Yessenov Science Journal
SCIENCE: RESEARCH AND PRACTICE

UDC 622.276.53
MRNTI 55.39.29
DOI 10.56525/BTUV4158

METHODOLOGY FOR AUTOMATIC FAULT RECOGNITION OF ROD-WELL PUMPING UNITS

***CHAZHABAYEVA M.**
Caspian University of Technologies
and Engineering named
after Sh. Yessenov
Aktau, Kazakhstan
E-mail: marzhan.chazhabayeva@yu.edu.kz
HUY-TUAN PHAM
Ho Chi Minh City University of
Technology and Education, Vietnam
***Corresponding author: marzhan.chazhabayeva@yu.edu.kz**

Annotation. One of the most effective ways of monitoring is dynamometry. Dynamograms are used to diagnose and monitor the operation of the recognition of rod-well pumping units (RDPU). A new approach to detecting and detecting failures, malfunctions in the operation of the RDPU system, excluding the human factor. The use of Fourier descriptors for the recognition of dynamometric maps and machine learning techniques for the classification of various pump conditions is proposed. With the help of the Fourier descriptor, it is possible to predict and diagnose malfunctions of the downhole pump. These descriptors simplify, normalize, and describe each map well. The proposed method is trained using data from real dynamometric maps.

Keywords: rod-well pumping unit (RDPU), monitoring, failure, malfunctions, dynamometric map, dynamogram, Fourier descriptors, machine learning technique, model, recognition algorithm, data analysis.

Introduction. Currently, the drop in demand for fuel due to the pandemic has caused a crisis in the oil industry. Oil-producing countries, including Kazakhstan, are being forced to reduce oil production in order to prevent oil prices from falling. This will inevitably lead to the suspension or complete shutdown of around 30% of the total number of wells. As 80 per cent of Kazakhstan's well inventory consists of sucker rod well pumping units (SRPUs), the issue of normal, uninterrupted operation of operating wells, in particular their equipment and facilities, becomes acute when 30 per cent of wells are shut down.

There are many ways to improve the efficiency of operation and extend the life of RDPU, one of the most important ways is timely diagnosis and monitoring of RDPU operation. It is necessary to develop methods to predict and diagnose well pump failures.

Materials and methods of research. One of the most effective monitoring methods is dynamometry, the construction of a dynamogram, which provides complete information on the condition of underground equipment and allows you to assess the operating mode of the well [1].

A supervised learning approach is used to train the model. The resulting model can predict the probability of some unfavourable conditions for pump operation and classify them.

Supervised learning algorithms are used to analyse the training data. These algorithms produce an inferred function that is able to represent training instances. In addition, they will be able to correctly define class labels for unseen instances. This is necessary for the learning algorithm to generalise the training data to unseen instances in a "reasonable" way.

Figure 1. Phases of the Knowledge Discovery in the Databases (KDD)

Dynamograms are used to diagnose and monitor the performance of the boom pump. Traditionally, diagnostics using rod pump dynamograms are performed with the help of industry experts. The results of analysing and interpreting dynamograms are directly dependent on the knowledge and experience of the industry experts. There is a huge influence of human factors on the analysis of RDPUs.

Monitoring all pumps in the field by experts is time-consuming and costly. These costs are relatively high compared to the cost of using an automatic diagnostic system. The use of machine learning in automatic pumping system diagnostics improves operational efficiency and enables faster and more accurate detection of equipment malfunctions.

The main challenge in rod pump fault detection is the visual interpretation of the information obtained. In most cases, the difficulty of visual interpretation depends on several factors. These factors include the behaviour of the system itself, which results in similar shapes of the resulting dynamograms. In addition, the problem encountered in analysing the dynamometer chart may be beyond the knowledge and experience of the expert performing the analysis.

A new approach is proposed to detect and identify failures and malfunctions in the operation of the RDPU system, eliminating the human factor (possible errors due to incorrect visual interpretation of dynamograms).

The main idea of the paper is to use Artificial Neural Network (ANN) to recognise and classify RDPU dynamometer map images. The approach involves using Fourier descriptors to recognise dynamometer maps, then uses machine learning techniques to classify different pump conditions. The model, by learning patterns from the input data, recognises the well pump conditions. This data is pre-entered by experts in dynamogram analysis.

This approach works in two ways. Firstly, this approach extracts the Fourier descriptors used in the analytical solution of the rod pump wave equation. Secondly, it aims to detect 13 different states of the dynamometer maps with a large data set collected from all the wells in the field.

It is proposed to divide the recognition algorithm into 4 stages. Each stage has a significant impact on the solution of each individual recognition case and covers all possible failures [3].

The 4 steps proposed are

- ✓ Data collection and analysis;

- ✓ extraction function;
- ✓ classification scheme;
- ✓ testing and evaluation.

The methodology includes the following procedures, which cover all 4 steps above:

- ✓ Review of previous approaches for extracting dynamometer map properties;
- ✓ Review of previous classification methods;
- ✓ The feature extraction algorithm that gives the best results;
- ✓ Collecting real dynamometer maps classified by experts;
- ✓ Preparing this data using a feature extraction algorithm;
- ✓ Using supervised machine learning algorithms;
- ✓ Proposing a specific model based on validation error;
- ✓ Evaluating this model with respect to different RDPU conditions;
- ✓ Testing the developed model.

The first step is to decide on the number of variables and which variables to use.

The choice of initial variables should be made by experts in the field. All variables that can affect the performance of the RDPU are important. At this stage, the experience and knowledge of the experts is important to imagine which data are significant.

Initial variables used: dynamometer chart data, pump fill, torque, polished rod load, total fluid flow, plunger (piston) stroke length, etc.

An important prerequisite at this stage is good dynamometry. Dynamometry is a powerful tool for monitoring the whole RDPU system. Analysis of the dynamometer results will be the main variable for this methodology.

Feature extraction is one of the most important elements for the successful feature recognition of dynamometer cards. By definition, a feature is a noticeable part or characteristic of an object for the purpose of distinguishing it. Feature extraction is the process of creating features [4].

It is proposed to transform the input data into a reduced set of features or a vector of features. This process is called feature selection. The selected features are expected to contain relevant information from the input data. This will facilitate the subsequent learning steps**.**

The model developed by this methodology performs a transformation of the dynamometer map data (x, y). This transformation transforms the map from observation space to an n-dimensional vector. The resulting vector contains characteristic elliptic Fourier descriptors. These descriptors simplify, normalise and describe each map well.

Elliptic Fourier functions were developed by Kuhl and Giardina (1982) [5]. They presented a new approach to the old problem of numerical characterisation of complex shapes.

In essence, elliptic Fourier descriptors are a parametric solution to the problem of numerically determining complex shapes. A two-dimensional contour (x, y) can be divided into two curves. The division is performed by plotting the contour function of the third variable (t). The first curve is a plot of the dependence of the x-axis on the map (t). The second curve is a plot of the dependence of the Y-axis on the map (t). A pair of equations is then obtained as a function of the third variable (t). The parameters of these two equations are used in the feature vector.

The closer the distance between the observed points, the more accurate the map. The X and Y axes can be defined as a function of time (third variable). These new functions are single-valued and periodic. They can therefore be fitted with Fourier functions [6].

Figure 2 shows an example of a real dynamometer graph of a piston pump. Consider the x and y coordinate values in Figure 2 plotted on a new horizontal 'x-axis', labelled the 't-axis' (time axis). These x and y "projections" are shown in Figures 3 and 4.

Examination of these figures shows that they will always be single digit and periodic. It is emphasised that Figures 3 and 4 are fitted with a Fourier function,[7]. One advantage of using this algorithm is that the map can be reconstructed. x- Coordinates and y-coordinates that were computed separately can be "reunited" with these coordinates (for identical values of t). The expected shape is then recreated.

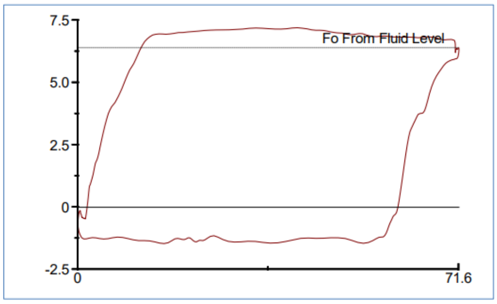

Figure 2: Example of a real dynamometer map

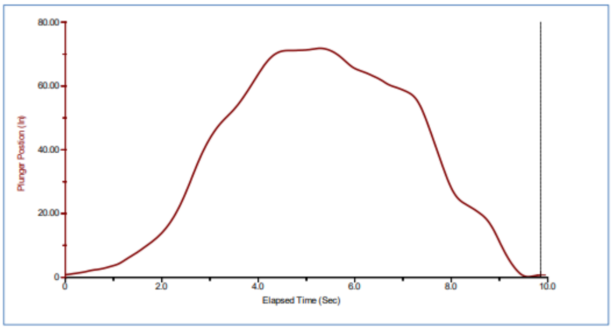

Figure 3: Projection of the x-axis of Figure 2 onto the third variable t

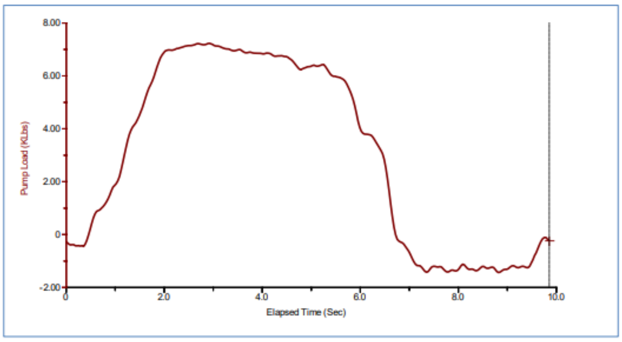

Figure 4: Projection of the y-axis of Figure 2 onto the third variable t

In summary, the elliptic Fourier descriptors represent a parametric solution. This implies the derivation of a pair of equations [8]. These two equations are a function of a third variable (X (t) $and\ Y\ (t)$), as shown in equation (1).

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \sum_{n=0}^{k} \begin{pmatrix} a_n b_n \\ c_n d_n \end{pmatrix} \begin{pmatrix} \sin\left(\frac{2n\pi t}{T}\right) \\ \cos\left(\frac{2n\pi t}{T}\right) \end{pmatrix} \tag{1}$$

Where:
t: the third assumed variable.
n: the harmonic number.
a_n, b_n, b_n, c_n, d_n: Fourier coefficients corresponding to nth-harmonics.
T: period (sum of all t increments).

The results of the study. *Step one: extracting the boundary*

The map file is represented as a list of (x, y) items of the entire map. xp and yp are defined as the summation of the links between each subsequent point. The sum of links between points can be calculated from Equations 2 and 3.

$$x_p = \sum_{i-1}^{p} \Delta x_i \tag{2}$$

$$y_p = \sum_{i-1}^{p} \Delta y_i \tag{3}$$

Where:
p: chain link index.
x_p: summation of links on the x-axis.
y_p: summation of y-axis links.

Step two: determination of the third parameter

The length of the links is called "time". The contribution of any diagonal link is $\sqrt{x^2 + y^2}$. "Time" of a path link is the sum of all times of previous links, as shown in Equation 4.

$$t_p = \sum_{i-1}^{p} \Delta t_i \tag{4}$$

Where:
p: the index of the chain link.
t_p: chain length in the track link.

Step three: Fourier descriptors

Equation (1) is solved to obtain a_n and b_n as shown in Equations 5 and 6:

$$a_n = \frac{T}{2n^2\pi^2} \sum_{i-1}^{p} \frac{\Delta x_p}{\Delta t_p} \left[cos\left(\frac{2n\pi t_p}{T}\right) - cos\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \tag{5}$$

$$b_n = \frac{T}{2n^2\pi^2} \sum_{i-1}^{p} \frac{\Delta x_p}{\Delta t_p} \left[sin\left(\frac{2n\pi t_p}{T}\right) - sin\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \tag{6}$$

Where:
n: harmonic number.
a_n, b_n: Fourier coefficients corresponding to the nth-harmonics.
T: period (sum of all t increments).
k: total number of links.
p: chain link index.
x_p: summation of links along the x-axis.

t_p: chain length in the path link.

The same procedures are applied for the y-axis. Accordingly, c_n and d_n are identified as shown in Equations 7 and 8.

$$c_n = \frac{T}{2n^2\pi^2} \sum_{i-1}^{p} \frac{\Delta y_p}{\Delta t_p} \left[cos\left(\frac{2n\pi t_p}{T}\right) - cos\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \quad (7)$$

$$d_n = \frac{T}{2n^2\pi^2} \sum_{i-1}^{p} \frac{\Delta y_p}{\Delta t_p} \left[sin\left(\frac{2n\pi t_p}{T}\right) - sin\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \quad (8)$$

Where:
n: harmonic number.
c_n, d_n: Fourier coefficients corresponding to the nth-harmonics.
T: period (sum of all t increments).
k: total number of links in the chain.
p: chain link index.
y_p: summation of links along the y-axis.
t_p: chain length of a path link.

Step four: machine learning input vector

It should be emphasized that the input parameters are 15 elliptic Fourier descriptors that provide a normalized set of coefficients a_n, b_n, c_n, and d_n.

These coefficients are invariants of rotation, displacement and scale. Hence, a matrix of values [15 × 4] is created. The first coefficient refers to the centroid of the input shape.

Therefore, this can be ignored. This leaves the remaining 59 EFD coefficients. These 59 descriptors can be used to compare shapes.

Hence, the feature vector $\chi = (a_2, b_2, c_2, d_2, ..., a_{15}, b_{15}, c_{15}, d_{15})$ can be constructed and used as input to the classification algorithm.

Conclusion. The model developed by this methodology performs an (x,y) transformation of the dynamometer map data. This transformation transforms the map from observation space to an n-dimensional vector. The resulting vector contains characteristic elliptic Fourier descriptors. These descriptors simplify, normalise and describe each map well. The proposed technique is trained using data from real dynamometer maps.

REFERENCES

[1] Chazhabaeva M.M./The technique of automatic fault recognition of the SSNU/Proceedings of the University No.3 (88), 2022. pp.166-171.

[2] Molchanov A. G. Puti dal`nejshego sovershenstvovaniya shtangovy`x skvazhinny`x nasosny`x ustanovok / Moskva: Specializirovanny`j zhurnal «Burenie i neft`», 2014 g. – str. 2-4.

[3] Bubnov M. V., Zyuzev A. M. Sredstva diagnostirovaniya oborudovaniya ustanovok shtangovy`x glubinny`x nasosov / UralE`NIN, FGAOU VPO «UrFU» - Konferenciya molody`x ucheny`x, 2016 – str. 2-3.

[4] Andrew, N., 2015, "Lecture Notes on Machine Learning, Dimensionality Reduction", Vol. 10.

[5]Anthony Allison (): Accurate Load & Position Measurement Is Critical to Quality Dynamometer Analysis. Weatherford, updated on 9/18/2015, checked on 1/22/2019. - 2015

[6]Bottone, S., Lee, Y., and O'Sullvin, M., "Failure Prediction and Diagnosis for Satellite Monitoring Systems Using Bayesian Network", SPIE 6736, Unmanned Sensors and Sensor Networks IV, 5 October., 2008

[7] Oshi, Prateek, Python. Real world machine learning: learn to solve challenging data science problems by building powerful machine learning models using Python / Prateek Joshi [and four others]. Birmingham, UK: Packt Publishing. – 2016.
[8] Ramez Maher Aziz Zaky Abdalla, «Automatic well failure analysis for the sucker rod pumping systems», Faculty of engineering, CAIRO UNIVERSITY GIZA, EGYPT, 2018.

Чажабаева М.М.[1], Huy-Tuan Pham[2]
[1]Ш. Есенов атындағы Каспий технологиялар және инжиниринг университеті, Казахстан, Ақтау
[2]Хо Ши Мин технологиялық және білім беру университеті, Вьетнам

ШТАНГАЛЫҚ-ҰҢҒЫМАЛЫҚ СОРҒЫ ҚОНДЫРҒЫЛАРЫНЫҢ АҚАУЛАРЫН АВТОМАТТЫ ТҮРДЕ ТАНУ ӘДІСТЕМЕСІ

Аңдатпа. Мониторингтің тиімді әдістерінің бірі-динамометрия. Динамограммалар ШТСҚ жұмысын диагностикалау және бақылау үшін қолданылады. Адам факторын болдырмайтын ШТСҚ жүйесінің жұмысындағы ақаулықтарды, істен шығуларды анықтаудың және анықтаудың жаңа тәсілі. Фурье дескрипторларын динамометриялық карталарды тану үшін және сорғының әртүрлі жағдайларын жіктеу үшін машиналық оқыту техникасын қолдану ұсынылады. Фурье дескрипторының көмегімен ұңғыма сорғысының ақауларын болжауға және диагноз қоюға болады. Бұл дескрипторлар әр картаны жеңілдетеді, қалыпқа келтіреді және жақсы сипаттайды. Ұсынылған әдіс нақты динамометриялық карталардың деректерін қолдана отырып оқытылады.

Кілт сөздер: штангалық тереңдік сорғы қондырғысы (ШТСҚ), мониторинг, істен шығу, ақаулар, динамометриялық карта, динамограмма, Фурье дескрипторлары, Машиналық оқыту техникасы, моделі, тану алгоритмі, деректерді талдау**.**

Чажабаева М.М.[1], Huy-Tuan Pham[2]
[1]Каспийский университет технологий и инжиниринга им. Ш.Есенова, Казахстан, Актау
[2]Хо Ши Мин технологический и образовательный университет, Вьетнам

МЕТОДИКА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ НЕИСПРАВНОСТЕЙ ШТАНГОВО-СКВАЖИННЫХ НАСОСНЫХ УСТАНОВОК

Аннотация. Одним из эффективных способов мониторинга является динамометрия. Динамограммы используются для диагностики и мониторинга работы ШСНУ. Новый подход обнаружения и выявления отказов, неполадок в работе системы ШСНУ, исключающий человеческий фактор. Предлагается использование дескрипторов Фурье для распознавания динамометрических карт и техники машинного обучения для классификации различных условий насоса. С помощью дескриптора Фурье можно спрогнозировать и диагностировать неисправности работы скважинного насоса. Эти дескрипторы упрощают, нормализуют и хорошо описывают каждую карту. Предложенная методика обучается с использованием данных реальных динамометрических карт.

Ключевые слова: штанговая глубинная насосная установка (ШСНУ), мониторинг, отказ, неполадки, динамометрическая карта, динамограмма, дескрипторы Фурье, техника машинного обучения, модель, алгоритм распознавания, анализ данных.