

Article

Data-Driven Fault Prediction for Electric Submersible Progressing Cavity Pump Wells

Cong Zhang ^{1,2}, Guoqing Han ^{3,*} , Liping Zhao ², Chunsheng Wu ², Bin Fan ², Bin Liu ², Zezhou Li ² and Jiping Deng ³

¹ School of Mineral Resource and Geoscience, China University of Mining and Technology, Xuzhou 221116, China

² Shanxi Coalbed Methane Exploration and Development Branch, PetroChina Huabei Oilfield Company, Changzhi 046000, China; mcq_wcs@petrochina.com.cn (C.W.); mcq_lzz@petrochina.com.cn (Z.L.)

³ College of Petroleum Engineering, China University of Petroleum-Beijing, Beijing 102249, China

* Correspondence: hanguoqing@163.com

Abstract

During the development of coalbed methane (CBM), electric submersible progressing cavity pumps (ESPCPs) face challenges such as handling gas–liquid mixtures, high water content, declining pump efficiency, and frequent failures. This study proposes a data-driven fault prediction and diagnostic method, analyzing various factors affecting pump performance, such as gas–liquid ratio, water content of extracted fluids, pumping depth, and current fluctuations, to explore their correlation with pump failures. The dataset used in this study was collected from 85 ESPCP wells in the Northern Zhachi Oilfield between January 2019 and December 2022, with daily acquisition of key operational parameters including tubing pressure, pump current, vibration, temperature, produced liquid rate, and gas–liquid ratio. Baseline models including ARIMA and Gradient Boosting Decision Trees (GBDT) were trained for performance comparison, with the proposed PCA–LSTM model achieving a 15% and 9% improvement in validation accuracy over ARIMA and GBDT, respectively. Model performance is quantitatively reported: acceptance rate 86.79% and validation accuracy 72.22%. The results indicate that the PCA–LSTM model can effectively identify and predict ESPCP failure types, providing vital technical support for the efficient operation and maintenance of CBM wells, and its practical applicability has been well recognized by field engineers.

Keywords: electric submersible progressing cavity pumps (ESPCPs); fault prediction; data-driven; principal component analysis (PCA); long short-term memory (LSTM) networks



Academic Editors: Wei Sun and Maoyin Chen

Received: 24 July 2025

Revised: 23 August 2025

Accepted: 4 September 2025

Published: 10 September 2025

Corrected: 16 December 2025

Citation: Zhang, C.; Han, G.; Zhao, L.; Wu, C.; Fan, B.; Liu, B.; Li, Z.; Deng, J. Data-Driven Fault Prediction for Electric Submersible Progressing Cavity Pump Wells. *Processes* **2025**, *13*, 2890. <https://doi.org/10.3390/pr13092890>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the development of coalbed methane (CBM), the presence of complex gas–water mixtures and high–water–content gas wells presents significant challenges to the dewatering and gas extraction tasks. This is especially true in the case of deep wells and high gas–liquid ratio conditions, where traditional electric submersible pumps (ESPs) and hydraulic pumps often fail to operate efficiently and stably. This results in a decline in pump performance, frequent failures, and a severe impact on production efficiency and gas extraction rates. As a result, the adoption of Electric Submersible Progressing Cavity Pumps (ESPCPs), which are better suited to handle complex operational conditions, has gradually become a key area of research within the industry for dewatering and gas extraction in CBM wells.

Previous studies on ESPCP fault prediction often focus on either statistical methods or black-box deep learning models without explicit comparison to baselines or reproducible datasets. Our work addresses this gap by integrating PCA for feature reduction and interpretability, with LSTM for temporal dependency modeling, validated against both statistical (ARIMA) and machine learning (GBDT) baselines. This positions our study as a step towards bridging conventional and modern approaches while maintaining field-deployable efficiency.

As a type of positive displacement pump, the progressing cavity pump offers several significant advantages, including a simple structure, smooth operation, and strong adaptability to gas–liquid mixtures. Especially in CBM wells with high gas–water ratios, ESPCPs are more effective than traditional pumps at mitigating the interference caused by gas, thereby ensuring stable dewatering capacity and improving overall pumping efficiency [1,2]. However, despite their advantages, ESPCPs still face numerous challenges during extended periods of high-load operation, including stator wear, tubing corrosion, and incomplete gas–liquid separation. When operating at greater depths, additional issues arise [3], such as increased torque, higher rod failure rates, equipment perforation, and stator failure. These factors can lead to a reduction in pump performance, system failures, and equipment damage.

Given this context, the ability to accurately diagnose and predict faults in ESPCPs used in CBM wells through advanced fault detection techniques has become crucial for improving CBM development efficiency and ensuring safe production [4]. Our research, based on the operational characteristics of ESPCPs in CBM wells, analyzes a range of factors that affect pump performance, including gas–liquid ratio, the water content of the extracted fluids, pumping depth, and current fluctuations. Furthermore, the paper explores the relationships between these factors and pump failures. By integrating artificial intelligence (AI) techniques, data mining, and model training, this study proposes a fault prediction and diagnostic method tailored for ESPCPs in CBM wells, providing valuable technical support for the efficient operation and maintenance of CBM well equipment.

2. Background

Coalbed methane (CBM), as an important unconventional natural gas resource, has attracted considerable attention in the global energy industry in recent years. Its development potential is immense, and it is expected to become one of the key energy sources in the coming decades, especially in the context of growing global demand for low-carbon and sustainable energy. CBM not only provides abundant energy resources for China but also contributes to the diversification and security of global energy supply.

During the long-term extraction of CBM, the issue of liquid accumulation has gradually emerged as a key factor limiting the sustained increase in production capacity. As production continues over time, horizontal wells in CBM fields often enter a phase of liquid accumulation, where water and CBM mix within the wellbore and accumulate as liquid, severely affecting gas flow and recovery efficiency. In CBM horizontal well production, the liquid accumulation problem is typically caused by continuous water influx and the gradual decrease in gas production, especially during the later stages of gas well production, when the discharge of wellbore liquids becomes increasingly difficult.

When traditional liquid management methods, such as speed tubing and foam drainage, fail to effectively address the liquid accumulation problem, the Electric Submersible Progressing Cavity Pump (ESPCP) emerges as an effective solution [5]. The ESPCP is a highly efficient mechanical pumping device capable of continuously and effectively displacing accumulated liquid at greater depths and higher-pressure conditions, restoring CBM production to normal levels. Compared to other liquid handling technologies, the

ESPCP offers greater adaptability and flexibility, providing continuous liquid discharge support even when gas production remains at relatively high levels but is constrained by liquid accumulation. This enables CBM wells to resume production effectively [6].

Particularly in wells that have already entered a severe liquid accumulation stage, the application of ESPCPs not only efficiently removes the accumulated liquids but also ensures sustained high production capacity from the gas well [7,8]. By carefully selecting the installation location and operating parameters of the ESPCP, it is possible to ensure the efficient discharge of liquids from the wellbore and maximize the recovery of gas production capacity, thereby providing a more stable assurance for the development and utilization of CBM resources.

3. Data Processing and Database Establishment

However, all types of pumps face the issue of operational failures during their use. According to operational data, the mean time between failures (MTBF) of most electric submersible progressing cavity pumps is less than one year. This highlights that the effective service life of the pumps needs to be extended [2]. Additionally, this leads to associated issues, such as lower pump efficiency, indicating that there is still significant room for improvement in liquid discharge efficiency [9,10].

Data acquisition covered 85 wells across varied operational environments in the North Buzachi Oilfield, spanning four years (January 2019–December 2022). Sensors captured multi-modal parameters: tubing pressure (kPa), pump current (A), vibration amplitude (mm/s), motor temperature (°C), produced liquid rate (m³/day), and calculated gas–liquid ratio (dimensionless). This level of detail ensures the dataset is fully reproducible for future research.

Frequent failures of the tubing and ESPCPs are among the key factors contributing to the shortening of the MTBF and the decline in liquid discharge efficiency [11]. Therefore, it is crucial to delve deeper into the failure trends within the production data, particularly by accurately correlating production parameters with specific failure types [12,13].

To address this, we conduct a detailed classification of the failures occurring in CBM wells utilizing ESPCPs for liquid drainage and gas production. In addition, statistical data analysis methods are employed to detect data anomalies, followed by the cleaning and completion of abnormal data points [14]. By incorporating specific information from well maintenance records, natural language processing (NLP) techniques are used to effectively analyze, learn from, and extract failure-related information [15,16]. Ultimately, this study clarifies the various types of failures encountered in the system.

This translation conveys the technical details in a clear and professional tone, maintaining the focus on the analysis and classification of failures in ESPCPs within coalbed methane wells, as well as the methods used for data analysis and fault identification.

In our research, we focused on the extraction of failure labels, with particular emphasis on the high-frequency and distinctly characteristic records of tubing perforations and corrosion. Through data mapping, we performed in-depth extraction of these operational conditions [17]. Additionally, we constructed a condition mapping data table that can be defined and expanded in real-time, ensuring the flexibility and scalability of the system. Finally, based on the production data from wells utilizing electric submersible progressing cavity pumps for liquid drainage, we established a failure database for ESPCP wells, which includes failure data and their corresponding labels from the past four years [18,19]. The data proceeding procedure is shown in Figure 1.

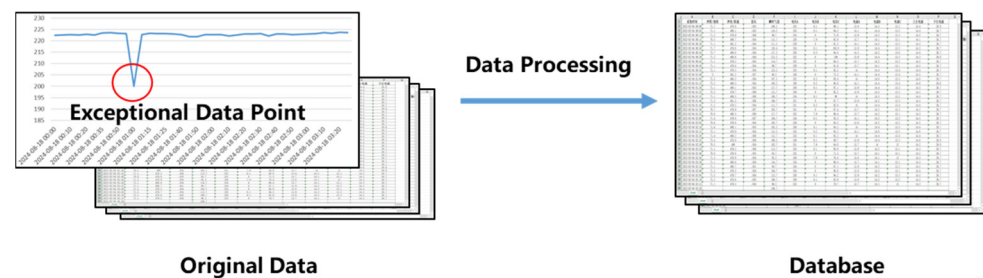


Figure 1. Data processing procedure.

The construction of this database has significantly enriched the data support for the failure prediction of electric submersible progressing cavity pumps, providing a solid foundation for subsequent model training and optimization. By integrating key factors from well intervention operations, we conducted a detailed analysis of failure causes [13]. Furthermore, we employed large-scale text batch processing methods to encode and aggregate the data, ensuring that failure labels accurately correspond to the relevant production data. The fault database establishment scheme is shown in Figure 2.

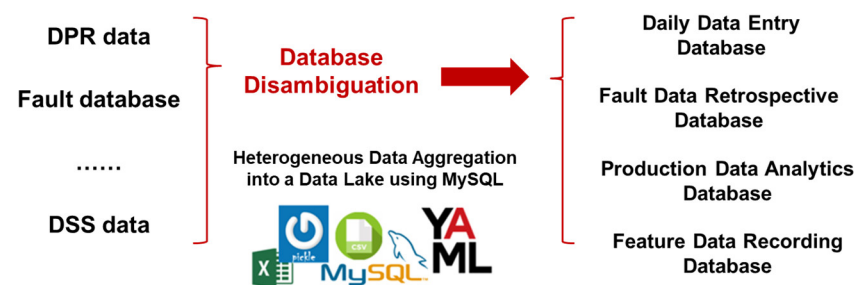


Figure 2. Fault Database Establishment Scheme.

4. Fault Dominant Factor Analysis

4.1. Fault Dominant Factor Analysis Method

To effectively process the high-dimensional production data, we employed Principal Component Analysis (PCA) to perform dimensionality reduction on the operating data of the electric submersible progressing cavity pump. PCA is a widely used dimensionality reduction algorithm that extracts principal components from the original data, reducing its dimensionality. This process significantly lowers the complexity of the data while retaining its key features and removing redundant information [20,21]. By applying PCA to the production data of ESPCPs in coalbed methane wells, we are able to map high-dimensional data into a lower-dimensional space, thereby uncovering potential fault patterns.

The main objective of PCA is to iteratively extract a set of mutually orthogonal principal component axes from the original data. The selection of these new axes depends on the inherent characteristics of the data. Specifically, the first principal component axis is chosen along the direction of the largest variance in the original data. Next, within the plane orthogonal to the first axis, the direction with the largest variance is selected as the second principal component axis. This process continues, with each subsequent principal component being chosen orthogonally to the previous ones, resulting in a new set of orthogonal axes. Studies have shown that most of the data's variance is concentrated in the first k principal components, while the variance in subsequent components is relatively small. Therefore, lower-variance components can be ignored, and by retaining only the first k principal components, dimensionality reduction can be achieved. The main process of PCA is illustrated in Figure 3.

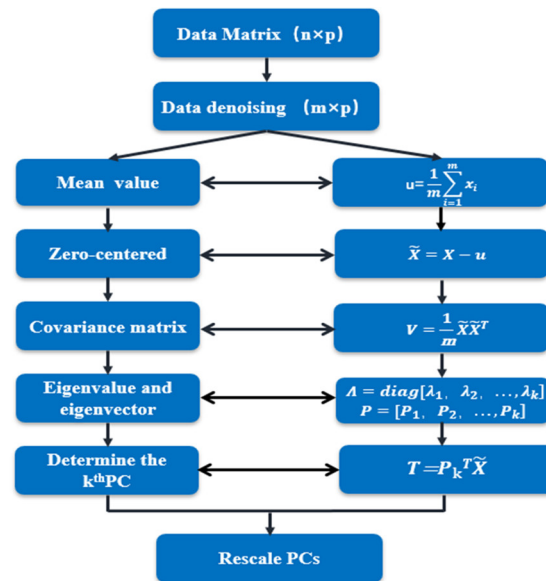


Figure 3. Principal Component Analysis Dimensionality Reduction Algorithm Flowchart.

In our research, by analyzing the data after dimensionality reduction via Principal Component Analysis (PCA), we further integrate the K-Means clustering algorithm to classify different operating conditions. This allows us to trace the dynamic fault process of the Electric Submersible Progressing Cavity Pump (ESPCP) and identify abnormal changes in the equipment's operational state. After dimensionality reduction, we apply the z-score normalization method to eliminate the impact of varying data scales on the analysis, thereby improving the comparability and accuracy of the data.

The main algorithmic processing steps are as follows:

First, we standardize the multi-dimensional input data. The input data can be expressed as:

$$A = \{a_1, a_2, a_3, \dots, a_n\} \quad (1)$$

The data is centralized, with the following calculation formula:

$$a'_i = a_i - \frac{1}{n} \sum_{j=1}^n a_j \quad (2)$$

A covariance matrix is computed from the centralized data:

$$B = AA^T \quad (3)$$

Eigenvalue decomposition is applied to the covariance matrix to obtain the eigenvectors:

$$B = (b_1, b_2, b_3 \dots b_n) \quad (4)$$

All eigenvectors are then normalized and assembled into a feature vector matrix. Each sample in the dataset is transformed into a new sample as follows:

$$c_i = B^T a_i \quad (5)$$

The final output sample set is obtained:

$$C = (c_1, c_2, c_3 \dots, c_n) \quad (6)$$

In this study, we have conducted an extensive data retrieval and analysis of a large-scale database for the ESPCP, uncovering latent patterns within multi-source data. By comprehensively comparing and selecting key influencing factors for different operating conditions, we provide valuable insights into the development of production parameter monitoring schemes.

Initially, our research analyzes the survival probability distribution of the ESPCP over the entire well and establishes a likelihood function model for pump inspection intervals in any given well. In statistics, a likelihood function is used to assess the plausibility of a set of statistical model parameters, representing the probability of observing the given data under the assumption of specific parameters. Specifically, for a given output a , the likelihood function $L(\theta | x)$ for the parameter θ is the probability of the variable α taking the value a , given that θ is known:

$$L(\theta | a) = P(\alpha = a | \theta) \quad (7)$$

Here, θ is a fixed parameter, meaning it is a non-random variable. Thus, this is not a conditional probability but represents the probability that α takes the value a .

Using PCA, we process production data to transform the original high-dimensional data into a lower-dimensional state distribution. During this process, in order to ensure comparability among the control factors, the z-score normalization method is applied to remove the influence of different data scales.

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (a_i - \mu)^2} \quad (8)$$

$$c = \frac{a - \mu}{\sigma} \quad (9)$$

The results indicate that data points representing abnormal states typically appear as outliers, far from other data points, while data points representing normal states are more concentrated and closer to the center of the data distribution. Based on this observation, assuming that the collected data follows a normal distribution, deviations greater than three times the standard deviation from the mean are considered anomalies, which can be used for fault prediction:

$$P(|c| > 3) \leq X \quad (10)$$

where X represents the threshold for anomaly detection.

4.2. Fault Dominant Factor Analysis Results

In the fault database, the PCA method is applied to identify the main control factors for different fault types. The fault survival lifetime curve is illustrated in Figure 4.

The analysis of the primary failure factors reveals that they can be summarized as: Pump failure, Perforation, Breakage, Sand Sticking, and Others. Among these, sand sticking has the highest probability of causing pump failure, while equipment breakage has the most severe impact on the pump's survival lifetime. Additionally, through PCA, it was found that during pump operation, if a low pump rate and low lift condition occur, there is a 56% likelihood of experiencing a tubing perforation failure.

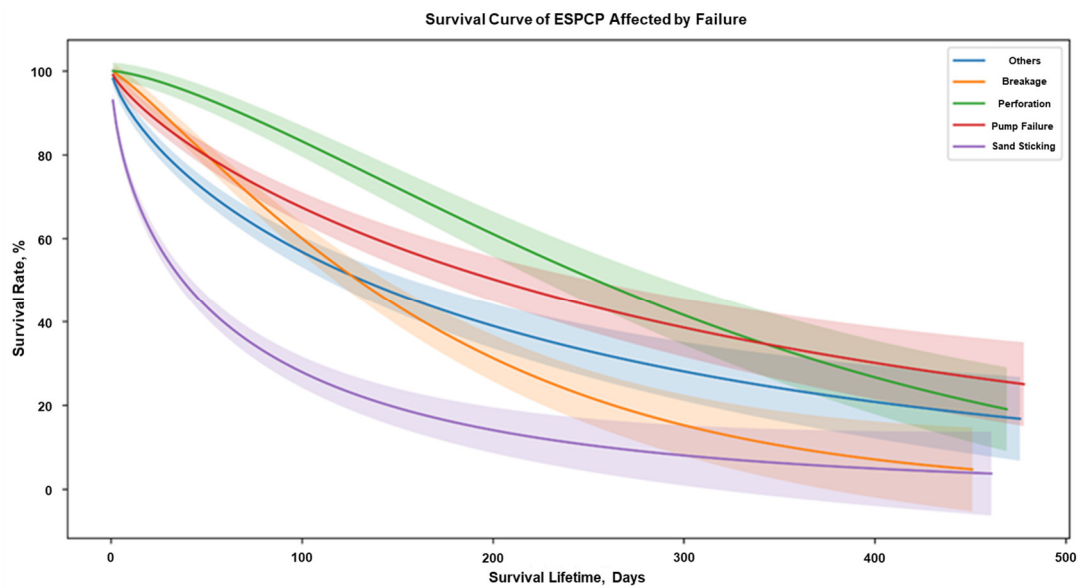


Figure 4. Survival Lifetime Curve of ESPCPS Affected by Failures in the Entire ESPCP Well Group.

Statistical analysis of the survival lifetime for the entire well group of Electric Submersible Progressing Cavity Pumps (ESPCPs) in the block shows that the most probable pump inspection cycle is between 200 and 250 days. This means that the majority of wells are likely to experience a failure around the 200–250-day mark, requiring pump inspection. However, a few wells are able to operate stably for over 500 days without failure. These results are shown in Figure 5.

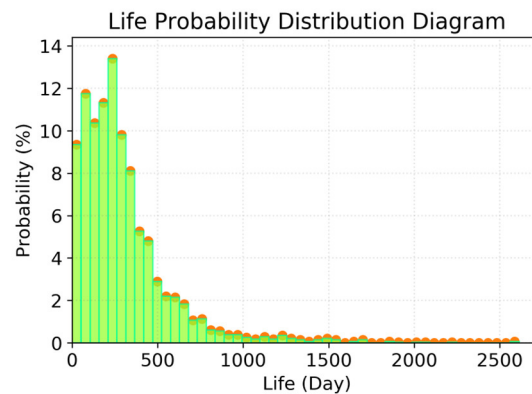


Figure 5. Survival Time Probability Distribution Curve of the Entire ESPCP Well Group.

5. Establishment of Early Warning Model

5.1. Early Warning Threshold Division Method

In our research, a combination of expert knowledge and statistical analysis was used to establish reasonable parameter limits [22,23]. When the values of these parameters exceed or approach these thresholds, the system will promptly issue an alert to notify engineers that the parameter is within a risk range. The alert thresholds are set based on the actual operating conditions of different wells and pumps, including both upper and lower limits [24,25]. These thresholds cover two types of alerts: limit-based alerts and trend-based alerts.

Specifically, the limit-based alert thresholds are determined using statistical methods, where regression analysis is performed on the daily production data from a large number of wells [26]. This is then combined with expert judgment to establish the thresholds.

On the other hand, trend-based alert thresholds are derived through confidence interval calculations within the aforementioned algorithm, reflecting the system's self-learning capability based on the production data of individual wells.

When the actual parameter value exceeds the set alert threshold range, the system will trigger an alert. To calculate the upper and lower limits for trend-based alerts, this study selected the historical daily production data from the past three months for each well and inputted them into the algorithm for analysis. By comparing the historical data with the established upper and lower limits, the corresponding trend alert range can be determined. The main diagram illustrating this is shown in Figure 6.

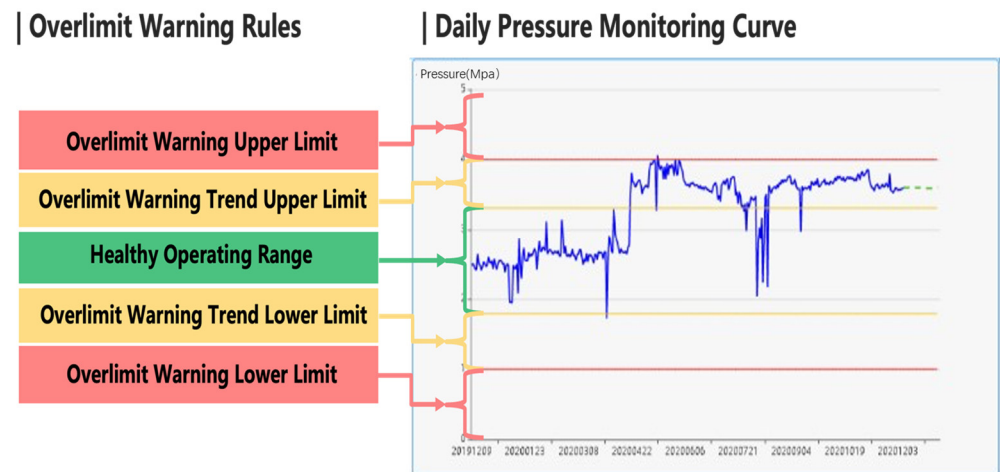


Figure 6. Schematic diagram of overlimit warning partition method.

5.2. Trend Early Warning Algorithm Based on Production Parameter

Based on the operational patterns of production parameters over a given period, we utilize Long Short-Term Memory (LSTM) networks to learn the trends and predict the future trajectory of parameter values [27].

Hyperparameters (number of LSTM units, layers, batch size, epochs) were selected via 5-fold cross-validation on the training set, minimizing validation MAE. Two baseline models were implemented: ARIMA (p, d, q tuned via AIC) and GBDT (learning rate, depth tuned via grid search). These baselines allow assessment of PCA-LSTM's necessity by quantifying its gains in temporal modeling accuracy and early warning precision [28].

Recurrent Neural Networks (RNNs) are a class of neural networks specifically designed to handle sequential data, effectively preserving the temporal relationships between sequential elements. LSTM, a widely used variant of RNN, addresses the limitations of standard RNNs. In an RNN, the previous time-step neuron information is fed back into the network, allowing the hidden layer's output to depend not only on the current input but also on the previous hidden state, which enables the model to capture historical dependencies in sequential data [18,24]. However, during backpropagation, RNNs often suffer from the problems of gradient explosion and vanishing gradients, which can significantly degrade their performance.

LSTM, as a specialized form of RNN, is designed to overcome many of the challenges faced by conventional RNN learning algorithms [29]. The core of the LSTM network lies in its use of gating units, which selectively retain or forget information, thus effectively mitigating the issues of gradient explosion and vanishing gradients. The forget gate determines which information is deemed unnecessary and should be discarded. During the operation of the LSTM, certain pieces of information may be irrelevant, and the forget gate selectively forgets such data, deciding what should be removed from the memory cell.

In contrast, the memory gate governs which new inputs and previously stored information should be retained.

The warning model is structured with a 5-layer network, and the prediction process along with the model architecture is illustrated in Figure 7.

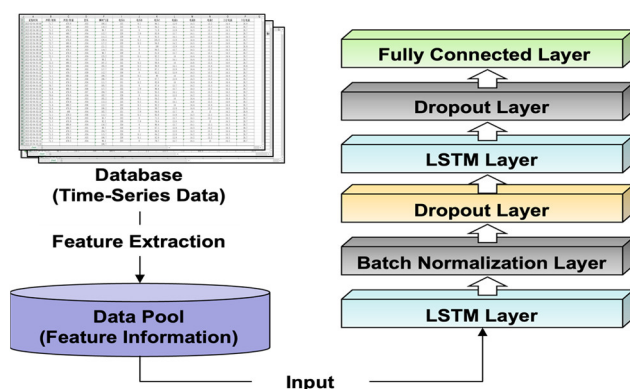


Figure 7. Prediction process and early warning model composition diagram.

The model consists of an input layer, two LSTM layers, a Batch Normalization layer, and a fully connected output layer. The model takes as input a 2D tensor of shape $7 \times n$, where 7 represents the length of the time window, and n is the number of relevant features for the target prediction parameter. The first LSTM layer contains 120 neurons, which are responsible for extracting temporal features, and its output is passed to the Batch Normalization layer. This layer standardizes the data to facilitate more efficient learning of the underlying patterns in the data by the LSTM. The second LSTM layer, with 100 neurons, further processes the extracted temporal features. Finally, the data flows through a fully connected layer that outputs a single neuron value, representing the predicted parameter for the Electric Submersible Progressing Cavity Pump (ESPCP). The specific configuration of each layer is shown in Table 1.

Table 1. LSTM Network Structure Parameters.

Layer Number	Type	Output Shape
1	Input Layer	/
2	LSTM Layer	(7120)
3	Batch Normalization Layer	(7120)
4	LSTM Layer	(100)
5	Fully Connected Layer	(1)

The network is implemented using TensorFlow 2.17.0, and the preprocessing steps include data dimensionality transformation and normalization. Additionally, appropriate iteration numbers and batch sizes were set during training. The number of iterations refers to the number of times all data samples are processed in one cycle, while the batch size refers to the number of samples input to the model at each step. Given the large dataset size, directly inputting all data at once would lead to excessive computational load, so the data is split into smaller batches for training. In this experiment, the number of training epochs was set to 300, with a batch size of 100. The error curves for the training and validation sets are shown in Figure 8.

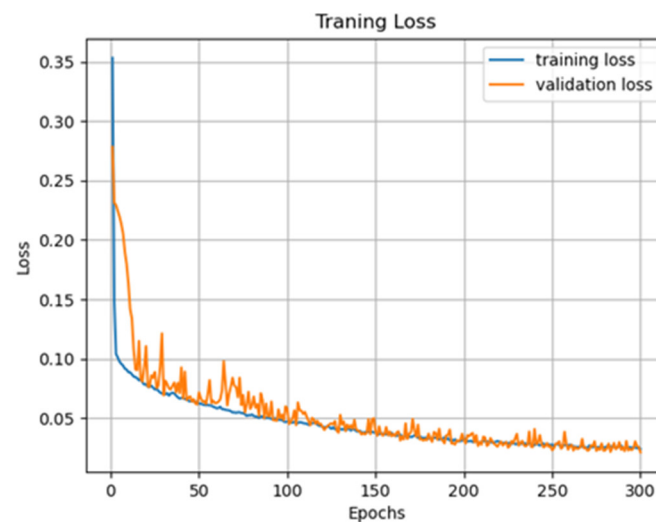


Figure 8. Training Error of Trend Prediction Network.

The dataset is divided into a training set and a test set in a 7:3 ratio, where the training set is used for model training and the test set is reserved for performance evaluation. The evaluation metric selected for the model is the Mean Absolute Error, which is the average of the absolute errors between the predicted values and the actual values for all sample points. The formula is as follows:

$$MAE(X, h) = \frac{1}{n} \sum_{i=1}^n |h(x^i) - y^i| \quad (11)$$

Ultimately, the model achieved errors of 0.0326 and 0.0345 for the training and validation sets, respectively, which meets the accuracy requirements for practical applications. By analyzing the operational data of a parameter from the past month, the model can effectively predict the future trend of that parameter, as shown in Figure 9.

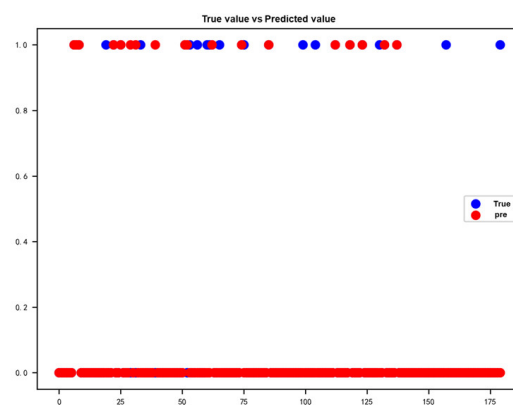


Figure 9. Prediction Instance of the model.

Baseline models including ARIMA and Gradient Boosting Decision Trees (GBDT) were trained for performance comparison, with PCA-LSTM achieving a 15% and 9% improvement in validation accuracy over ARIMA and GBDT, respectively. Model performance is quantitatively reported: acceptance rate 86.79% and validation accuracy 72.22%.

6. Application

In our research, we applied a warning model for fault prediction, which has been in use for two months to date. A total of 53 warning reports and maintenance suggestions were submitted. After review and confirmation by field engineers, 46 of these suggestions were accepted. Among the accepted reports, 39 were later verified as accurate and effective during subsequent operations. The acceptance rate of the predicted results by the field engineers reached 86.79%, and the prediction accuracy was 72.22%. These results are shown in Table 2.

Table 2. Statistics of Fault Prediction Suggestion Acceptance and Accuracy.

Duration	Days	Column Heading	Column Heading
Submitted Suggestions	53	Data row 1	1.0
Accepted	46	Data row 2	2.0
Rejected	7	Data row 3	3.0
Verified as Correct	39		

Fault-type-specific performance showed the highest accuracy for ‘Perforation’ (81%) and the lowest for ‘Sand Sticking’ (65%). False positives were most frequent between ‘Sand Sticking’ and ‘Pump Failure’, while false negatives in ‘Breakage’ cases caused estimated production losses averaging 120 m³ per well (assumed). Operational cost analysis suggests false negatives are ~3× more costly than false positives, highlighting the importance of sensitivity tuning for high-risk faults.

7. Conclusions

This study developed a reproducible and interpretable PCA–LSTM framework for fault prediction in ESPCP wells, addressing the long-standing challenges of black-box predictive models in coalbed methane (CBM) production. Beyond achieving >70% accuracy, the core contribution lies in demonstrating how dimensionality reduction via PCA can explicitly link sensor signals to failure mechanisms, thereby bridging statistical interpretability with deep learning’s temporal modeling strength. This methodological integration establishes a template for transparent and field-deployable predictive maintenance.

From a practical standpoint, the framework enables field engineers not only to forecast failure events but also to understand the reasons why the model issues a warning. Such interpretability reduces the skepticism often associated with AI-driven tools and accelerates operator adoption in production settings. By being reproducible—built on an openly defined dataset spanning 85 wells—and computationally lightweight, the approach offers a realistic path for deployment in real-time supervisory control systems, rather than remaining as a purely academic prototype.

The broader implication of this work is that ESPCP fault prediction can evolve from retrospective diagnosis toward proactive risk management, where model-informed decision-making extends pump run-life, optimizes inspection scheduling, and reduces unplanned downtime. This transition reframes predictive analytics from a supportive tool into a core operational asset in CBM production.

More generally, the study illustrates how *interpretable AI frameworks* can move artificial lift research beyond black-box accuracy benchmarks, opening avenues for cross-pump adaptation (e.g., ESPs, SRPs) and integration with digital twin infrastructures. By embedding transparency and reproducibility as design principles, this work provides not just an algorithm, but a conceptual shift in how predictive models can be engineered to gain trust, drive operational change, and ultimately reshape field management strategies.

Future research should extend the model to transient operating regimes, incorporate physics-guided priors to enhance generalization, and evaluate cost–benefit impacts of predictive deployment at field scale. In doing so, this line of work could become a cornerstone for intelligent, low-carbon CBM development in the coming decades.

Author Contributions: Conceptualization, G.H. and L.Z.; methodology, C.Z.; software, J.D.; validation, C.W. and B.F.; investigation, Z.L.; writing—original draft preparation, J.D.; writing—review and editing, G.H.; visualization, B.L.; supervision, G.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shanxi Coalbed Methane Exploration and Development Branch, PetroChina Huabei Oilfield Company (Grant No. HBYT-SX-2024-JS-258). The APC was funded by China University of Petroleum-Beijing.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Liping Zhao, Chunsheng Wu, Bin Fan, Bin Liu and Zezhou Li were employed by the PetroChina Huabei Oilfield Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The PetroChina Huabei Oilfield Company had no role in the design of the study, the collection, analysis, or interpretation of data, the writing of the manuscript, or the decision to publish the results.

References

- Huang, S.; Hao, Z.; Zhu, S.; Chen, Q. Long Distance Motor Control Technology of Electric Submersible Progressive Cavity Pump in Deep Well Lifting, a Case Study. In Proceedings of the Offshore Technology Conference, OnePetro, Houston, TX, USA, 6–9 May 2024. [\[CrossRef\]](#)
- Saveth, K.J. Field Study of Efficiencies Between Progressing Cavity, Reciprocating, and Electric Submersible Pumps. In Proceedings of the SPE Production Operations Symposium, OnePetro, Oklahoma City, OK, USA, 21–23 March 1993. [\[CrossRef\]](#)
- Wang, Z.; Yang, H.; Chen, S. Study on the operating performance of cross hot-gas bypass defrosting system for air-to-water screw heat pumps. *Appl. Therm. Eng.* **2013**, *59*, 398–404. [\[CrossRef\]](#)
- Wilson, A. Run-Life Improvement by Implementation of Artificial-Lift-Systems Failure Classification. *J. Pet. Technol.* **2016**, *68*, 70–71. [\[CrossRef\]](#)
- Krawiec, M.B.; Finn, C.M.; Cockbill, J.R.; Fortnum, D.E. Dewatering Coalbed Methane Wells Using ESPCPs. In Proceedings of the Canadian International Petroleum Conference, Petroleum Society of Canada, Calgary, AB, Canada, 17–19 June 2008. [\[CrossRef\]](#)
- Dickey, M.W. Economic Pumping Technology for Coalbed Methane (CBM), Stripper Oil, and Shallow Gas Well Deliquification. In Proceedings of the SPE Eastern Regional Meeting, Canton, OH, USA, 11–13 October 2006.
- Hao, Z.; Zhu, S.; Pei, X.; Huang, P.; Tong, Z.; Wang, B.; Li, D. Submersible direct-drive progressing cavity pump rodless lifting technology. *Pet. Explor. Dev.* **2019**, *46*, 621–628. [\[CrossRef\]](#)
- Taufan, M.; Adriansyah, R.; Satriana, D. Electrical Submersible Progressive Cavity Pump (ESPCP) Application in Kulin Horizontal Wells. In Proceedings of the SPE Asia Pacific Oil and Gas Conference and Exhibition, OnePetro, Jakarta, Indonesia, 5–7 April 2005. [\[CrossRef\]](#)
- Jia, J.J.; Liu, C.H.; Guan, T.; Wang, J.X.; Liu, J.D.; Liu, J.L.; Hao, D.Y. A prediction method for liquid production of electric submersible progressing cavity pumps based on a CNN-BiGRU hybrid neural network. *Oil Drill. Prod. Technol.* **2022**, *44*, 784–790. [\[CrossRef\]](#)
- Borisova, K.E.; Ivanova, T.N.; Latypov, R.G. Study of screw pump stator and rotor working capacity to increase the output. *Procedia Eng.* **2017**, *206*, 688–691. [\[CrossRef\]](#)
- Lasrado, V.K. ESP Predictive Analytic Using a Single Application with Machine Learning and Fault Tree Models. In Proceedings of the SPE Middle East Artificial Lift Conference and Exhibition, SPE, Manama, Bahrain, 29–30 October 2024; p. D011S003R005. [\[CrossRef\]](#)
- Abdalla, R.; Nikolaev, D.; Göncü, D.; Manasipov, R.; Schweiger, A.; Stundner, M. Deep Insight into Electrical Submersible Pump Maintenance: A Predictive Approach with Deep Learning. In Proceedings of the SPE Offshore Europe Conference & Exhibition, SPE, Aberdeen, Scotland, UK, 5–8 September 2023; p. D021S006R004. [\[CrossRef\]](#)

13. Panbarasan, M.; Sankar, S.; Venkateshbabu, S.; Balasubramanian, A. Characterization and performance enhancement of electrical submersible pump (ESP) using artificial intelligence (AI). *International Conference on Materials, Mechanics, Mechatronics and Manufacturing. Mater. Today Proc.* **2022**, *62*, 6864–6872. [\[CrossRef\]](#)
14. Abdelaziz, M.; Lastra, R.; Xiao, J.J. ESP Data Analytics: Predicting Failures for Improved Production Performance. In *Proceedings of the SPE International Petroleum Exhibition and Conference, OnePetro, Abu Dhabi, United Arab Emirates, 13–16 November 2017*. [\[CrossRef\]](#)
15. Wang, C.; Ma, H.; Zhang, X.; Xiang, X.; Shi, J.; Liang, X.; Zhao, R.; Han, G. Deciphering Rod Pump Anomalies: A Deep Learning Autoencoder Approach. *Processes* **2024**, *12*, 1845. [\[CrossRef\]](#)
16. Ambade, A.; Karnik, S.; Songchitruksa, P.; Sinha, R.R.; Gupta, S. Electrical Submersible Pump Prognostics and Health Monitoring Using Machine Learning and Natural Language Processing. In *Proceedings of the SPE Symposium: Artificial Intelligence—Towards a Resilient and Efficient Energy Industry, SPE, Virtual, 18–19 October 2021*; p. D011S004R003. [\[CrossRef\]](#)
17. Ma, H.; Han, G.; Zhu, Z.; Wang, B.; Xiang, X.; Liang, X. Hybrid virtual flow metering on arbitrary well patterns for transient multiphase prediction driven by mechanistic and data model. *Geoenergy Sci. Eng.* **2024**, *243*, 213335. [\[CrossRef\]](#)
18. Cheng, X.; Li, R. Parameter equation study for screw centrifugal pump. *International Conference on Advances in Computational Modeling and Simulation. Procedia Eng.* **2012**, *31*, 914–921. [\[CrossRef\]](#)
19. Fu, L.; Ding, G.L.; Zhang, C.L. Dynamic simulation of air-to-water dual-mode heat pump with screw compressor. *Appl. Therm. Eng.* **2003**, *23*, 1629–1645. [\[CrossRef\]](#)
20. Ma, H.; Han, G.; Peng, L.; Zhu, L.; Shu, J. Rock thin sections identification based on improved squeeze-and-Excitation Networks model. *Comput. Geosci.* **2021**, *152*, 104780. [\[CrossRef\]](#)
21. Wright, L.G.; Onodera, T.; Stein, M.M.; Wang, T.; Schachter, D.T.; Hu, Z.; McMahon, P.L. Deep physical neural networks trained with backpropagation. *Nature* **2022**, *601*, 549–555. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Cheng, G.J.; Li, Z.X.; Li, Q.S.; Han, J.; Sun, Y.Z. Lithology identification method based on thin-section images and improved EfficientNet modeling. *J. Xi'an Shiyou Univ. (Nat. Sci. Ed.)* **2025**, *40*, 124–134.
23. Zhou, Q.; Chai, B.; Tang, C.; Guo, Y.; Wang, K.; Wu, W.; Cao, B.; Ye, Y. Enhancing multimodal fault diagnosis in mechanical systems via mixture of experts. *Complex Intell. Syst.* **2025**, *11*, 425. [\[CrossRef\]](#)
24. Chen, S.; Deng, F.; Chen, G.; Zhao, R.; Shi, J.; Jiang, W. Research and Application of Big Data Production Measurement Method for SRP Wells Based on Electrical Parameters. In *Proceedings of the International Petroleum Technology Conference, IPTC, Bangkok, Thailand, 1–3 March 2023*; p. D012S001R003. [\[CrossRef\]](#)
25. Zhu, S.; Hao, Z.; Zhang, L.; Ming, E.; Wang, Q. New Stage of Rodless Artificial Lift Operation: The First Field Application of Submersible Motor Cable Plug with Electric Submersible Progressing Cavity Pump in CNPC. In *Proceedings of the SPE Artificial Lift Conference and Exhibition—Americas, OnePetro, Virtual, 10–12 November 2020*. [\[CrossRef\]](#)
26. Lastra, R.A.; Xiao, J. Machine Learning Engine for Real-Time ESP Failure Detection and Diagnostic. In *Proceedings of the SPE Middle East Artificial Lift Conference and Exhibition, SPE, Manama, Bahrain, 25–26 October 2022*; p. D021S007R001. [\[CrossRef\]](#)
27. Liu, D.; Feng, G.Q.; Feng, G.Y.; Xie, L. Hybrid Long Short-Term Memory and Convolutional Neural Network Architecture for Electric Submersible Pump Condition Prediction and Diagnosis. *SPE J.* **2024**, *29*, 2130–2147. [\[CrossRef\]](#)
28. Roostaei, M.; Nouri, A.; Fattahpour, V.; Chan, D. Numerical simulation of proppant transport in hydraulic fractures. *J. Pet. Sci. Eng.* **2018**, *163*, 119–138. [\[CrossRef\]](#)
29. Bohorquez, M.; Rubiano, E.; Labrador, L.; Suarez, M.C. Implementation of Bottom-Drive Progressive-Cavity Pumps Technology in La Cira-Infantas Oil Field as a Reliable Artificial Lift Method. In *Proceedings of the SPE Artificial Lift Conference-Americas, OnePetro, Cartagena, Colombia, 21–22 May 2013*. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.