Journal of Petroleum Exploration and Production Technology (2024) 14:343–363
https://doi.org/10.1007/s13202-023-01710-6

ORIGINAL PAPER-PRODUCTION ENGINEERING

Anomaly detection in oil-producing wells: a comparative study of one-class classifiers in a multivariate time series dataset

Wander Fernandes Jr.[1,2] · **Karin Satie Komati**[2] · **Kelly Assis de Souza Gazolli**[2]

Received: 2 January 2023 / Accepted: 26 September 2023 / Published online: 8 November 2023

Abstract

Anomalies in oil-producing wells can have detrimental financial implications, leading to production disruptions and increased maintenance costs. Machine learning techniques offer a promising solution for detecting and preventing such anomalies, minimizing these disruptions and expenses. In this study, we focused on detecting faults in naturally flowing offshore oil and subsea gas-producing wells, utilizing the publicly available 3W dataset comprising multivariate time series data. We conducted a comparison of different anomaly detection methods, specifically one-class classifiers, including Isolation Forest, One-class Support Vector Machine (OCSVM), Local Outlier Factor (LOF), Elliptical Envelope, and Autoencoder with feedforward and LSTM architectures. Our evaluation encompassed two variations: one with feature extraction and the other without, each assessed in both simulated and real data scenarios. Across all scenarios, the LOF classifier consistently outperformed its counterparts. In real instances, the LOF classifier achieved an F1-measure of 87.0% with feature extraction and 85.9% without. In simulated instances, the LOF classifier demonstrated superior performance, attaining F1 measures of 91.5% with feature extraction and 92.0% without. These results show an improvement over the benchmark established by the 3W dataset. Considering the more challenging nature of real data, the inclusion of feature extraction is recommended to improve the effectiveness of anomaly detection in offshore wells. The superior performance of the LOF classifier suggests that the boundaries of normal cases as a single class may be ill-defined, with normal cases better represented by multiple clusters. The statistical analysis conducted further reinforces the reliability and robustness of these findings, instilling confidence in their generalizability to a larger population. The utilization of individual classifiers per instance allows for tailored hyperparameter configurations, accommodating the specific characteristics of each offshore well.

Keywords Oil well monitoring · 3W dataset · LSTM (Long short-term memory) · One-class Support Vector Machine (OCSVM) · Local Outlier Factor (LOF) · Elliptical envelope

✉ Karin Satie Komati
kkomati@ifes.edu.br

Wander Fernandes Jr.
wanderfj@gmail.com

Kelly Assis de Souza Gazolli
kasouza@ifes.edu.br

[1] Petrobras, Av. Nossa Sra. da Penha, 1688 - Barro Vermelho, Vitória, ES 29057-550, Brazil

[2] Graduate program in Applied Computing (PPComp), Instituto Federal do Espírito Santo, Campus Serra, Av. dos Sabiás, 330 - Morada de Laranjeiras, Serra, ES 29166-630, Brazil

List of symbols

Latin symbols

b	Input vector of bias in Autoencoder neural network, Dimensionless
b'	Output vector of bias in Autoencoder neural network, Dimensionless
c	Vector of bias in Autoencoder neural network, Dimensionless
$\tilde{c}_t$	Is a temporary variable that contains the relevant information in the current timestep t
f	Forgetting gate of LSTM, Dimensionless
h	State of the hidden units of the network, Dimensionless
i	Input gate of LSTM, Dimensionless
k	Number of neighbors to be considered in LOF algorithm, Dimensionless

M	Length of multivariable time series, Dimensionless
n	Subset of observations in Elliptical Envelope algorithm, Dimensionless
o	Output gate of LSTM, Dimensionless
p	For the Wilcoxon test, a p-value is the probability of getting a test statistic as large or larger assuming both distributions are the same
P	Precision, Dimensionless
R	Recall, Dimensionless
t	Value of threshold in Elliptical Envelope algorithm selected by "fastmcd" algorithm, Dimensionless
T	Length of univariate time series, Dimensionless
U	Weight matrices for output layer in RNN, Dimensionless
V	Weight matrices for hidden layer in RNN, Dimensionless
W	Input weight matrices for input layer in RNN, Dimensionless
W'	Output weight matrices for input layer in RNN, Dimensionless
x_i	Real value in univariate time series, unit of measure
X	Input data of an autoencoder neural network, unit of measure
X'	Output data of an autoencoder neural network, unit of measure
X^i	Univariate time series, unit of measure
y_t	Output of softmax function in RNN
Z	Reduced or latent dimension in Autoencoder neural network. Dimensionless

Greek symbols

γ	OCSVM parameter that influences the radius of the Gaussian hypersphere, Dimensionless
μ	Mean, unit of measure
ν	OCSVM parameter that is used to control the sensitivity of support vectors, Dimensionless
ρ	A statistical measurement used to validate a hypothesis against observed data, Dimensionless
σ	Sigmoid function or Standard deviation, unit of measure
σ^2	Variance, square of unit of measure
Σ	Covariance matrix, Dimensionless
φ	Activation function in Autoencoder neural network, Dimensionless

Abbreviations

ABOD	Angle-based Outlier Detection
AI	Artificial intelligence
BSW	Basic sediment and water
CBM	Condition-based monitoring
CKP	Choke for production
DHSV	Down Hole Safety Valve
DLSTM	Deep long short-term memory
ELM	Extreme learning machine
F1	F1 score or the F1 measure
FN	False negative
FP	False positive
LOF	Local Outlier Factor
LSTM	Long short-term memory
MCD	Minimum Covariance Determinant
MEMD	Multiple empirical mode decomposition
ML	Machine learning
OCSVM	One-class Support Vector Machine
PDG	Permanent downhole manometer
P-JUS-CKGL	Fluid pressure downstream of gas lift
P-MON-CKP	Fluid pressure upstream to valve CKP
P-PDG	Fluid pressure at PDG
P-TPT	Fluid [Pressure at TPT
QGL	Flow of gas lift
RNN	Recurrent neural network
SVM	Support Vector Machine
T-JUS-CKGL	Fluid temperature downstream of gas lift
T-JUS-CKP	Fluid temperature downstream to CKP valve
TN	True negative
TP	True positive
TPT	Temperature transducer
T-TPT	Fluid temperature at TPT

Introduction

During oil and gas production, unwanted events called anomalies can cause significant financial impacts. Considering the average production per well in the pre-salt of 18 Mbbl/d (thousands of barrels per day) (ANP 2020) and the average oil price of U$110.93 (MacroTrends 2022), the loss of revenue in the event of an anomaly that interrupts the production of a well is on the order of U$1 million dollars per day in Brazil. Additionally, vessels that carry out repairs in damaged wells (called rigs) have high costs that reach U$500 thousand dollars per day.

Condition-based monitoring (CBM) is a strategy that verifies the condition of a system or equipment during its continuous operation (Marins et al. 2021). Understanding the behavior of fluid flow in porous media, including the phenomena of channeling and fault effects, is crucial in CBM oil and gas applications. Fluid flow in porous media

refers to the movement of fluids, such as liquids or gases, through a porous material, such as soil, rock, or sediment (Soltanmohammadi et al. 2021). Channeling, in the context of fluid flow in porous media, refers to the preferential flow of fluid through specific pathways or channels within the porous material. Instead of uniformly spreading throughout the porous media, the fluid concentrates its flow in certain areas, resulting in uneven distribution and potentially reduced effectiveness of fluid transport or extraction processes. Faults are fractures or planes of weakness in the Earth's crust where the movement has occurred. They can affect fluid flow in several ways. Firstly, faults can act as conduits or barriers for fluid movement, either facilitating or obstructing the flow of fluids through the porous media. Secondly, fault zones can introduce permeability variations, leading to preferential fluid flow along the fault surfaces or altered flow patterns within the porous media. Lastly, faults can also influence the overall structure and geometry of the porous media, affecting the flow dynamics and distribution of fluids.

Automatic monitoring of the oil production process could detect and prevent anomaly events. Detecting anomalies is a hard task because it does not have a set of characteristics or rules that aggregate them and there are a lot of challenges in fault detection for process signals of traditional methods (Alrifaey et al. 2021). An anomaly can be occasional, a single extreme value (such as a temperature) above a threshold can be enough to characterize an anomaly. A sudden change in temperature during an industrial process can be regarded as abnormal, even if the initial and final values of the change are not atypical in isolation (Chandola et al. 2009). In many industrial processes, we seek to detect irregular patterns, in which most observations refer to typical situations, and the minority, to rare situations that we want to identify (Santos and Kern 2016).

Anomaly detection is a binary classification problem (as normality and abnormality) (Castro et al. 2021). A possible solution to predict an anomaly is the application of multivariate statistics (Soriano-Vargas et al. 2021) and machine learning (ML) methods (D'Almeida et al. 2022). In industrial processes, the input data for monitoring come from several sensors indexed by time, that is, they are multivariate time series. As written by Fawaz et al. (2019), time series classification has been considered one of the most challenging problems in data mining. The detection of new patterns (novelty detection) can be done with classifiers of a single class, in which only data associated with the common class (normality) are used in the training (Khan and Madden 2014). The challenges associated with petroleum time series data, as mentioned by (Sagheer and Kotb 2019), include excessive noise, defects, anomalies, high dimensionality, non-stationarity, variable trends, and the nonlinear and heterogeneous nature of reservoir properties.

This work applied and compared machine learning techniques to detect anomalies in oil-producing wells, using the 3W dataset composed of multivariate time series. The present work used the benchmark for anomaly detection proposed by Vargas et al. (2019) and extended it with more classifiers Local Outlier Factor (LOF), Elliptical Envelope, and neural networks of the type Autoencoder with layers feedforward and recurrent LSTM type (Long Short-Term Memory), besides the hyperparameter calibration step.

Tariq et al. (2021)'s work provided a detailed and comprehensive review of data science and ML roles in different petroleum engineering. The work of Tariq et al. (2021) also brings a discussion about the limitations of the ML model, and one of the limitations is the availability of data addressed through the use of a public and annotated database. Vargas et al. (2019) made the 3W database public, which is composed of multivariate time series. The 3W dataset contains 1984 time series instances of the production of surge-type offshore oil wells (wells that manage to flow the produced fluids to the platform with their pressure). These instances are separated into normal conditions and anomalies, and the anomalies are organized into eight classes. This base can be used both for detecting and classifying anomalies in oil wells. In addition to the base with real production data, Vargas et al. (2019) also developed two specific benchmarks, one for evaluating the impact of using simulated and hand-drawn instances and another for detecting anomalies. In the benchmark for anomaly detection, the Isolation Forest (with F1 metric of 0,727) and OCSVM - One-class Support Vector Machine (with 0,47 of F1 metric) techniques were used.

Experiments are carried out with and without the feature extraction step. In the experiments with feature extraction, the median, mean, standard deviation, variance, maximum, and minimum are extracted for each variable. In the experiments without feature extraction, the time series themselves are the input for the classifiers.

Friedman's and Wilcoxon's statistical tests assess whether the tested classifiers generate performance metrics whose average is different from the others (Demšar 2006).

The main contributions of the present work are:

1. Anomaly detection in oil wells using one-class classifiers;
2. Comparison of one-class classifiers with and without feature extraction to assess their performance;
3. Comparison of multiple one-class classifiers in anomaly detection using the 3W database, including five classifiers for experiments with feature extraction and six for experiments without feature extraction;
4. Evaluation of the effectiveness of different one-class classifiers in detecting anomalies in both real-world and simulated scenarios. This approach enabled us to examine the impact of feature extraction on detection

performance across datasets with diverse characteristics; and
5. Utilization of statistical tests at the 5% significance level by Wilcoxon's statistical tests with Bonferroni adjustment to validate the results and ensure their generalizability to a larger population.

This paper is organized as follows: Section 2 describes related works; Section 3 details the material (3W database) and methods for each step of the experimental methodology; Section 4 presents the results and analyzes the performance of the proposed system, and Section 7 closes the paper emphasizing its conclusions.

Related works

This section presents two subsections. The first subsection describes articles that use anomaly detection techniques in several applications and the second, articles that used the 3W dataset.

Anomaly detection

The work of Chandola et al. (2009) is an important survey article on anomaly detection. It presents contributions and discussions about the concept of anomaly, and its different aspects in each application domain, providing a structured overview, grouping existing techniques into different categories, and identifying the advantages and disadvantages of each one. It also discusses the computational complexity of the techniques. While Chandola et al. (2009) extensively discuss the concept of anomalies and provide a structured overview of existing techniques across various application domains, our research focuses specifically on the detection of faults in naturally flowing offshore oil and subsea gas-producing wells. By narrowing the scope, we aim to evaluate the performance of different one-class classifiers in this specific domain. Unlike previous studies, our work compares the performance of these classifiers with and without feature extraction, examines their effectiveness in both simulated and real instances, and conducts rigorous statistical analysis to validate the results and ensure their generalizability.

Barbariol et al. (2019) proposed an anomaly detection approach in metering modules. This equipment is an important tool in the oil and gas sector, as it simultaneously provides real-time data on the flows of oil, gas, and water. The Cluster-Based Local Outlier Factor and Isolation Forest algorithms were used to detect quality changes in the measurements performed, using a semi-synthetic dataset. In contrast to their work, our research addresses the detection of faults in naturally flowing offshore oil and subsea gas-producing wells. The application is different but used the same one-class classifiers, LOF and Isolation Forest, as this work.

Chan et al. (2019) performed anomaly detection in programmable logic controllers that make up supervisory control and data acquisition systems. Such equipment manages sensor-based industrial equipment operations and is exposed to cyber threats. A case study involving a traffic light simulation was carried out which demonstrated that anomalies are detected with high precision using One-class SVM. The application is different but used the same one-class classifiers, OCSVM, as this work.

Khan et al. (2019) applied anomaly detection techniques in unmanned aerial vehicles. In the experiments, the Aero-Propulsion System Simulation database was used and then, more experiments were carried out in a real vehicle. The anomaly detection technique was the Isolation Forest algorithm. The application is different but uses the same one-class classifier, Isolation Forest, as this work.

Tan et al. (2020) compared the performance of several classifiers for anomaly detection in marine vessel machines. The safety and reliability of navigation depend on the performance of these machines, and intelligent condition monitoring is essential for maintenance activities. A dataset from a ship's gas turbine propulsion system was the input in the experiments. They investigated the performance of single class classifiers: OCSVM, Support Vector Data Description, Global K-Nearest Neighbors, LOF, Isolation Forest, and ABOD (Angle-based Outlier Detection). The OCSVM algorithm obtained the best accuracy results. As this work, Tan et al. (2020) specifically investigated the performance of one-class classifiers. Their findings revealed that the OCSVM algorithm achieved the best accuracy results, which is used in this experiment.

Grashorn et al. (2020) describe the use of neural networks to detect anomalies in the operation of the Columbus module of the International Space Station. It is a scientific laboratory that transmits around 17,000 telemetry parameters per second to Earth. The Columbus Control Center operations team, in collaboration with Airbus, monitors these parameters and uses autoencoder-type algorithms with LSTM-type cells to support the detection of anomalies during the center's workflow of control. This research highlights the successful application of deep learning techniques in anomaly detection. We also evaluate the performance of LSTM architecture.

Said Elsayed et al. (2020) used a combination of neural network structure autoencoder with LSTM type cells, together with the OCSVM algorithm, to model the normal data flow in a computational network. The experiments showed that the proposed model can efficiently detect the anomalies presented in the network traffic data. As the work of Said Elsayed et al. (2020), we also evaluate the performance of the OCSVM algorithm and LSTM type cells.