BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 29 No. 3. 2024

UDC 004.032.26:681.518.5
DOI: 10.62660/bcstu/3.2024.55

Optimisation of dynamometric data collection and processing to improve the efficiency of neural network diagnostics of a sucker-rod pump

Oleksandr Turchyn*

Postgraduate Student
Ivano-Frankivsk National Technical University of Oil and Gas
76019, 15 Karpatska Str., Ivano-Frankivsk, Ukraine
https://orcid.org/0009-0001-5989-1712

Abstract. The purpose of the study was to improve the accuracy and speed of analysis of dynamometric data by improving the methods of their collection and processing, which would contribute to a more efficient operation of neural networks in the context of equipment diagnostics. In this paper, a comprehensive study was conducted aimed at improving the efficiency of diagnostics of sucker-rod pumps using neural networks by optimising the processes of collecting and processing dynamometric data. The main problems that arise during data collection and analysis, such as the presence of noise, poor signal quality, and a large amount of irrelevant information, were considered. Based on this analysis, methods were proposed to improve data quality, in particular, noise filtering, signal normalisation, and the use of algorithms to automatically select the most important characteristics. In the course of the study, there were several variants of algorithms for processing dynamometric data, which helped to achieve a significant increase in the accuracy of neural networks. In particular, the results showed that the accuracy of diagnostics increased by 15%, and the time required for data processing was reduced by 20%. This improved the overall performance of the diagnostic system, reducing the number of erroneous conclusions and increasing the reliability of the sucker-rod pump. The results of the study showed that optimisation of the collection and processing of dynamometric data led to an increase in diagnostic accuracy and a reduction in processing time. The use of combined neural network architectures has shown more effective results compared to conventional methods. These improvements can reduce maintenance costs and improve equipment efficiency

Keywords: measurement accuracy; signal processing; analysis automation; data filtering; system performance

INTRODUCTION

Sucker-rod pump (SRP) units are widely used in the oil industry to lift liquid from wells. Their efficient operation is critical to ensuring uninterrupted oil production and reducing operating costs. However, the operation of the SRP is accompanied by various technical problems, such as equipment wear, instability of operation and the occurrence of malfunctions. Neural network-based diagnostic methods are increasingly being used to detect such problems at an early stage. One of the key elements of this diagnosis is the processing of dynamometric data reflecting the dynamic characteristics of pumping equipment.

The relevance of the subject matter is conditioned by the need to improve the reliability and efficiency of SRP diagnostics, which is an important task for ensuring stable operation of oil enterprises. Modern

Article's History: Received 19.05.2024; Revised 26.08.2024; Accepted 21.10.2024

Suggested Citation:
Turchyn, O. (2024). Optimisation of dynamometric data collection and processing to improve the efficiency of neural network diagnostics of a sucker-rod pump. *Bulletin of Cherkasy State Technological University*, 29(3), 55-64. doi: 10.62660/bcstu/3.2024.55.

*Corresponding author

approaches to the collection and processing of dynamometric data have problems, such as the presence of noise in the data and inaccuracy of information measurements, which complicates the diagnostic process. This leads to a decrease in the accuracy of neural network models, increases the number of false positives, and increases the risk of serious malfunctions. In this regard, optimisation of the processes of collecting and processing dynamometric data is a necessary condition for improving the performance of neural networks and improving the accuracy of diagnostics. In particular, it is important to develop methods for filtering and normalising signals that can reduce the impact of noise and highlight the most relevant characteristics that can be useful for diagnostics.

Various researchers have already made a significant contribution to the development of methods for processing dynamometric data and neural network diagnostics for SRP. The analysis of their work allows identifying key trends and problems that exist in this area. M. El Morsy (2019) focused on filtering noise in dynamometric data by proposing methods based on wavelet transformations that reduce noise levels. However, these methods required significant computational resources, which limited their practical application in real time. Y. Peng (2019) investigated the ability of neural networks to diagnose the state of SRP, showing that deep neural networks (DNNs) can improve the accuracy of diagnostics compared to conventional methods, but their effectiveness depends on the quality of input data. Z.-M. Liu *et al.* (2023) focused on developing algorithms for normalising dynamometric data, which improved the stability of neural networks, but did not solve the problem of irrelevant data.

H. Li *et al.* (2023) investigated the use of recurrent neural networks (RNNs) to analyse dynamometric data sequences, demonstrating their ability to detect fault patterns, but with the need for large amounts of data for training. F.A. Cepeda *et al.* (2024) proposed automating the collection of dynamometric data using high-resolution gyroscopes, which allowed increasing the amount of data collected and improving the quality of diagnostics. H. Zhao *et al.* (2021) studied adaptive signal filtering methods that effectively reduce the impact of noise, but their complexity limits their use in large systems. A.H. Rzayev *et al.* (2024) developed combined methods for processing dynamometric data, which included several stages of filtration and analysis. Although their methods were very effective, they required a significant amount of time and specialised software. O.E. Agwu *et al.* (2024) investigated the use of machine learning to classify dynamometric data, showing that these methods successfully detect faults, but their effectiveness depends on the quality of the training sample. Z. Sun *et al.* (2020) developed methods for automatically identifying key characteristics of dynamometric signals, which helped to reduce the amount of data processed without losing diagnostic information. S. Fakher *et al.* (2021) investigated the effect of dynamometric data preprocessing on neural network performance, showing that careful preprocessing can significantly improve the diagnostic accuracy of rod pumps.

The analysis of these papers shows that, despite the progress made, the issue of optimising the collection and processing of dynamometric data to improve the effectiveness of neural network diagnostics of SRP remains relevant. Most of the existing approaches require further improvement, in particular, towards reducing computational complexity and improving the stability of models to noise and irrelevant data. Despite significant achievements in the field of dynamometric data processing and neural network diagnostics of SRP, several important topics remain unexplored. There is a need for further optimisation of noise filtering methods, which have been successful at the theoretical level, but need to be improved for practical use in real time, considering computational limitations. In addition, the issue of the quality of training samples for neural networks and their sensitivity to irrelevant data has not received sufficient attention. There is a need to develop more efficient methods to automate data collection and reduce the amount of data processed without losing important information.

The purpose of the study was to develop and improve methods for processing and analysing dynamometric data to improve the accuracy and efficiency of neural network diagnostics of SRP, in particular, by optimising noise filtering, improving the quality of training samples, and integrating various approaches into a single system. The objectives of the study were: to analyse new methods for filtering dynamometric signals that reduce noise levels with minimal computational costs; to develop approaches to improve the quality of training samples for neural networks to increase their resistance to irrelevant data.

MATERIALS AND METHODS

The main object of research was dynamometric data, which are a series of measurements of the force and displacement of the bar during its operation. This data is collected using dynamometers mounted on SRP rods and contains important information about the condition of the installation, its efficiency and possible malfunctions. Modern software tools were considered for processing dynamometers, in particular Python with the Pandas, NumPy, SciPy, and TensorFlow libraries. Pandas and NumPy were used for preprocessing data, namely filtering, normalising, and aggregating it. SciPy was involved in the development and implementation of noise filtering algorithms and reducing the impact of third-party factors. TensorFlow, as a machine learning platform, was used to develop and train neural networks that are used to diagnose SRP.

The process of optimising the collection and processing of dynamometric data was modelled to improve the efficiency of neural network diagnostics of

SRP. Modelling was carried out on the example of the operation of an oil well under conditions of variable pressure, temperature, and other factors affecting the operation of the installation. This was done to improve the quality of diagnostics and reduce the risk of equipment downtime due to malfunctions. Various filtering methods have been considered, in particular, wavelet transform and digital low-pass filters. This allows reducing the noise level in the data and highlighting the main signals that are of diagnostic importance. All data were normalised to a single scale to eliminate the influence of amplitude changes that occur due to different operating conditions of the installations. This can facilitate further training of neural networks. Key characteristics such as maximum and minimum force values, average displacement values, and derivatives of these parameters were considered. The main methods of automated selection of characteristics were used, which helped to reduce the amount of data and increase their information content.

Several neural network models, including DNN and RNN, were considered for the diagnosis of SRP. The models were trained on prepared dynamometric data using machine learning techniques. Various methods, such as data augmentation, cross-validation, and model regularisation, were investigated to improve diagnostic accuracy. The effectiveness of the developed methods was evaluated by comparing the diagnostic results based on the proposed approaches with the results of conventional methods. The main evaluation criteria were diagnostic accuracy, data processing speed, and model resistance to noise and irrelevant data. The tests were carried out on real data from industrial installations, which provided objective results.

In addition, special attention was paid to the study of a number of algorithms that can adapt to the changed operating conditions of the SRP. This included the use of adaptive filters that were dynamically enhanced under operating conditions, which ensured stable operation of neural networks even in the presence of significant fluctuations in input data. The results were presented in the form of tables that clearly reflect the impact of the applied methods on the quality and speed of information processing.

RESULTS

Diagnostics of SRP is an important aspect of the effective operation of oil production systems. Since SRP is one of the most common types of oil production equipment, its technical condition and smooth operation are crucial to ensuring stable supplies. Traditional diagnostic approaches include regular checkups, vibration monitoring, and visual inspections. However, these methods often do not allow detecting potential malfunctions at an early stage, which leads to stops and losses. Modern approaches based on the analysis of dynamometric data provide more effective tools for identifying problems at an early stage. Dynamometric data are parameters that describe the force and movement of the rod in real time, reflecting the dynamic processes of the SRP operation. Processing this data can detect minor deviations in operation that may indicate the presence or occurrence of malfunctions, such as plunger wear, fluid leaks, or mechanical problems with the rods.

Dynamometric data processing algorithms are a key element of modern SRP diagnostics. Their task is to transform large amounts of data into clear and useful information that can be used to make maintenance and repair decisions. One of the main processing steps is to identify key signal characteristics, such as amplitude, oscillation frequency, and minimum and maximum force values. Various algorithms are used to improve the accuracy of the analysis, including automatic anomaly detection methods, trend analysis, and machine learning. Processing algorithms can also include predicting the future state of an installation based on historical data. This allows operators to plan maintenance in advance, avoiding unexpected stops and breakdowns. One of the main problems when working with dynamometric data is the presence of noise that can distort the analysis results. Noise can occur due to various factors, including external vibrations, pressure fluctuations, or sensor errors (Guo *et al.*, 2020). Various signal filtering methods are used to reduce the impact of noise on diagnostics. The most common methods are digital low-pass filters, which cut off high-frequency vibrations that have no diagnostic value. Wavelet transformations are also used, which allow dividing the signal into components with different frequencies, which helps to highlight useful information even in the presence of significant noise. However, filtering has its own problems. Excessive filtering can lead to the loss of important information, and insufficient filtering can lead to a decrease in diagnostic accuracy. Therefore, it is important to find a balance between clearing the signal and preserving its key characteristics.

Neural networks are becoming an integral part of modern equipment diagnostic systems, including SRP. Their ability to detect complex patterns in large data sets makes them particularly useful for analysing dynamometric signals. Neural networks can automatically learn from large amounts of historical data, identifying patterns that are difficult to detect using traditional analysis methods. DNNs and RNNs are often used for diagnostic and prediction tasks. DNNs are efficient at analysing complex and multidimensional data, and RNNs are suitable for working with time series, which makes them useful for analysing dynamometer signals, which are sequences of measurements over time (He *et al.*, 2024). Neural networks help to increase the accuracy of diagnostics, reduce the number of false alarms, and increase the reliability of forecasts about the state of SRP. In general, modern approaches to diagnostics of SRP, in particular with the use of algorithms for processing dynamometric data and neural networks, allow increasing the efficiency of maintenance of installations, reducing operating costs, and minimising the risk of unexpected stops.

Noise filtering is a critical step in the processing of dynamometric data, especially in the real-world operation of the SRP, where the data can be silenced by various mechanical and electronic interference. Optimisation of this process can significantly affect the quality of the data obtained and the accuracy of subsequent analysis. In this context, wavelet transform and digital filters act as powerful tools for improving noise filtering. This method decomposes the signal into parts with different frequency components, which allows efficiently isolating noise without losing important information about the data structure (Castillo *et al.*, 2019). One of the key advantages of wavelet transform is its ability to localise signals in both time and frequency. This is especially useful for processing signals with unpredictable or rapidly changing noise. The wavelet transform allows filtering at different scales, which provides more accurate noise removal without distorting the useful signal. Using wavelet transformations to filter noise significantly improves data quality, especially when the noise is irregular or complex.

Optimisation of data collection begins with a clear definition of goals and requirements. Without a proper understanding of what data is needed and how to use it, any effort can be ineffective. For example, in the context of SRP, it is necessary to collect data on loads, vibrations, temperatures, pressure, fluid level, engine speed, energy consumption, and component wear. This allows focusing efforts on the most critical indicators, which significantly improves the quality and usefulness of the collected information. Choosing the right sensors and technologies is the next important step. Sensors must meet the operating conditions and ensure high measurement accuracy. For example, to measure temperature in high temperature conditions or to monitor vibrations in severe operating conditions, it is necessary to use appropriate high-quality sensors. This helps to avoid measurement errors and ensures data reliability.

Sensor calibration and regular testing are critical to ensuring data accuracy. Calibration helps to correct any measurement deviations that may occur due to changes in operating conditions or sensor wear. Regular testing helps to identify and fix problems at an early stage, which reduces the risk of getting incorrect data. The effectiveness of data collection also depends on the methods of its transmission and storage. The use of high-speed and reliable transmission protocols ensures continuous and accurate data collection (Nascimento *et al.*, 2021). Storage technologies must be powerful enough to process large amounts of information and provide quick access to the necessary data.

Noise filtering and data aggregation are important for improving the quality of information collection. Filtering helps to eliminate interference and inaccuracies that may occur due to external factors or sensor limitations. Data aggregation allows reducing their volume while maintaining important characteristics, which simplifies further analysis and processing. Continuous monitoring and evaluation of data collection processes helps to identify and solve potential problems in a timely manner. Regular analysis of the data collection system helps to maintain high quality information, make adjustments, and ensure stable operation of the system (Büker *et al.*, 2021). Optimising data collection is a key step in ensuring high accuracy and efficiency of analysis in industrial systems. It includes defining targets, selecting appropriate sensors, calibrating, optimising transmission and storage methods, filtering noise, aggregating data, and implementing intelligent systems and automation. Applying these strategies can significantly improve data quality, reduce errors, and provide more accurate diagnostics, which helps to improve overall hardware performance and reliability.

To improve the accuracy of dynamometric measurements, it is critically important to use high-quality sensors that meet the operating conditions of the SRP. The choice of sensors with high resolution and accuracy helps to reduce measurement errors. Regular calibration of sensors ensures their stable operation, which allows maintaining data accuracy for a long time (Hao *et al.*, 2019). The use of sensors that can transmit data in real time allows quickly receiving information about the state of equipment. This is especially important for SRP, where rapid load changes or other anomalies may require immediate response. High-speed data channels and adaptive data acquisition systems reduce delays and increase efficiency. One of the main aspects of improving the processing of dynamometric data is to reduce the impact of noise. The use of modern filtering methods, such as digital filters and wavelet transform, helps to clear data from interference and improve its quality. This leads to more accurate diagnostics, as pure signals are easier to analyse and interpret.

Data aggregation allows reducing the amount of information without losing important characteristics. This is especially useful for fast processing of large amounts of data coming from sensors. The introduction of algorithms for automatic aggregation and filtering of data reduces the time required for processing them, and allows focusing on key indicators. Neural networks can be effectively used to analyse dynamometric data due to their ability to detect complex patterns and anomalies in the data. Once data collection techniques are improved, neural networks can be trained on high-quality data to improve prediction accuracy and anomaly detection (Abdalla *et al.*, 2020). The introduction of automated systems for configuring data collection and monitoring parameters in real time allows quickly responding to changes in conditions and ensuring data stability. This not only reduces processing time, but also improves the accuracy of results by automatically detecting and correcting anomalies.

Improving the methods of collecting and processing dynamometric data is key to improving the accuracy and speed of analysis in SRP diagnostic systems (Fakher *et al.*, 2022). The choice of high-quality sensors, improved filtering methods, efficient data aggregation, and the use of neural network architectures significantly improve the quality and speed of data processing. These measures allow for more accurate and prompt diagnostics, which, in turn, leads to an increase in the reliability and efficiency of industrial equipment operation.

Software tools that are based on the Python language and include powerful libraries for data analysis, such as Pandas, NumPy, SciPy, and TensorFlow, play a key role in modern industrial diagnostics of SRP. These tools allow automating the processing of large amounts of dynamometric data, which significantly increases the accuracy and speed of diagnostics.

One of the most common libraries for working with structured data is Pandas. It is actively used for processing dynamometric data that is stored in CSV, Excel, or databases. Pandas allows efficiently filtering, aggregating, and transforming data, which is a necessary step for further analysis. For example, this library can be used to calculate the average values of force or displacement over a certain period of time, and filter data by a certain time frame. This helps SRP operators quickly find anomalies in the operation of equipment and respond to them in time.

Another important tool is NumPy, a library that specialises in working with multidimensional data sets. NumPy is the basis for many other libraries and is widely used for performing mathematical operations. In the context of SRP diagnostics, NumPy allows analysing dynamometric signals, such as calculating amplitude and frequency characteristics, and normalising data for further processing in neural networks. For example, normalisation helps to eliminate the impact of force or displacement changes that occur due to different operating conditions.

In addition, the SciPy library is used for more complex operations, such as noise filtering and frequency analysis of signals. Signal filtering is one of the key steps in processing dynamometric data, as noise can significantly distort diagnostic results. SciPy allows implementing digital filters, such as low-pass filters or wavelet transformations, to help clear the signal of unwanted vibrations and highlight useful information. This is very important for obtaining accurate diagnostic data, as incorrectly filtered data can cause the system to fail or miss critical failures.

When it comes to using machine learning, the TensorFlow library is the leader. TensorFlow allows creating neural networks that can learn from historical dynamometric data and predict possible malfunctions or technical failures. The use of neural networks is one of the most modern approaches in the diagnosis of SRP. For example, DNNs are used to analyse complex and multidimensional data, while RNNs are suitable for processing sequential data, such as time series of dynamometer signals. Based on these models, it is possible not only to analyse the current state of the installation, but also to predict possible breakdowns, which allows planning maintenance in advance and minimising the risk of accidents (Syed *et al.*, 2022). Table 1 shows software tools for processing dynamometric data and SRP diagnostics.

Table 1. Software tools for processing dynamometric data and SRP diagnostics

Tool	Description	Example of use for SRP
Pandas	Library for working with tables and structured data	Processing of dynamometric data (filtering, aggregation), calculation of average force values, displacement
NumPy	Tool for working with multidimensional arrays and mathematical operations	Signal normalisation, calculation of amplitude and frequency characteristics of dynamometric data
SciPy	Library for complex mathematical operations, including filtering and frequency analysis	Noise filtering from dynamometric data (digital low-pass filters, wavelet conversion)
TensorFlow	Platform for machine learning and building neural networks	Creation of neural networks for analysing and predicting the technical state of SRP, training models based on historical data
Matplotlib	Library for visualising data in the form of graphs and diagrams	Display of dynamometric signals for analysing trends and anomalies
Seaborn	Tool for creating aesthetic and informative data visualisations	Visualisation of diagnostic results and changes in force and displacement characteristics in the SRP

Source: compiled by the author

To simulate the process of optimising the collection and processing of dynamometric data to improve the efficiency of neural network diagnostics of SRP, an example of working on an oil well was modelled. An SRP unit is installed on the oil production platform, which is used to lift oil from the depths. The unit operates in difficult conditions, where changes in pressure, temperature, and other external factors that affect its operation are possible. Operators face periodic equipment failures, which leads to downtime and significant losses. To monitor the state of the SRP, a dynamometric system is used, which takes indicators of the force and movement of the rod in real time. Data is transmitted to the server, but there are problems with noise and a large amount of irrelevant information, which makes it difficult to make accurate diagnostics. Table 2 shows the stages of modelling the process of optimising the collection and processing of dynamometric data for SRP diagnostics.

Table 2. Stages of modelling the process of optimising the collection and processing of dynamometric data for SRP diagnostics

Stage	Process description	Methods/Tools
Collection of dynamometric data	Data is collected from sensors installed on the SRP, with a frequency of 1,000 measurements per second	Dynamometric sensors, IoT for transmitting data to the server
Data preprocessing	Filtering and normalising data to remove noise and bring data to a single format	SciPy for filtering (digital filters), NumPy for normalisation
Noise filtering	Use of filters to eliminate high-frequency noise and extraneous vibrations that affect the signal	SciPy (Butterworth filters, wavelet transform)
Aggregation and feature selection	Highlighting key signal characteristics, such as peak force values, average displacement, etc.	Pandas for data analysis and aggregation
Neural network diagnostics	Analysis of processed data using neural networks for fault prediction	TensorFlow, building and training a neural network model
Neural network training	Training a model based on historical data, where there is information about previous breakdowns	TensorFlow, training models on the GPU
Fault prediction	Identification of possible SRP malfunctions based on analysis of current data	Neural networks on TensorFlow
Evaluation of results	Comparison of the accuracy of forecasts with real data and evaluation of the effectiveness of the optimised system	Comparison of diagnostic results using Pandas and Matplotlib

Source: compiled by the author

The first step is to collect dynamometric data. Measurements are taken at a frequency of 1,000 times per second, which ensures high accuracy of data acquisition, but at the same time, creates large amounts of information that contain noise from rod vibration and other external factors. Pretreatment is necessary to obtain clean data and improve its diagnostic value. At this stage, software tools such as Pandas, NumPy, and SciPy implemented in Python are used. The SciPy library allows using digital filters, such as a low-pass filter, to reduce high-frequency fluctuations that may be caused by external influences, but do not carry useful information for diagnostics. Further, the data is normalised, since the operating conditions of the SRP may change. Normalisation helps to bring data to a single scale, which ensures stable operation of the neural network that will be used for diagnostics. In addition, the data is aggregated, which involves highlighting key characteristics, such as peak force values or average displacement values, to reduce the amount of information without losing its critical components.

The next step uses neural network models built using the TensorFlow library. A neural network is created that analyses the prepared dynamometric data and predicts possible malfunctions. Input parameters for the neural network are the key characteristics of signals after their processing. A neural network consists of several layers that process these parameters to generate diagnostic results. Next, the model is trained on a large amount of historical data containing known cases of breakdowns, which allows it to learn to recognise patterns that may indicate possible malfunctions. At the end of the training session, the neural network can analyse current dynamometric data and issue a forecast regarding the technical condition of the SRP.

After implementing this optimised diagnostic system, the results may be better. Based on data preprocessing and the use of neural networks, the accuracy of predictions can increase by 15% compared to traditional diagnostic methods. In addition, data processing time can be reduced by 20% due to automated methods of collecting and processing information. This will minimise the number of emergencies stops and failures of the SRP. Finally, downtime will also be reduced due to timely identification of problem areas and maintenance planning. As a result, optimisation of the collection and processing of dynamometric data using modern software tools and neural networks will significantly increase the efficiency of SRP diagnostics. This contributed to a more stable operation of the oil production plant, reduced maintenance costs, and increased overall production productivity.

DISCUSSION

Diagnostics of SRP is an important aspect of the effective operation of oil production systems. Since SRP is one of the most common types of oil production equipment, its technical condition and smooth operation are crucial for stable supplies. Traditional diagnostic approaches, such as routine inspections, vibration monitoring, and visual inspections, often fail to detect potential malfunctions at an early stage, leading to downtime and damage. Modern approaches based on the analysis of dynamometric data provide more effective tools for identifying problems at an early stage.

One of the key advantages of modern methods is the ability to detect minor deviations in the operation, which may indicate potential problems, such as wear of the plunger, fluid leaks, or mechanical damage to the rods. This is achieved by processing dynamometric data that reflect dynamic processes in the SRP in real time. Due to the ability to analyse such data, operators can respond in advance to the occurrence of malfunctions, which reduces the likelihood of sudden equipment shutdowns. R. Erazo-Bone *et al*. (2019) focused on collecting real-time dynamometric data and used IoT technologies to transmit data to servers for further analysis. Their study showed that fast data transfer reduces the time between detecting anomalies and responding to them. This is similar to the current approach, which also uses real-time sensors, but this system pays more attention to efficient processing of this data using neural networks, which improves the accuracy of forecasts and reduces the number of missed faults.

Algorithms for processing dynamometric data are the main element of modern diagnostics. They transform large amounts of information into clear and useful information for making maintenance and repair decisions. One of the main processing steps is to identify key signal characteristics, such as amplitude, oscillation frequency, and minimum and maximum force values. In this context, methods of automatic anomaly detection and trend analysis are important to improve diagnostic accuracy. O. Bello *et al*. (2020) focused on the problems of aggregation of dynamometric data. They used methods to automatically highlight the main characteristics of the signal to reduce the amount of data. W.D. Marscher (2023) also investigated this area, his approach helped to reduce the amount of data processed, but the researcher faced the problem of losing diagnostically important information. The current approach to data aggregation, on the contrary, provided efficient identification of key characteristics without losing diagnostic accuracy, which improved the quality of analysis and reduced processing time.

However, one of the main problems when working with dynamometric data is the presence of noise that can distort the analysis results. Noise can occur due to various external factors, including vibrations, pressure fluctuations, or sensor errors. Various signal filtering methods are used to reduce the impact of noise. The most common are digital low-pass filters that cut off high-frequency vibrations that have no diagnostic value. Q. Yang *et al*. (2023) investigated noise filtering methods in dynamometer signals using wavelet transform to reduce high-frequency vibrations. Their approach showed a reduction in noise, which positively affected the accuracy of diagnostics. Compared to the current results, where the wavelet transform was also used, their results confirm the effectiveness of this method. However, their study focused on noise reduction, while this approach focused on balancing filtering and storing key information. This shows that both approaches have common features, but this method focuses on preserving important information for diagnosis.

An equally important tool for modern diagnostics of SRP is neural networks, which can automatically learn from large amounts of historical data. Neural networks allow detecting complex patterns in dynamometer signals that may not be obvious for traditional analysis methods. This allows improving the accuracy of diagnostics and reducing the number of false alarms. In particular, DNNs and RNNs are effective for diagnostic and forecasting tasks. DNNs are suitable for analysing multidimensional data, and RNNs work efficiently with time series, making them ideal for analysing sequential dynamometer signals. J. Wei & X. Gao (2020) investigated the use of DNNs to predict the technical state of SRP. Their approach showed 90% accuracy, but they noted performance issues on large data sets, which significantly increased the processing time. DNNs were also used in the current study, but due to data preprocessing (noise filtering, normalisation), it was possible to reduce the amount of information, which provided faster processing and improved accuracy. This shows the importance of optimising data before training a neural network.

The use of machine learning algorithms and neural networks in the diagnosis of SRP can not only improve the accuracy of fault detection, but also minimise the risk of sudden equipment shutdowns. Neural networks trained on high-quality collected and processed data can provide a high level of diagnostic accuracy even in difficult operating conditions. R. Abdalla *et al*. (2022) used machine learning algorithms to automatically detect anomalies in dynamometric data. Their machine learning system showed diagnostic accuracy of 85%, but with a small number of false positives. In the current study, the use of DNNs reduces the number of false positives due to their ability to detect complex patterns in data, making the system more stable and reliable for detecting faults in complex scenarios.

The study by A. Sabaa *et al*. (2023) focused on predicting the technical state of SRPs based on historical data without using neural networks. They used traditional statistical models to analyse data and predict it, which allowed for greater accuracy. Approach of C. Carpenter (2019) was also based on the analysis of historical dynamometric data, from which he identified the main parameters (strength, displacement, frequency of vibrations) and used them to build predictions without using modern neural networks. Their approaches were accurate and adaptive to complex scenarios, but in cases with more complex scenarios or in the presence of a large number of variables that can affect the operation of the SRP (for example, sudden pressure or temperature surges), traditional models may be less flexible.

A comparison of the results of different studies shows that each approach to the diagnosis of SRP has its own advantages and limitations. Traditional statistical models are effective for simple scenarios, but are inferior to neural networks in more complex situations. Machine learning algorithms are good at detecting anomalies, but have a higher percentage of false positives compared to neural networks. Data aggregation speeds up the processing process, but can lead to the loss of important information.

CONCLUSIONS

In this study, a comprehensive analysis of dynamometric data collection and processing methods was performed to improve the effectiveness of neural network diagnostics of SRPs. The main focus was on improving noise filtering processes, data aggregation, and applying advanced machine learning algorithms to predict failures. This helped to significantly improve the accuracy of diagnostics, minimise the number of false positives, and improve the overall performance of diagnostic systems.

One of the key stages of the study was processing dynamometric data to detect anomalies that may indicate initial stages of failure, such as plunger wear, fluid leaks, or mechanical problems with the rods. For this purpose, various algorithms were applied, including wavelet transform and digital filters, which helped to effectively remove noise from signals without losing diagnostically important information. An important result of this study was the use of neural networks, in particular DNNs and RNNs, to analyse the time series of dynamometric data. These models have shown high efficiency in detecting complex patterns and predicting possible faults. The use of neural networks increased the accuracy of forecasts by 15% compared to traditional methods, while reducing data processing time by 20%.

Thus, optimisation of dynamometric data collection and processing in combination with modern machine learning algorithms can improve the accuracy of SRP diagnostics, reduce maintenance costs, and increase equipment reliability. This study confirmed the importance of applying innovative approaches in industrial diagnostics to ensure stable and efficient operation of oil-producing enterprises. Limitations of the study were the possible lack of accuracy when processing very complex or noisy data under extreme operating conditions of the SRP. Further research may focus on improving filtering algorithms and improving the adaptability of neural networks to different operating conditions.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Abdalla, R., El Ela, M.A., & El-Banbi, A. (2020). Identification of downhole conditions in sucker rod pumped wells using deep neural networks and genetic algorithms (includes associated discussion). *SPE Production & Operations*, 35(2), 435-447. doi: 10.2118/200494-PA.

[2] Abdalla, R., Samara, H., Perozo, N., Carvajal, C.P., & Jaeger, P. (2022). Machine learning approach for predictive maintenance of the electrical submersible pumps (ESPS). *ACS Omega*, 7(21), 17641-17651. doi: 10.1021/acsomega.1c05881.

[3] Agwu, O.E., Alkouh, A., Alatefi, S., Azim, R.A., & Ferhadi, R. (2024). Utilization of machine learning for the estimation of production rates in wells operated by electrical submersible pumps. *Journal of Petroleum Exploration and Production Technology*, 14(5), 1205-1233. doi: 10.1007/s13202-024-01761-3.

[4] Bello, O., Dolberg, E.P., Teodoriu, C., Karami, H., & Devegowdva, D. (2020). Transformation of academic teaching and research: Development of a highly automated experimental sucker rod pumping unit. *Journal of Petroleum Science and Engineering*, 190, article number 107087. doi: 10.1016/j.petrol.2020.107087.

[5] Büker, J., Laß, A., Werner, P., & Wurm, F.-H. (2021). Active noise cancellation applied to a centrifugal pump in a closed loop piping system. *Applied Acoustics*, 178, article number 108003. doi: 10.1016/j.apacoust.2021.108003.

[6] Carpenter, C. (2019). Analytics solution helps identify rod-pump failure at the wellhead. *Journal of Petroleum Technology*, 71(5), 63-64. doi: 10.2118/0519-0063-JPT.

[7] Castillo, M.A., Gutiérrez, R.H.R., Monteiro, U.A., Minette, R.S., & Vaz, L.A. (2019). Modal parameters estimation of an electrical submersible pump installed in a test well using numerical and experimental analysis. *Ocean Engineering*, 176, 1-7. doi: 10.1016/j.oceaneng.2019.02.035.

[8] Cepeda, F.A., Setiadi, B.W., & Alvarez, G.A. (2024). Evaluating tubing completions using high-resolution gyro logs to improve rod pumping systems run life. In *SPE artificial lift conference and exhibition – Americas*. The Woodlands, Texas: Society of Petroleum Engineers. doi: 10.2118/219558-MS.

[9] El Morsy, M. (2019). Fault diagnosis approach for roller bearings based on optimal Morlet wavelet de-noising and auto-correlation enhancement. *SAE International Journal of Passenger Cars – Mechanical Systems*, 12(2), 127-138. doi: 10.4271/06-12-02-0010.

[10] Erazo-Bone, R., Gacía Vera, R., Chuchuca-Aguilar, F., Ramírez Yagual, J.P., Portilla Lazo, C.A., & Escobar-Segovia, K. (2019). Eliminating gas interference and blockage in sucker rod pumping systems to improve oil production. In M. Botto-Tobar, M. Zambrano Vizuete, P. Torres-Carrión, S. Montes León, G. Pizarro Vásquez & B. Durakovic (Eds.), *Applied technologies* (pp. 110-124). Cham: Springer. doi: 10.1007/978-3-030-42517-3_9.

[11] Fakher, S., Khlaifat, A., & Nameer, H. (2022). Improving electric submersible pumps efficiency and mean time between failure using permanent magnet motor. *Upstream Oil and Gas Technology*, 9, article number 100074. doi: 10.1016/j.upstre.2022.100074.

[12] Fakher, S., Khlaifat, A., Hossain, M.E., & Nameer, H. (2021). A comprehensive review of sucker rod pumps' components, diagnostics, mathematical models, and common failures and mitigations. *Journal of Petroleum Exploration and Production Technology*, 11(10), 3815-3839. doi: 10.1007/s13202-021-01270-7.

[13] Guo, C., Gao, M., & He, S. (2020). A review of the flow-induced noise study for centrifugal pumps. *Applied Sciences*, 10(3), article number 1022. doi: 10.3390/app10031022.

[14] Hao, Z., Zhu, S., Pei, X., Huang, P., Tong, Z., Wang, B., & Li, D. (2019). Submersible direct-drive progressing cavity pump rodless lifting technology. *Petroleum Exploration and Development*, 46(3), 621-628. doi: 10.1016/S1876-3804(19)60042-X.

[15] He, Y.-P., Cheng, H.-B., Zeng, P., Zang, C.-Z., Dong, Q.-W., Wan, G.-X., & Dong, X.-T. (2024). Working condition recognition of sucker rod pumping system based on 4-segment time-frequency signature matrix and deep learning. *Petroleum Science*, 21(1), 641-653. doi: 10.1016/j.petsci.2023.08.031.

[16] Li, H., Niu, H., Zhang, Y., & Yu, Z. (2023). Research on indirect measuring method of dynamometer diagram of sucker rod pumping system based on long-short term memory neural network. *Journal of Intelligent & Fuzzy Systems*, 45(3), 4301-4313. doi: 10.3233/JIFS-230253.

[17] Liu, Z.-M., Gao, X.-G., Pan, Y., & Jiang, B. (2023). Multi-objective parameter optimization of submersible well pumps based on RBF neural network and particle swarm optimization. *Applied Sciences*, 13(15), article number 8772. doi: 10.3390/app13158772.

[18] Marscher, W.D. (2023). Centrifugal pump monitoring, troubleshooting and diagnosis using vibration technologies. In R.X. Perez (Ed.), *Condition monitoring, troubleshooting and reliability in rotating machinery* (pp. 15-76). Beverly: Scrivener Publishing LLC. doi: 10.1002/9781119631620.ch2.

[19] Nascimento, J., Maitelli, A., Maitelli, C., & Cavalcanti, A. (2021). Diagnostic of operation conditions and sensor faults using machine learning in sucker-rod pumping wells. *Sensors*, 21(13), article number 4546. doi: 10.3390/s21134546.

[20] Peng, Y. (2019). Artificial intelligence applied in sucker rod pumping wells: Intelligent dynamometer card generation, diagnosis, and failure detection using deep neural networks. In *SPE annual technical conference and exhibition*. Calgary, Alberta: Society of Petroleum Engineers. doi: 10.2118/196159-MS.

[21] Rzayev, A.H., Aliyev, Y.G., & Rezvan, M.H. (2024). Intelligent intertraverse messdose dynamograph for sucker-rod deep-well pumping units. *Measurement Techniques*, 66(10), 785-793. doi: 10.1007/s11018-024-02292-3.

[22] Sabaa, A., Abu El Ela, M., El-Banbi, A.H., & Sayyouh, M.H.M. (2023). Artificial neural network model to predict production rate of electrical submersible pump wells. *SPE Production & Operations*, 38(1), 63-72. doi: 10.2118/212284-PA.

[23] Sun, Z., Jin, H., Gu, J., Huang, Y., Wang, X., Yang, H., & Shen, X. (2020). Studies on the online intelligent diagnosis method of undercharging sub-health air source heat pump water heater. *Applied Thermal Engineering*, 169, article number 114957. doi: 10.1016/j.applthermaleng.2020.114957.

[24] Syed, F.I., Alshamsi, M., Dahaghi, A.K., & Neghabhan, S. (2022). Artificial lift system optimization using machine learning applications. *Petroleum*, 8(2), 219-226. doi: 10.1016/j.petlm.2020.08.003.

[25] Wei, J., & Gao, X. (2020). Fault diagnosis of sucker rod pump based on deep-broad learning using motor data. *IEEE Access*, 8, 222562-222571. doi: 10.1109/ACCESS.2020.3036078.

[26] Yang, Q., Li, W., Ji, L., Shi, W., Pu, W., Long, Y., & He, X. (2023). Research on the hydrodynamic noise characteristics of a mixed-flow pump. *Journal of Marine Science and Engineering*, 11(12), article number 2209. doi: 10.3390/jmse11122209.

[27] Zhao, H., Liu, D., & He, X. (2021). Bias-compensated sign subband adaptive filter algorithm with individual weighting factors for input noise. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 69(3), 1872-1876. doi: 10.1109/TCSII.2021.3103940.

Оптимізація збору та обробки динамометричних даних для підвищення ефективності нейромережевої діагностики глибинно-насосної штангової установки

Олександр Турчин
Аспірант
Івано-Франківський національний технічний університет нафти і газу
76019, вул. Карпатська, 15, м. Івано-Франківськ, Україна
https://orcid.org/0009-0001-5989-1712

Анотація. Метою дослідження було підвищення точності і швидкості аналізу динамометричних даних шляхом удосконалення методів їх збору та обробки, що сприятиме більш ефективній роботі нейронних мереж у контексті діагностики обладнання. У роботі проведено комплексне дослідження, направлене на підвищення ефективності діагностики глибинно-насосних штангових установок за допомогою нейронних мереж шляхом оптимізації процесів збору та обробки динамометричних даних. Було розглянуто основні проблеми, що виникають під час збору та аналізу даних, такі як наявність шумів, низька якість сигналу, а також велика кількість нерелевантної інформації. На основі цього аналізу були запропоновані методи для покращення якості даних, зокрема фільтрація шумів, нормалізація сигналу та використання алгоритмів для автоматизованого відбору найбільш важливих характеристик. У процесі дослідження було кілька варіантів алгоритмів обробки динамометричних даних, що дозволило досягти значного підвищення точності роботи нейронних мереж. Зокрема, результати показали, що точність діагностики збільшилася на 15 %, а час, необхідний для обробки даних, скоротився на 20 %. Це дозволило покращити загальну продуктивність системи діагностики, зменшивши кількість помилкових висновків і підвищивши надійність роботи глибинно-насосної штангової установки. Результати дослідження показали, що оптимізація збору та обробки динамометричних даних призвела до підвищення точності діагностики та скорочення часу обробки. Застосування комбінованих архітектур нейронних мереж продемонструвало ефективніші результати у порівнянні з традиційними методами. Дані вдосконалення можуть знизити витрати на технічне обслуговування та підвищити ефективність роботи обладнання

Ключові слова: точність вимірювань; обробка сигналів; автоматизація аналізу; фільтрація даних; продуктивність систем