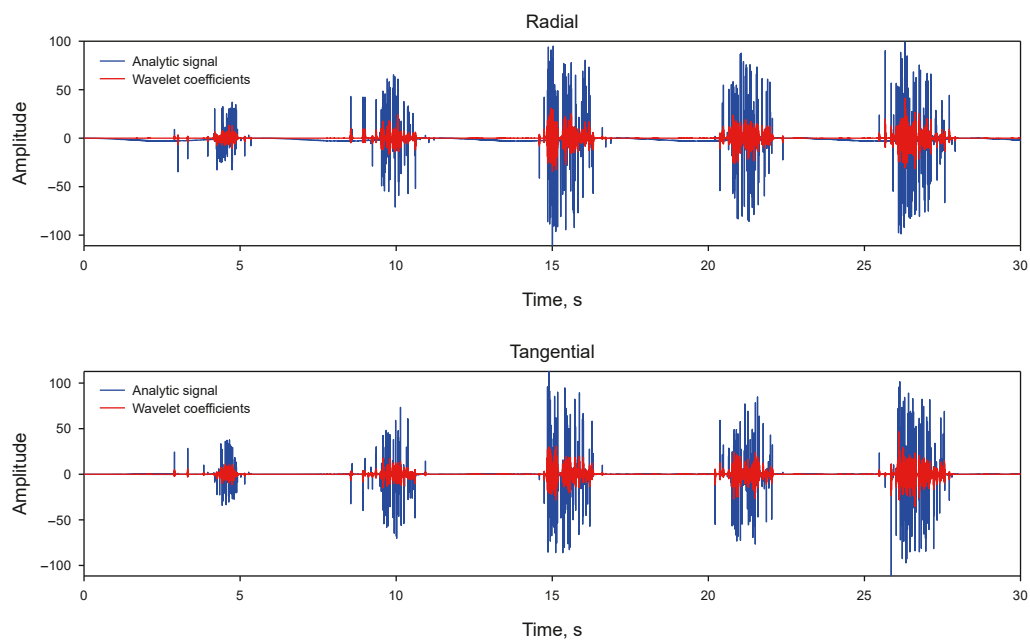


Fig. 13. Morse wavelet coefficients compared with the original signal.

ensuring that each segment encompasses at least one complete cycle of backward whirl or stick-slip, as illustrated in Fig. 15. Reckmann et al. (2018) indicate that a 5 s interval is adequate for capturing backward whirl; however, stick-slip necessitates a minimum duration of 20 s to maintain cycle integrity due to its lower frequency. To eradicate dimensional discrepancies, all segmentation signals are normalized to the [0,1] range by min-max scaling, thereby improving training stability and model generalization.

4.2. Scalogram transformation

A scalogram is a graphical depiction of a one-dimensional signal projected onto a two-dimensional time-frequency plane using CWT. It is based on the absolute values of the CWT coefficients and represents the energy distribution of the signal across time and frequency (indirectly shown by scale), facilitating multi-frequency analysis. Given that time-domain signals are finite and frequently exhibit noise, their analysis

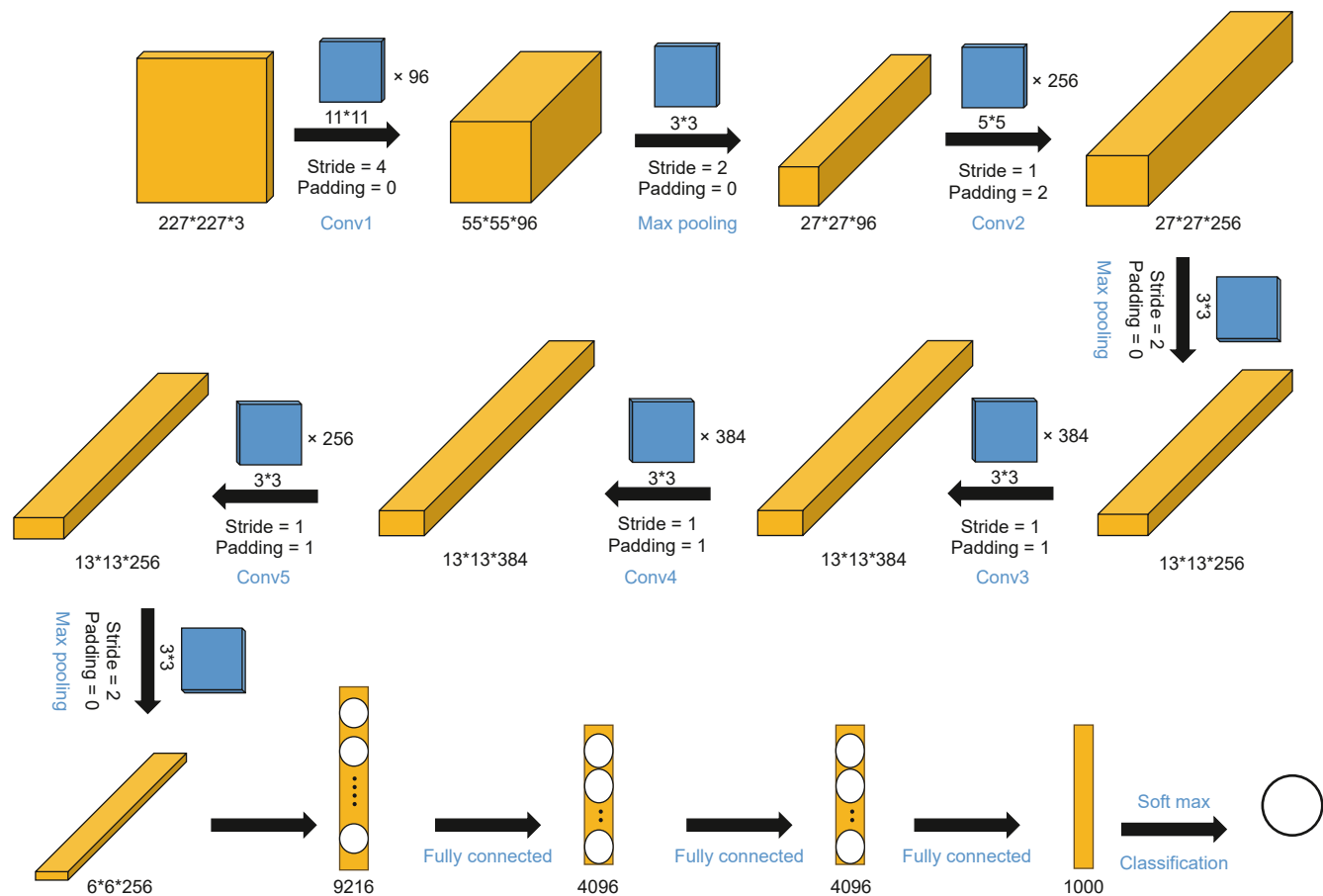


Fig. 14. Architecture of the AlexNet model.

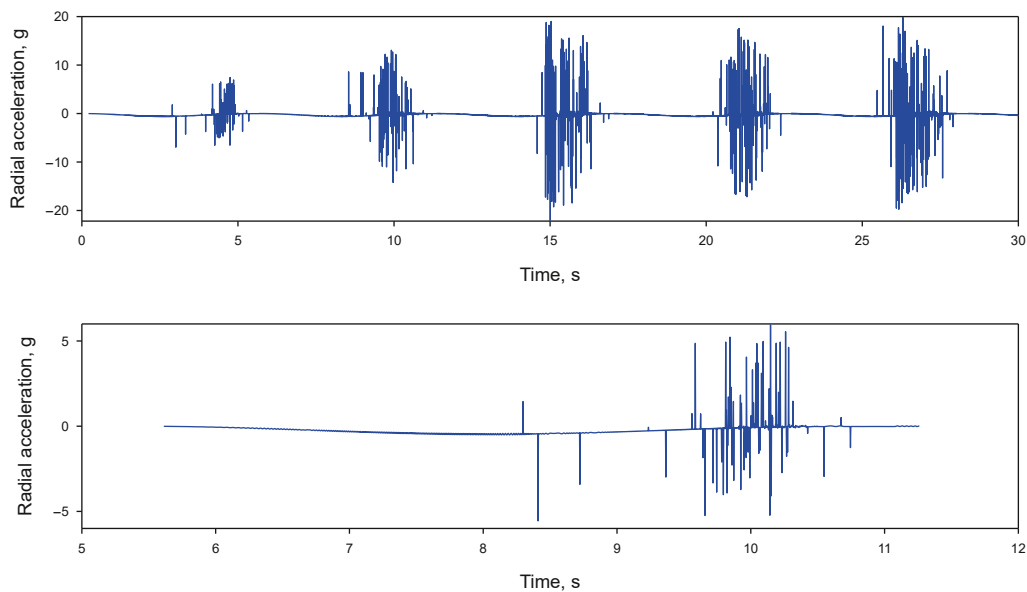


Fig. 15. Acceleration signals of stick-slip vibration before and after segmentation.

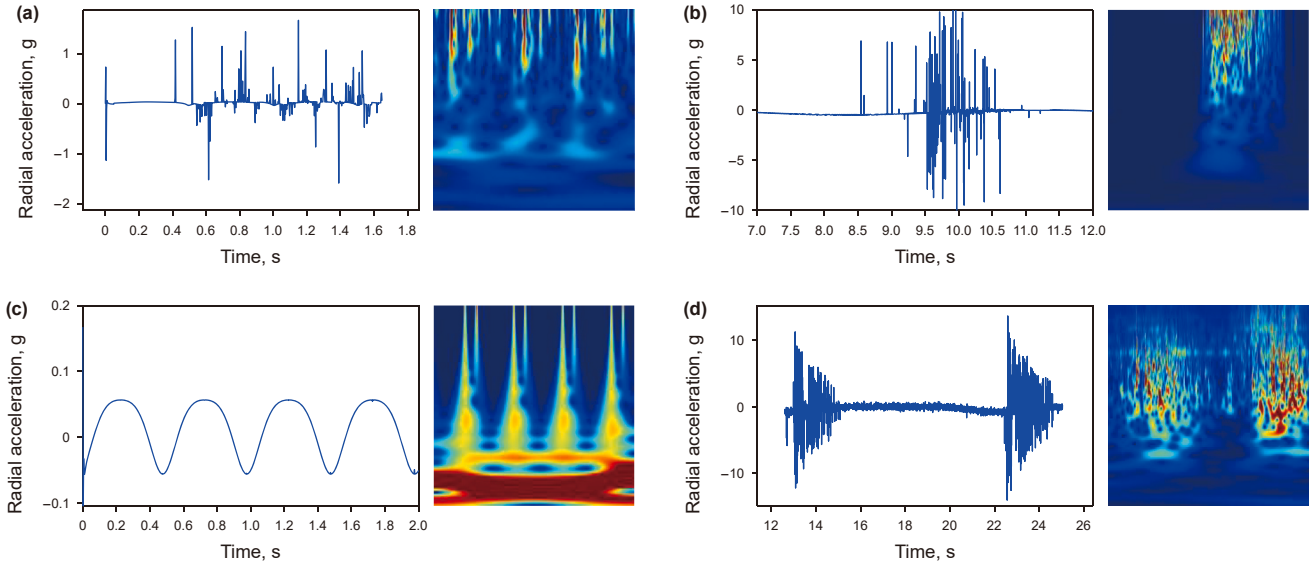


Fig. 16. Scalograms of different vibration types. (a) Acceleration of backward whirl. (b) Acceleration of typical stick-slip. (c) Acceleration under normal conditions. (d) Acceleration of pseudo stick-slip.

in the time-frequency domain provides a more lucid and succinct perspective on intricate behaviors. CWT offers a clear method for visualizing signal aspects across different scales and frequencies, facilitating the discovery of concealed characteristics. The scalograms are converted into RGB images with a resolution of 227×227 , utilizing the 128-color jet colormap for input into the CNN model. Fig. 16 illustrates scalograms of several vibration kinds. In typical circumstances, the scalogram demonstrates consistent amplitude fluctuations, but stick-slip signals reveal the distinctive "sticking-sliding" pattern.

4.3. Backward whirl recognition

In whirl signal detection, the data is categorized into two types: backward whirl (Bw) and non-backward whirl (Non Bw), with the Bw category encompassing both slip and no-slip instances. Each category comprises two attributes (radial and tangential acceleration), yielding 288 samples per feature, resulting in a total of 576 samples per class and 1152 overall. The model's performance is assessed using train-test splits of 8:2, 7:3, and 6:4. Fig. 17 illustrates the accuracy and loss trajectories for each scenario, while Fig. 18 presents the associated confusion matrices. The evaluation metrics comprise accuracy, sensitivity, specificity, false positive rate (FPR), and false negative rate (FNR), which are derived from true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), as delineated below (Choi et al., 2016).

$$\text{Accuracy} = \frac{CC}{CC + WC} \quad (12)$$

$$CC = TP + TN \quad (13)$$

$$WC = FP + FN \quad (14)$$

Here, accuracy refers to the comprehensive classification accuracy, reflecting the model's capacity to accurately classify all samples. CC denotes the count of accurately classified samples, WC signifies the number of inaccurately classified samples, TP

represents the number of positive samples accurately classified, TN indicates the number of negative samples accurately classified, FP refers to the number of negative samples erroneously classified as positive, and FN denotes the number of positive samples erroneously classified as negative.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (15)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (16)$$

$$\text{FRP} = \frac{FP}{FP + TN} \quad (17)$$

$$\text{Specificity} = \frac{FN}{FN + TP} \quad (18)$$

Table 2 illustrates that the model achieves the maximum validation accuracy with a training-to-testing ratio of 7:3. The relevant confusion matrix metrics are displayed in Table 3. The model has exceptional performance, with an overall accuracy of 96.25%, signifying substantial predictive reliability. A sensitivity of 96.02% indicates the model's proficiency in accurately identifying backward whirl instances, whereas a specificity of 96.48% demonstrates its effectiveness in correctly differentiating non-backward whirl cases. The false positive rate (FPR) and false negative rate (FNR) are both low, signifying a negligible rate of misclassification.

To enhance the validation of the categorization outcomes, a planar dynamic model was constructed in ABAQUS, as shown in Fig. 19. The model comprises two components: the outside circle represents the wellbore, defined as a rigid body, while the inner disc denotes the drill tool. A 0.002 kg mass was added to the periphery of the disc to establish a displacement between the center of mass and the geometric center. The disc possesses a material density of 1300 kg/m^3 , a Young's modulus of 3.5 GPa, a Poisson's ratio of 0.5, and a diameter of 6 mm; the wellbore diameter measures 7 mm. Fig. 20 shows the tangential acceleration data of backward whirl simulated by the model, which exhibit a certain level of noise and

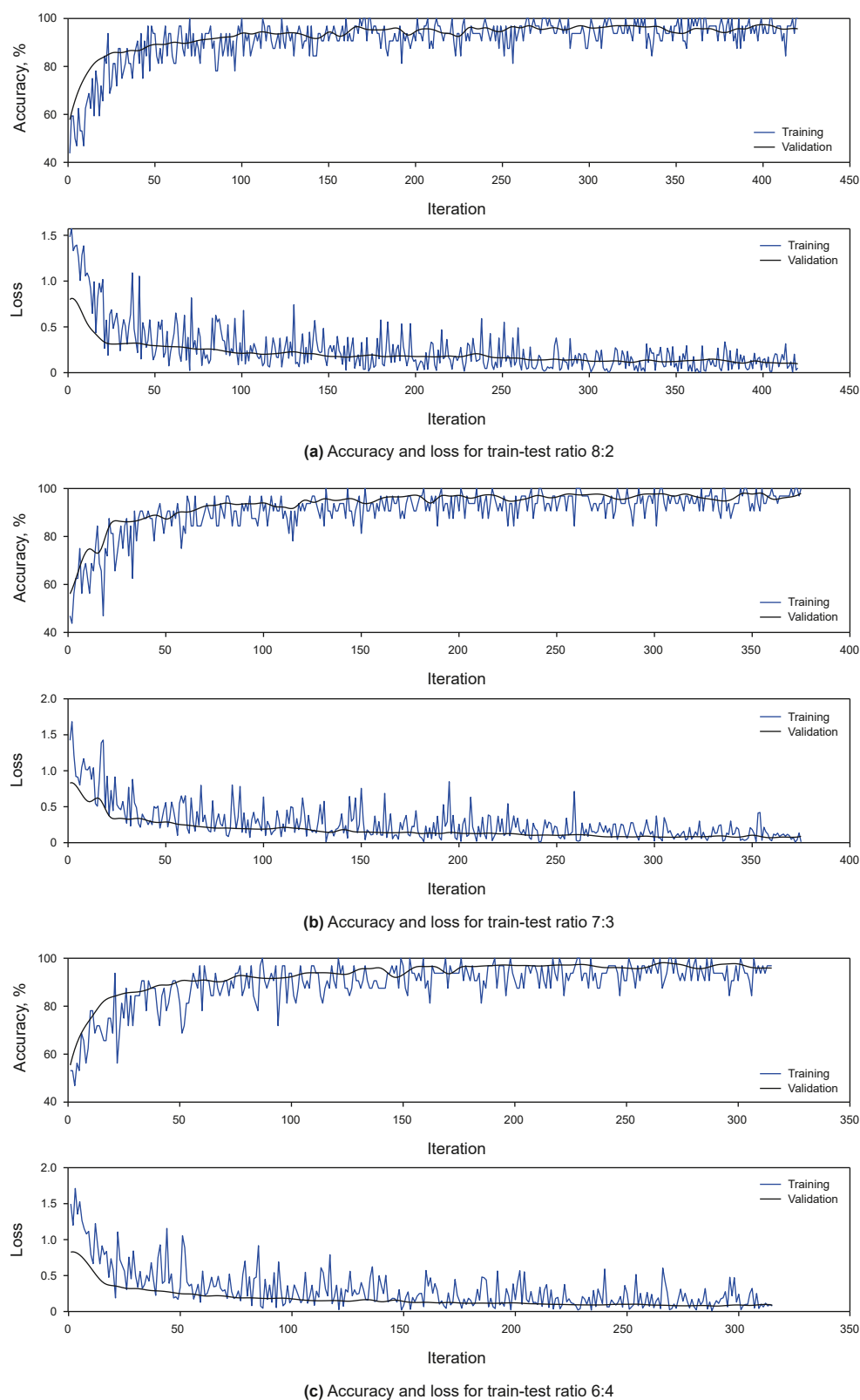


Fig. 17. Accuracy and loss curves for backward whirl signal classification under different data splits. **(a)** Accuracy and loss for train-test ratio 8:2. **(b)** Accuracy and loss for train-test ratio 7:3. **(c)** Accuracy and loss for train-test ratio 6:4.

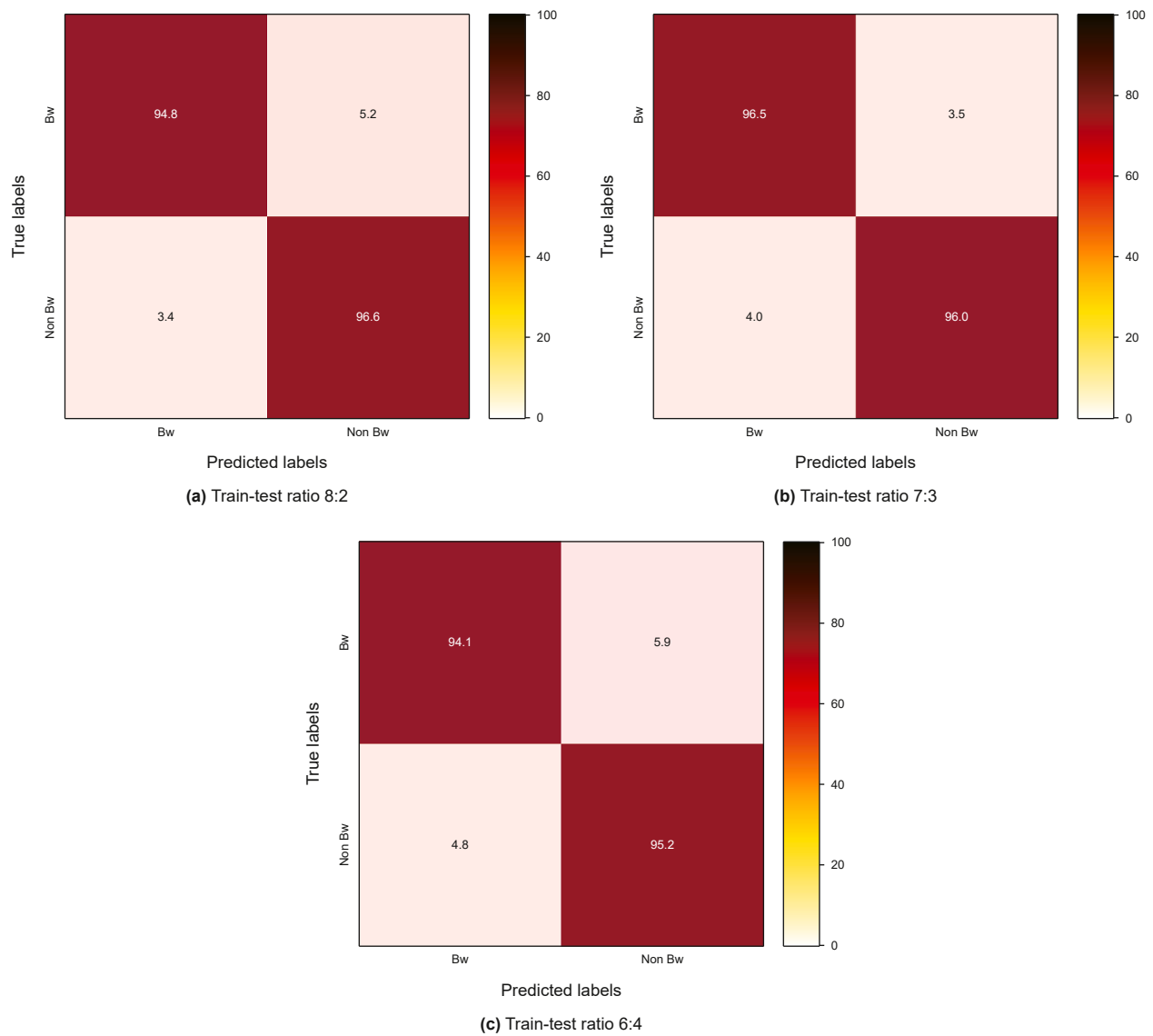


Fig. 18. Confusion matrices for backward whirl signal classification with different data splits. (a) Train-test ratio 8:2. (b) Train-test ratio 7:3. (c) Train-test ratio 6:4.

Table 2
AlexNet training results for backward whirl signals under different data splits.

Method	Train:Val	LR, %	Epoch	Batch	Solver	Functions	Accuracy, %
AlexNet	8:2	0.01	15	32	Sgdm	ReLU	95.62
AlexNet	7:3	0.01	15	32	Sgdm	ReLU	96.57
AlexNet	6:4	0.01	15	32	Sgdm	ReLU	94.38

Table 3
Confusion matrix metrics for 7:3 train–test ratio (%).

Accuracy	Sensitivity	Specificity	FPR	FNR
96.25	96.02	96.48	3.52	3.98

irregularity, thereby realistically reproducing the characteristics of the actual downhole environment. A total of 352 validation samples (176 backward whirl and 176 non-backward whirl) were utilized,

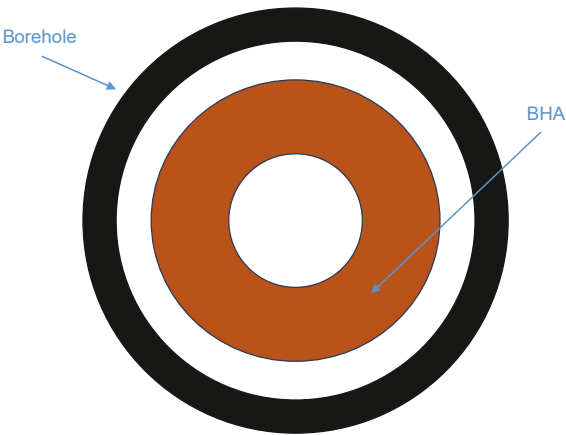


Fig. 19. Finite element rotor model.

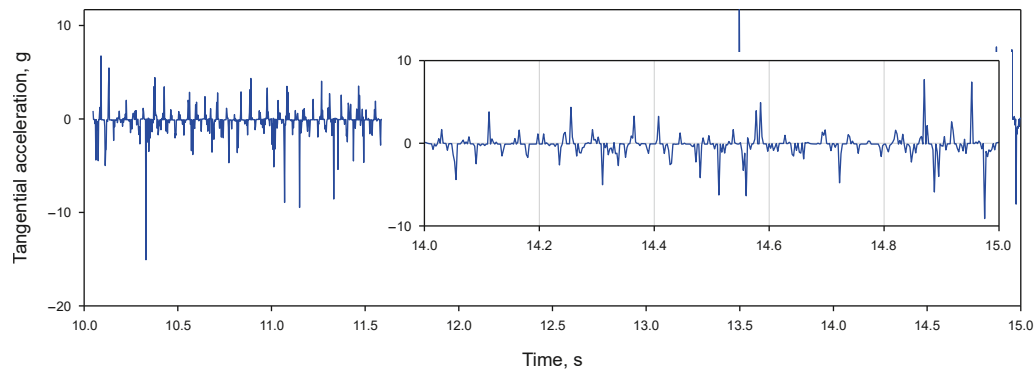


Fig. 20. Backward whirl acceleration signals simulated in ABAQUS (Wang et al., 2024).

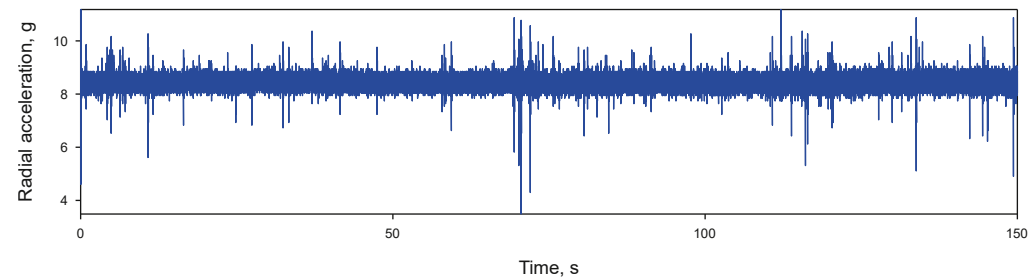


Fig. 21. Field-measured backward whirl acceleration signal.

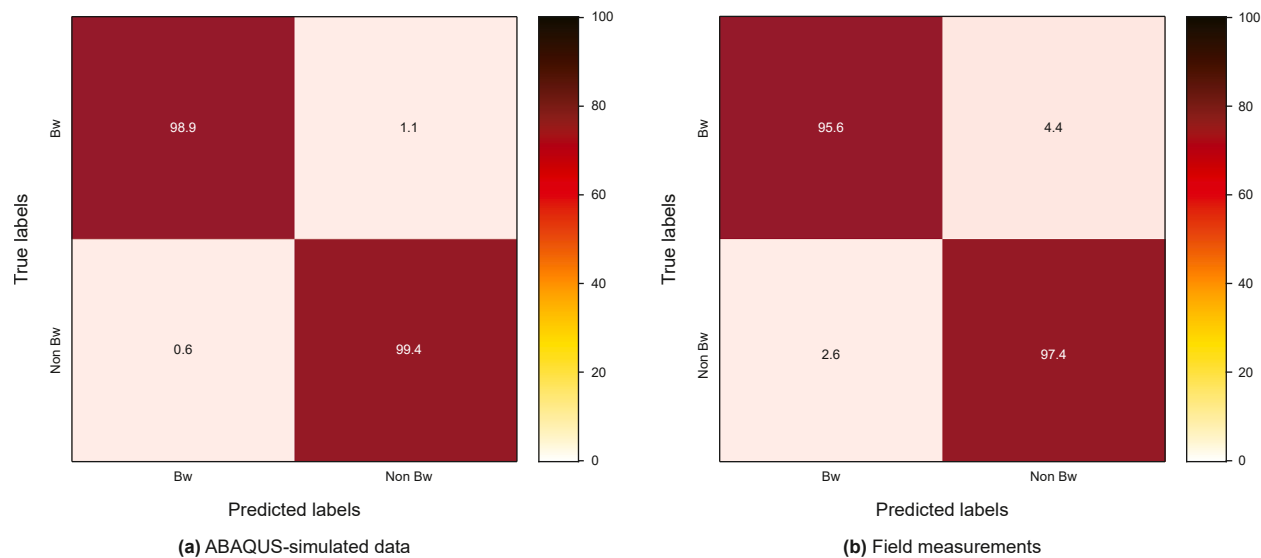


Fig. 22. Confusion matrix of the validation set. (a) ABAQUS-simulated data. (b) Field measurements.

Table 4
Comparative performance of ML classifiers for backward whirl identification (%).

Classifier	Accuracy	Recall	F1 Score
CWT + AlexNet	95.81	93.52	94.62
Decision tree	88.23	85.13	86.64
Naïve bayes	85.74	82.34	83.93
SVM	90.40	86.90	88.64

Table 5
Classification of stick-slip severity levels (Yezid et al., 2011)

SSI	Severity level
SSI <1.0	Non stick-slip
1.0< SSI <2.0	Low severity
2.0< SSI <2.5	Medium severity
SSI >2.5	High severity

representing 30% of the entire dataset. As shown in Fig. 22(a), the model had a classification accuracy of 99.15%.

In practical applications, backward whirl demonstrates pronounced periodic properties that remain mostly impervious to downhole ambient noise, allowing the trained model to sustain excellent accuracy. Fig. 21 illustrates that, despite the measured acceleration data exhibiting greater random noise than the simulation results, the model effectively captures its periodic characteristics. Fig. 22(b) presents the confusion matrix, demonstrating a classification accuracy of 96.50% under the specified field conditions.

We employed different ML classifiers to identify backward whirl, and Table 4 presents a comparative evaluation based on overall accuracy, typical recall, and typical F1 Score.

4.4. Stick-slip recognition

For stick-slip recognition, the stick-slip severity is first categorized based on the SSI, which is calculated using the following formula (Macpherson et al., 2015):

$$SSI = \frac{\omega_{\max} - \omega_{\min}}{\omega_{\text{avg}}} \tag{19}$$

where $\omega_{\max}$ is the maximum BHA rotational speed over a given time period, $\omega_{\min}$ is the minimum BHA rotational speed during the same period, and ω_{avg} is the average. BHA rotational speed over that period.

Based on field experience, stick-slip are categorized into four types, as shown in Table 5: non-stick-slip and three levels of severity—low, medium, and high, where each severity level includes both conventional stick-slip and pseudo stick-slip. During training, the non-stick-slip category is labeled as “Others”, while the stick-slip levels are labeled as “Low”, “Medium”, and “High”. The dataset is split into training and testing sets with an 8:2 ratio. For each type, 400 samples are used per feature, resulting in 800 samples per class and a total of 3200 samples. Fig. 23 shows the accuracy and loss curves for training and testing sets, and Fig. 24 presents the corresponding confusion matrix. Stick-slip vibration can be identified with high accuracy because, in practical drilling scenarios, stick-slip often manifests as a dominant vibration mode,

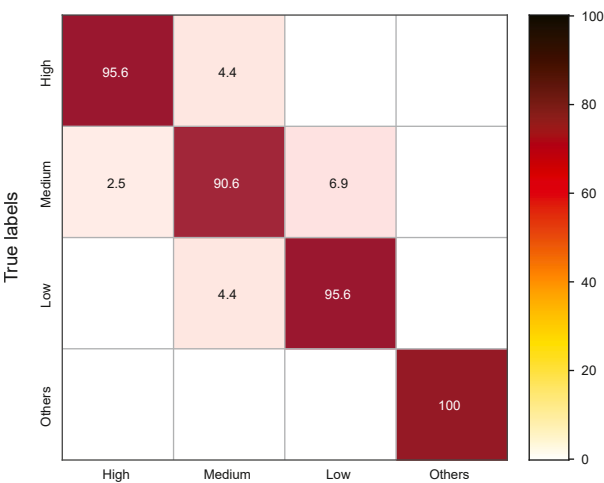


Fig. 24. Confusion matrix of the testing set.

Table 6
AlexNet training results for stick-slip signals.

Method	Train:Val	LR, %	Epoch	Batch	Solver	Functions	Accuracy, %
AlexNet	8:2	0.05	20	64	Sgdm	ReLU	95.67

Table 7
Confusion matrix metrics with an 8:2 train–test split (%).

Category	Accuracy	Sensitivity	Specificity	FPR	FNR
High	98.28	95.63	99.17	0.83	4.39
Medium	95.47	90.63	97.08	2.92	9.38
Low	97.19	95.63	97.71	2.29	4.38
Others	100.00	100.00	100	0.00	0.00

effectively overshadowing other concurrent vibration signals. As a result, during stick-slip events, the signal-to-noise ratio tends to be higher, with minimal interference from other vibrational components or external excitations, thereby enhancing the

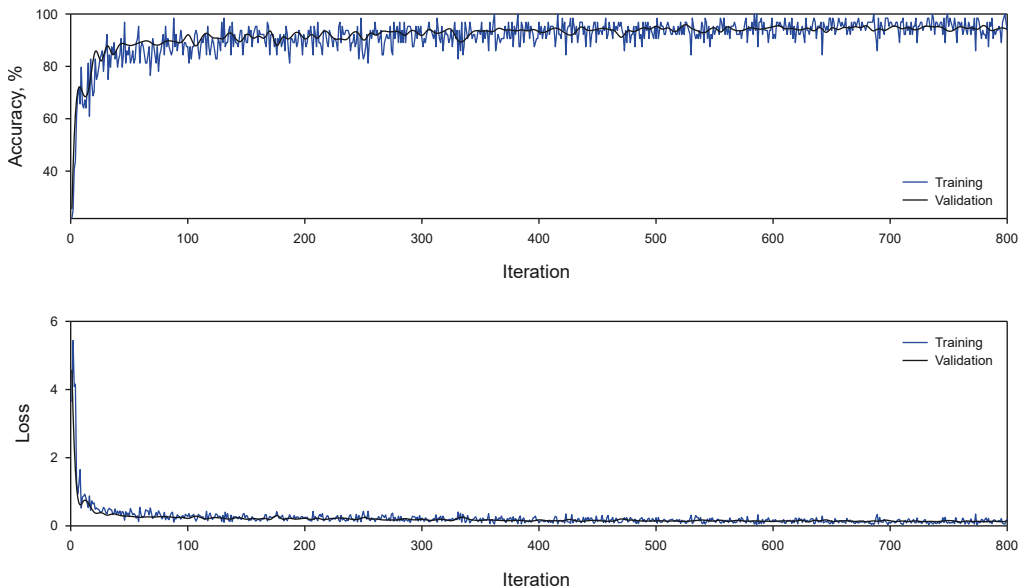


Fig. 23. Accuracy and loss curves of the training and testing sets.

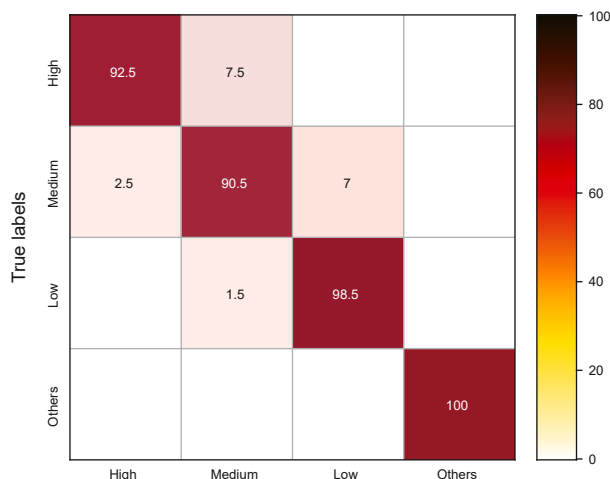


Fig. 25. Confusion matrix of validation set.

model's ability to accurately identify and classify stick-slip patterns.

Table 6 demonstrates that the achieved testing accuracy was 95.67%. The results of the confusion matrix, presented in Table 7, indicate that the “Others” category achieved the highest accuracy of 100%, reflecting a distinct contrast between non-stick-slip and stick-slip features. Accuracies for all stick-slip severity levels reached 95%, demonstrating considerable differences in their acceleration characteristics. The sensitivity for the “Others” category reached 100%, signifying that all positive samples were correctly identified. The specificity remained consistently high across categories, ranging from 97.08% to 100%, indicating strong effectiveness in reducing erroneous predictions. Low FPR and FNR values signify that the model produced few misclassifications, particularly with negative cases.

The trained model underwent additional validation for generalization utilizing field measurements, as illustrated in Fig. 11. The validation set comprised 200 samples per class, aligning with the size of the test set. Fig. 25 illustrates that the field data validation achieved an accuracy of 95.38%, illustrating the model's adaptability to intricate downhole conditions, encompassing fluctuations in formation, drilling fluid properties.

Upon completion of training, the finalized model can precisely classify unknown acceleration data. For stick-slip identification specifically, the model necessitates solely acceleration input, obviating the requirement for rotational speed measurements. This improves identification efficiency and substantially lowers measurement expenses.

5. Conclusions

This work proposes a method for classifying and identifying whirl and stick-slip using CWT and the AlexNet deep CNN. Acceleration data for whirl and stick-slip were generated using a single-disk rotor dynamics model, preprocessed, and transformed into CWT-based scalograms as input for the AlexNet model. Based on the results of this process, the following conclusions can be drawn.

- (1) In whirl recognition, three train-test split ratios were assessed, with the 7:3 split achieving the maximum accuracy of 96.57%. The remaining splits likewise sustained accuracy exceeding 94%, signifying consistent model performance and effective generalization. Sensitivity and specificity both attained 96%, confirming that acceleration features facilitate the accurate identification of backward whirl, even in slip situations.

- (2) In comparison to whirl, non-stick-slip was identified with 100% accuracy, indicating a clear differentiation in acceleration characteristics. This emphasizes acceleration as a key characteristic for differentiating stick-slip from non-stick-slip vibrations.
- (3) The model exhibited commendable performance on signals produced by ABAQUS, indicating that simulation data derived from the rotor dynamics model can be effectively utilized to categorize data from alternative sources. This signifies the model's robust generalization capabilities.
- (4) This work illustrates that whirl and stick-slip phenomena may be precisely identified and diagnosed with solely single-point acceleration data. This diminishes dependence on downhole sensors, substantially decreases the complexity and expense of downhole dynamic measures, and enhances overall drilling efficiency.

CRedit authorship contribution statement

Xue-Ying Wang: Writing – review & editing, Investigation, Funding acquisition, Formal analysis, Data curation. **Ming Liu:** Writing – original draft, Investigation, Funding acquisition, Formal analysis, Data curation. **Jian-Min Wang:** Formal analysis, Data curation. **Wei Li:** Resources, Project administration. **Fei-Fei Zhang:** Validation, Supervision, Software.

Conflict of interest

The authors assert that they possess no recognized competing financial interests or personal affiliations that may have seemingly influenced the work presented in this study.

Acknowledgements

This study was sponsored by the National Natural Science Foundation of China (Grant No. 52304002) and Qingdao Natural Science Foundation (Grant No. 24-4-4-zrjj-65-jch).

References

- Besaisow, A.A., Payne, M.L., 1988. A study of excitation mechanisms and resonances inducing bottomhole-assembly vibrations. *SPE Drill. Eng.* 3 (1), 93–101. <https://doi.org/10.2118/15560-pa>.
- Bailey, J.R., Biediger, E.A.O., Gupta, V., et al., 2008. Drilling vibrations modeling and field validation. In: IADC/SPE Drilling Conference, Richardson. Society of Petroleum Engineers, pp SPE 112650–MS. <https://doi.org/10.2118/112650-ms>.
- Baumgartner, T., van Oort, E., 2014. Pure and coupled drill string vibration pattern recognition in high frequency downhole data. In: SPE Annual Technical Conference and Exhibition, Richardson. Society of Petroleum Engineers, pp SPE 170955–MS <https://doi.org/10.2118/170955-ms>.
- Bowler, A., Harmer, R., Logesparan, L., et al., 2016. Continuous high-frequency measurements of the drilling process provide new insights into drilling-system response and transitions between vibration modes. *SPE Drill. Complet.* 31 (2), 106–118. <https://doi.org/10.2118/170713-pa>.
- Byeon, Y.H., Pan, S.B., Kwak, K.C., 2019. Intelligent deep models based on scalograms of electrocardiogram signals for biometrics. *Sensors* 19 (4), 935. <https://doi.org/10.3390/s19040935>.
- Choi, H.S., Lee, B., Yoon, S., 2016. Biometric authentication using noisy electrocardiograms acquired by mobile sensors. *IEEE Access* 4, 1266–1273. <https://doi.org/10.1109/ACCESS.2016.2548519>.
- Chen, S., Wisinger, J., Dunbar, B., et al., 2020. Identification and mitigation of friction- and cutting-action-induced stick/slip vibrations with PDC bits. *SPE Drill. Complet.* 35 (4), 576–587. <https://doi.org/10.2118/199639-pa>.
- Du, X., Jia, W., Yu, P., et al., 2022. A remaining useful life prediction method based on time-frequency images of the mechanical vibration signals. *Measurement* 202, 111782. <https://doi.org/10.1016/j.measurement.2022.111782>.
- Esmaili, A., Elahifar, B., Fruhwirth, R.K., et al., 2012. ROP modeling using neural network and drill string vibration data. In: SPE Kuwait International Petroleum Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/163330-ms>.

- Gowida, A., Elkatatny, S., 2024. Novel machine learning approach for predicting the gel strength of the drilling fluid while drilling. In: Offshore Technology Conference. OTC. <https://doi.org/10.4043/35433-ms>.
- Gabry, M.A., Ye, Z., Ramadan, A., et al., 2025a. Calibration of continuous wavelet transform (CWT) fracture closure detection using laboratory hydraulic fracturing experiments. In: Proceedings of the 2025 Unconventional Resources Technology Conference. American Association of Petroleum Geologists. <https://doi.org/10.15530/urtec-2025-4221618>.
- Gabry, M.A., Ramadan, A., Soliman, M.Y., 2025b. Hydraulic fracture characterization using continuous wavelet transform for treating pressure, calibrated with fiber optics data. In: SPE/AAPG/SEG Unconventional Resources Technology Conference (URTeC). <https://doi.org/10.15530/urtec-2025-4203446>.
- Gabry, M.A., Ramadan, A., Soliman, M.Y., 2025c. Innovative method for analyzing inter-well connectivity with continuous wavelet transform. In: Offshore Technology Conference. <https://doi.org/10.4043/35951-ms>.
- Jones, S., Sugiura, J., Rose, K., et al., 2017. Drilling dynamics data recorders now cost-effective for every operator - compact embedded sensors in bit and BHA capture small data to make the right decisions fast. In: SPE/IADC Drilling Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/184738-ms>.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, vol. 25. NIPS, pp. 1097–1105. <https://doi.org/10.1145/3065386>.
- Lilly, J.M., Olhede, S.C., 2009. Higher-order properties of analytic wavelets. IEEE Trans. Signal Process. 57 (1), 146–160. <https://doi.org/10.1109/TSP.2008.2007607>.
- Minett-Smith, D.J., Stroud, D.R., Pagett, J.M., 2010. Real-time whirl detector aids drilling optimization. In: SPE Annual Technical Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/135110-ms>.
- MacPherson, J.D., Paul, P., Behounek, M., et al., 2015. A framework for transparency in drilling mechanics and dynamics measurements. In: SPE Annual Technical Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/174874-ms>.
- Millan, E., Ringer, M., Boualleg, R., et al., 2019. Real-time drillstring vibration characterization using machine learning. In: SPE/IADC International Drilling Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/194061-ms>.
- Moya, C., Prado, R., Caycedo, H., et al., 2024. Machine learning-assisted prediction of electrical properties for heterogeneous reservoirs: case study in mamey field, northern Colombia. In: International Petroleum Technology Conference. IPTC. <https://doi.org/10.2523/iptc-23381-ms>. D031S124R006.
- Nussbaumer, H.J., 1982. The fast Fourier transform. In: *Fast Fourier Transform and Convolution Algorithms*. Springer, Berlin, Heidelberg, pp. 80–111.
- Olukoga, T.A., Feng, Y., 2021. Practical machine-learning applications in well-drilling operations. SPE Drill. Complet. 36 (4), 849–867. <https://doi.org/10.2118/205480-pa>.
- Popp, T., Stibbe, H., Heinisch, D., et al., 2018. Backward whirl testing and modeling with realistic borehole contacts for enhanced drilling tool reliability. In: IADC/SPE Drilling Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/189600-ms>.
- Reckmann, H., Hohl, A., Oueslati, H., et al., 2018. Step change in BHA reliability and performance due to utilization of downhole drilling dynamics services. In: IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/191096-ms>.
- Shyu, R.J., 1989. Bending Vibration of Rotating Drill Strings. Massachusetts Institute of Technology, Cambridge, MA, USA. Ph.D. dissertation. <http://hdl.handle.net/1721.1/14205>.
- Safavian, L.S., Kinsner, W., Turanli, H., 2005. Aquantitative comparison of different mother wavelets for characterizing transients in power systems. In: Canadian Conference on Electrical and Computer Engineering, 2005. IEEE, pp. 1461–1464. <https://doi.org/10.1109/CCECE.2005.1557253>.
- Sifuzzaman, M., Islam, M.R., Ali, M.Z., 2009. Application of wavelet transform and its advantages compared to Fourier transform. J. Phys. Sci. 13, 121–134. <https://ir.vidyasagar.ac.in/handle/123456789/779>.
- Stroud, D.R., Lines, L.A., Minett-Smith, D.J., 2011. Analytical and experimental backward whirl simulations for rotary steerable bottom hole assemblies. In: SPE/IADC Drilling Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/140011-ms>.
- Sheng, G., Li, Z., Wang, W., 2015. A new automatic detection method of micro-seismic event based on CWT and HOS. In: SEG Technical Program Expanded Abstracts 2015. Society of Exploration Geophysicists, pp. 2625–2629. <https://doi.org/10.1190/segam2015-5881516.1>.
- Saadeldin, R., Gamal, H., Elkatatny, S., 2023. Detecting downhole vibrations through drilling horizontal sections: machine learning study. Sci. Rep. 13, 6204. <https://doi.org/10.1038/s41598-023-33411-9>.
- Townsend, T., Moss, W., Heinisch, D., et al., 2021. Advanced high frequency in-bit vibration measurement including independent, spatially separated sensors for proper resolution of vibration components including lateral, radial, and tangential acceleration. In: Abu Dhabi International Petroleum Exhibition & Conference, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/208110-ms>.
- Toma, R.N., Toma, F.H., Kim, J.M., 2021. Comparative analysis of continuous wavelet transforms on vibration signal in bearing fault diagnosis of induction motor. In: 2021 International Conference on Electronics, Communications and Information Technology (ICECIT). IEEE, pp. 1–4. <https://doi.org/10.1109/icecit54077.2021.9641199>.
- Wang, X., Wang, X., Lu, J., et al., 2024. Automated classification of drill string vibrations using machine learning algorithms. Geoenergy Sci. Eng. 239, 212995. <https://doi.org/10.1016/j.geoen.2024.212995>.
- Yezid, I.A., Yonnellybeth, M., Andre, N., 2011. Vibration risk index offers tool for preventing drillstring failure. In: World Trends and Technology for Offshore Oil and Gas Operations. Offshore Magazine. <https://www.offshore-mag.com/deepwater/article/16755347/quantifying-drilling-vibration-challenges>.
- Zannoni, S.A., Cheatham, C.A., Chen, C.D., et al., 1993. Development and field testing of a new downhole MWD drillstring dynamics sensor. In: SPE Annual Technical Conference and Exhibition, Richardson. Society of Petroleum Engineers. <https://doi.org/10.2118/26341-ms>.
- Zhu, X., Liu, W., 2013. The effects of drill string impacts on wellbore stability. J. Petrol. Sci. Eng. 109, 217–229. <https://doi.org/10.1016/j.petrol.2013.08.004>.
- Zhao, C., Jiang, L., Wang, X., et al., 2021. Prestack seismic noise attenuation based on 3D CWT. In: First International Meeting for Applied Geoscience & Energy Expanded Abstracts. Society of Exploration Geophysicists, pp. 2834–2838. <https://doi.org/10.1190/segam2021-3594268.1>.