Fault detection with Stacked Autoencoders and pattern recognition techniques in gas lift operated oil wells

Rodrigo Scoralick Fontoura do Nascimento[1], Bruno Henrique Groenner[2], Ricardo Emanuel Vaz Vargas[3], Ismael Humberto Ferreira dos Santos[3]

[1]*Programa de Pós-Graduação em Engenharia de Sistemas e Automação*
Universidade Federal de Lavras, MG, Brasil
rodrigo.nascimento2@estudante.ufla.br
[2]*Departamento de Automática*
Universidade Federal de Lavras, MG, Brasil
brunohb@ufla.br
[3]*Petróleo Brasileiro S.A., RJ, Brasil*
ricardo.vargas@petrobras.com.br
ismaelh@petrobras.com.br

Abstract. The offshore industry is responsible for most of the oil and gas production in Brazil. When the level of complexity in this industry is high, it has been a precursor to new technologies in recent years. The main objective of the present work is the development of a system for the detection and classification of failures in oil production wells operated with elevation by gas lift. Stacked autoencoders are used and pattern recognition techniques for fault classification, verifying performance metrics and applying cross-validation to check the generalization of the models for the available observations. After the development of the classifiers, high recall values were obtained (much higher than 0.88), which shows the great applicability of the proposed system in detecting failures in offshore production wells.

Keywords: Fault detection; Oil well monitoring; Multivariate time series classification; Cross validation; Pattern recognition.

1 Introduction

In the current scenario, the oil and gas industry has become more demanding in all areas of engineering, including safety and production. Several aspects must be taken into consideration in the oil and gas area, as it is a very complex industrial area, encompassing several engineering areas that are related in search of better quality processes and products, adding technology and innovation over the years development, as addressed by [1].

The oil extraction industry is divided into two types of production, onshore and offshore. The first is based on production on land, on the mainland. In the second modality, production is carried out offshore through oil extraction platforms, normally far from the continent and in deep waters. The application focus of this work is offshore oil and gas wells, that is, offshore wells.

Deepwater oil wells are classified in two ways, upstream and non-upstream. Non-emergent wells need methods to assist fluid flow (water, oil, gas and sediments). The surgers, on the other hand, are able to carry out the flow of production fluids with their own pressure. In other words, in emerging wells there is a natural rise in fluids [2].

The process of artificial lifting by gas lift consists of the gasification of the production column using natural gas in order to reduce the average density of the fluid being produced in the reservoir [3]. It is a complex process and subject to several failures, whether in actuators and sensors, and also according to the characteristics of each production well.

The occurrence of failures in oil production wells with gas lift can generate losses of thousands of dollars for producing companies, in addition to the complex operation that follows such an occurrence, so that normal operation is reestablished. The oil industry believes that prognosis of anomalies in oil-producing wells can help

reduce maintenance costs, as well as prevent production losses and environmental and human life accidents [4].

One way to predict the occurrence of these failures is the implementation of pattern recognition systems based on Computational Intelligence techniques. Failure detection seeks to expose possible deviations presented in a process, based on its monitored variables. With the advent of measurement systems, process variable values could be obtained and stored in large quantities and with greater precision, allowing for more efficient monitoring [5].

As an example, [6] implemented a failure detection system in gas lift wells based on Artificial Immune Systems, dividing into two operating patterns, normal and abnormal. [7] have adopted techniques such as PCA (Principal Component Analysis) in the treatment of data from oil wells and the Random Forests classification technique to detect hydrate accumulation failures in production lines or injection of emerging wells. these failures cataloged by experts in the field of Petroleum Engineering. These works are similar to this article, as they all aim to find disturbances that cause an abnormality in the operation of real oil production wells.

Several other Computational Intelligence techniques can be implemented to detect failures in industrial processes. As shown by [8] which uses the output of two autoencoders as inputs to a multilayer perceptron network (MLP) for detection of broken bars in three-phase induction motors. In the work presented by Wen et al. [9] autoencoders were applied together as an MLP neural network in the detection of faults that cause unbalance in marine current turbines.

This work aims to implement a pattern recognition system based on Computational Intelligence techniques, making the process more analytical and less operational, which is one of the main objectives of Industry 4.0, through disruptive technologies [10]. The main idea of the article is the identification of failures in production wells with artificial elevation by gas lift. Where their origins are not known, but were determined through inference from process operators due to production losses. These failures are distinct from the normal and stable operation of a well, through data collected from an offshore oil extraction platform.

This work is structured as follows. In the next section, the gas lift oil extraction process and the theoretical foundation of the computational tools used in the study are presented. Section 3 presents the work development methodology. In Section 4 the results and discussions are presented. Finally, in Section 5, the conclusions are presented.

2 Theoretical Foundation

2.1 Process

The Figure 1 shows a simplified diagram of the piping and instrumentation diagram P&ID (Piping and Instrument Diagram) of a production well that uses artificial lifting by gas lift. Nascimento et al. [11] presents variables and their units of magnitude in a table. Being the high pressure gas coming from the gas header on the platform (instruments marked by 4) which is injected through the ring between the piping and the coating chain until reaching an orifice valve located downstream at the bottom of the piping.

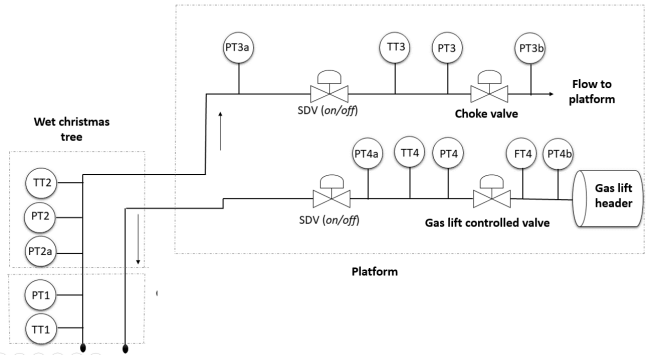

Figure 1. Simplified P&ID diagram of a production well operating with artificial elevation by gas lift

The fluid density is then reduced and the reservoir pressure raised enough to transport the mixture of oil, gas, water and sediment to the platform. On the ocean floor, a set of valves and adapters known as a wet christmas tree (WCT) control the flow of production to the platform. On the platform, an SDV (shutdown valve) is available to interrupt production during an emergency situation and a choke valve, which regulates the production flow rate.

Different flow dynamics are obtained depending on the gas rise pressure (PT4 and PT4a) and downhole pressure (PT1) values.

The WCT is a subsea set consisting of several valves remotely operated by means of hydraulic commands, which contain, for example, the TPT (Pressure and Temperature Transmitter) sensors and the sensor that measures the annular pressure in the gas lift valve (PT2a). Pressures and temperatures are also measured as needed, as well as upstream and downstream of the SDV, choke valve and gas lift injection valve.

The choke valve is a downstream flow control instrument. The gas lift valve, on the other hand, has the function of controlling the pressure received from the inlet header. These high pressures come from the gas compressors of the installation itself, these equipments mentioned in this paragraph are installed on the offshore production platform. A more detailed description of this process is provided by [12].

2.2 Autoencoder

The autoencoder is an Artificial Neural Network (ANN) type that is formed by three layers, the encoder consisting of the first two layers and the last two configuring the decoder, as shown in Figure 2. The autoencoder has the function of mapping as close as possible the input to its output layer. Usually autoencoders have in their hidden layer a lower number of neurons compared to their input and output layers. This is beneficial in relation to reducing the dimensionality of the data, which makes the autoencoder use only the main characteristics of the input data, in order to eliminate descriptors of little relevance to the models. In addition to reducing the dimension, the autoencoder also transforms the data non-linearly, providing the maximization of differences between classes.

In Figure 2 the structure of an autoencoder is presented, where the input data are $x = (x_1, x_2, ...x_n)$, the autoencoder output values are $z = (z_1, z_2, ...z_n)$, with n being the number of neurons in both the input and output layers. The vector $h = (h_1, h_2, ...h_m)$ is the representation of the input x in the hidden layer after using a sigmoid activation function (sf) and m is the number of neurons in the hidden layer. The equations that govern this type of model are described by: eliminating descriptors of little relevance to the models. In addition to reducing the dimension, the autoencoder also transforms the data non-linearly, providing the maximization of differences between classes. The autoencoder optimization functions are presented in [13], [8], [14] and [11]

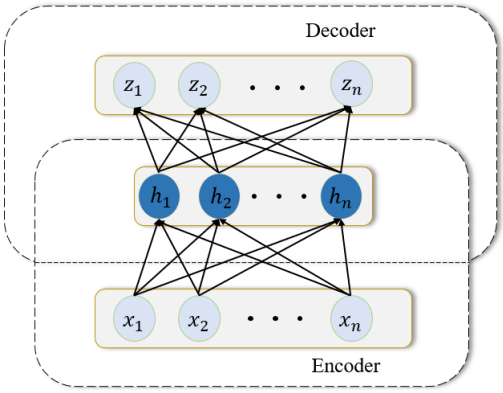

Figure 2. Structure of an autoencoder

3 Materials and methods

3.1 Experimental Data from a Well

Data acquisition was performed on an oil platform offshore, extracted from the plant information system PI System by OSIsoft, widely used in the oil industry. PI System consists of a system that stores process plant information. Through this system, data were collected in approximately 90 consecutive days, making up an observation window of an oil production well. Data extraction took place between 12/06/2018 to 03/06/2019, totaling 129592 observations, with a sampling interval of 1 minute.

Sensors and actuators in the process plant are divided into top variables and bottom variables in the supervisory. Since, for the top variables, their sensors are located on the oil production platform, whereas for the bottom

variables, the instruments are installed on the seabed and in the production column, which may be more susceptible to noise and failures .

Failure diagnosis can be identified from certain variables inherent to the process. These variables, which change in the occurrence of failures according to the intrinsic characteristics of the wells, as well as the physical and chemical conditions, can vary from well to well. For the well studied in this work, an undefined Soft Fault was observed, that is, its origin is unknown, between 02/07/2019 and 02/19/2019, in which the fluid flow decreases, slowly reducing production to complete cessation, as can be seen in Figure 3.

The characteristics of PDG pressure, temperature in the wet Christmas tree and temperature upstream of the choke valve, in order to exemplify this type of anomaly with these three sensors, are presented in Figure 4, where they indicate an abnormality in the behavior of the well described as below:

- PDG Pressure (PT1): tends to be high from the occurrence of a failure, due to an obstruction along the fluid flow path, increasing the pressure in the production column;
- Wet Christmas tree temperature (TT2): tends to balance with the seabed temperature, in the case of fluid flow failures, in deep water these temperatures normally vary between 2°C and 6°C;
- Temperature upstream of the choke valve (TT3): tends to equal the temperature on the surface of the sea;

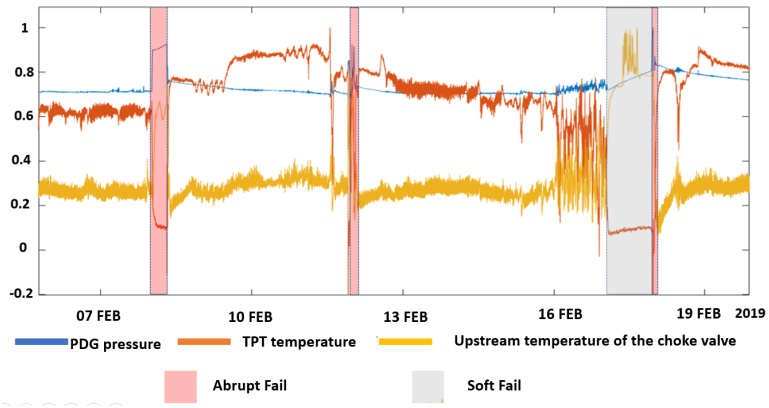

Figure 3. Faults occurred between 02/07/2019 to 02/19/2019

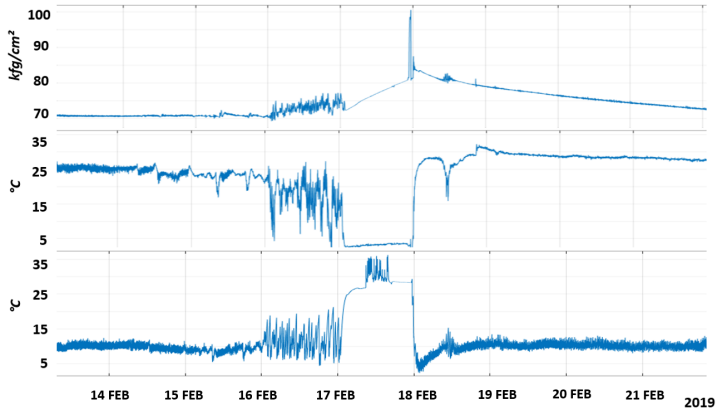

Figure 4. Behavior of PT1 Pressure, TT2 Temperature and TT3 Temperature, respectively between 02/14/2019 to 02/21/2019

The sensor values are normalized in the figures to facilitate the visualization of the graphics and the classes were defined as Failure and No Fail.

Labeling was proposed by the authors, according to information provided by offshore oil well production operators, both field operators and control room operators (supervisory). Through experiences throughout their careers in the oil and gas field. The failures were found in the production monitoring supervisory. In the Failure class there are two patterns of occurrence, Abrupt Failure and Soft Failure. The number of observations for the Non-Fail class is 126577 and 3015 for the Failure class, with Abrupt Failure and Soft Failure containing 2143 and 872 observations respectively. For the training and testing rounds, they were separated according to Table 1, taking the smallest amount of data as a parameter. Information about classes is described as follows:

- No Failure: in this class there is no variation that causes damage to the process or production, the operation of the wells behaves normally, within a known stability;
- Abrupt Failure: this type of failure results from a quick event, which may be the actuation of a final element, such as valves or process actuators. This type of failure can generate an alert in the platform supervisory, it is quickly observed by the process operators, acting abruptly on the PT1 variable, PDG pressure, for example;
- Soft Failure: the occurrence of this type of failure is not noticeable by the process supervisory operators in an agile way, the PT1 pressure rise happens slowly until an Abrupt Failure occurs.

Table 1. Division of classes in data windows

Data	Fail (Abrupt Fail + Soft Fail)	No Fail
Quantity	872+872	1744

3.2 Construction of Detectors

In the construction of the fault detectors (or models) two stacked autoencoders are used [8], the first has a hidden layer with 9 neurons and the second with 5 neurons. The training of the first autoencoder is performed with the 16 variables contained in [11] as input and the training of the second autoencoder uses the output of the hidden layer of the first autoencoder as input. The hidden layer output of the second autoencoder is used as input for the training of detectors, that is, the autoencoders are used as data pre-processing and dimensionality reduction from 16 to 5 input variables. The authors previously defined a reduction of characteristics of approximately 70% of the initial amount of input variables, with the intention of reducing the computational complexity of the training of detectors.

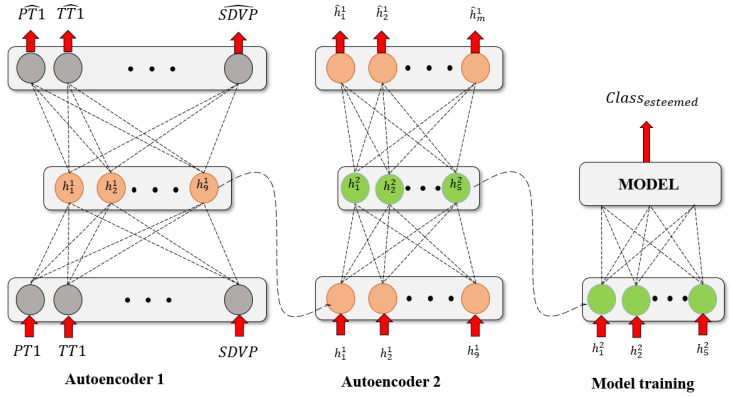

Figure 5. Process of building a stacked autoencoders model for fault classification in production wells with gas lift lift

Figure 5 shows the construction process of the proposed model, where $v = \{PT1, TT1, ..., SDVP\}$ represents the model's input variables . $\hat{v} = \{\hat{v_1}, \hat{v_2}, ...\hat{v_{11}}\}$ are the variables input values estimated in the output of the *autoencoders*, $h^1 = \{h_1^1, h_2^1, ...h_9^1\}$ are the values of the vector h^1 in the hidden layer of the first *autoencoder* and $\hat{h}_1 = \{\hat{h}_1^1, \hat{h}_2^1, ...\hat{h}_9^1\}$ the estimated output of h^1 in the second *autoencoder*. The values of the hidden layer of the second *autoencoder* are $h^2 = \{h_1^2, h_2^2, ...h_5^2\}$ which are the input parameters of the detectors, which outputs the estimated class. Models without dimensionality reduction are also created, using the 16 available input characteristics, in order to compare the computational cost and accuracy of the proposed technique.

Four models commonly used in pattern recognition are developed in this work: Decision Tree [15], Linear Discriminant Analysis [16], Support Vector Machine [17] and K nearest neighbors [18]. The recall, precision and f1-score [19] metrics are used to analyze the performance of the classifiers.

For the development of the proposed models, classifiers available in the Classification Learner tool, a Matlab® software application, are used. For the creation of the models, the hyperparameters were adjusted in the tool itself, being displayed in the Table 2.

Table 2. Hyperparameters of the developed models

Models	Hyperparameters
Decision Tree	Maximum number of divisions: 25
	Impurity Criteria: Gini Index
Discriminant	Covariance Structure: Total
SVM	Kernel function: Gaussian
	Kernel Scale: 0.56
	Multiclass Method: *One-vs-One*
k-NN	Number of Neighbors: 7
	Distance Metric: Euclidean
	Distance Weights: Equal

4 Results and Discussions

Most models satisfactorily classified the set of observations. The metrics recall, precision and f1-score of the developed models, obtained from the test data, can be observed in Table 3 . This table presents the results of the SVM and k-NN models with all 16 inputs, that is, without implementing the autoencoders, as these classifiers obtained a superior performance in the f1-score metric.

Table 3. Values of the metrics for the test data

Models	recall	precission	f1-score
Decision Tree	0.8233	0.7524	0.7863
Discriminant	0.8767	0.7412	0.8033
SVM	0.8234	0.8104	0.8297
k-NN	0.8946	0.8526	0.8731
SVM without reduction	0.8896	0.8752	0.8823
k-NN without reduction	0.9114	0.9049	0.9081

The k-NN models obtained, in general, better performances than the other techniques in the calculated metrics. Although the k-NN model without the application of autoencoders has achieved better metrics compared to its version with dimensionality reduction, it is observed that in the recall item this difference in performance is even smaller. For the problem of this work, this technique is more important because the presence of a false negative is more problematic. The cost of a false negative is usually higher, and these costs being different, an abnormal event classified as normal is more harmful than a normal event classified as abnormal.

In other words, in addition to the fact that the use of autoencoders does not greatly affect performance indices, the dimensionality reduction provides a reduction in the training time of the models. When using the k-NN technique, this factor is in the order of approximately eight times. The training rate is 12,000 observations per second for the scaling model and 1,600 observations per second without scaling.

The models trained with the SVM technique, when compared to each other, are within six percentage points of recall difference. Demonstrating again the effectiveness of autoencoders.

The classifiers based on decision tree and discriminant achieved good results. These classifiers obtained lower results for the F1-score metric when compared to SVM and k-NN.

5 Conclusions

The failure detection models presented in this work satisfactorily classified the observations. In particular, the k-NN model achieved the best results, especially in the recall metric, which is considered the most important for this type of problem.

The cascade autoencoder network used together with other techniques for the classification of failures in oil wells with artificial elevation by gas lift, has a great applicability in reducing the dimensionality of the data. The

autoencoders made it possible to reduce the training time, keeping the performance of the models close when compared to models without its use. This shows that their use can be viable in the petrochemical industry, where a fast response time to abnormalities in the production process monitoring systems is needed. A more agile response tends to reduce the complexity of direct actions to return to normality in the process plants.

For future work, it is recommended the application of other failure detection models, such as other techniques for reducing data characteristics, verifying their performance compared to those developed. The use of the models in other production wells with elevation by gas lift will allow the verification of generalization and understanding of the results obtained, in addition to simulations and studies of applicability in real-time systems.

References

[1] M. T. Schiavi and W. A. M. Hoffmann. Cenário petrolífero: sua evolução, principais produtores e tecnologias. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, vol. 13, n. 2, pp. 259–278, 2015.

[2] J. Thomas. *Fundamentos de engenharia de petróleo*. Interciência, 2004.

[3] H. S. R. Filho. A otimização de gás lift na produção de petróleo: Avaliação da curva de performance do poço. Master's thesis, Universidade Federal do Rio de Janeiro, Rio de Janeiro - RJ, 2011.

[4] R. E. V. Vargas. *Base de Dados e Benchmarks para Prognóstico de Anomalias em Sistemas de Elevação de Petróleo*. PhD thesis, Universidade Federal do Espírito Santo, Vitória - ES, 2019.

[5] Xuewu Dai, Guangyuan Liu, and Zhengji Long. Discrete-time robust fault detection observer design: A genetic algorithm approach. In *2008 7th World Congress on Intelligent Control and Automation*, pp. 2843–2848, 2008.

[6] M. Araujo, J. Aguilar, and H. Aponte. Fault detection system in gas lift well based on artificial immune system. In *Proceedings of the International Joint Conference on Neural Networks, 2003.*, volume 3, pp. 1673–1677 vol.3, 2003.

[7] I. H. Santos, H. F. Lisboa, de T. S. Feital, M. M. Câmara, R. M. Soares, M. A. Marins, B. D. Barros, de T. M. Prego, de A. A. Lima, and S. L. Netto. Hydrate failure detection in production and injection lines using model and data-driven approaches. In *Rio Oil Gas Expo and Conference 2018*, Rio de Janeiro - RJ, 2018.

[8] S. Abdellatif, C. Aissa, A. A. Hamou, S. Chawki, and B. S. Oussama. A deep learning based on sparse auto-encoder with mcsa for broken rotor bar fault detection and diagnosis. In *2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)*, pp. 1–6, 2018.

[9] P. Wen, T. Wang, B. Xin, T. Tang, and Y. Wang. Blade imbalanced fault diagnosis for marine current turbine based on sparse autoencoder and softmax regression. In *2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp. 246–251, 2018.

[10] R. Nilchiani, C. M. Edwards, and A. Ganguly. Introducing a tipping point measure in explaining disruptive technology. In *2019 International Symposium on Systems Engineering (ISSE)*, pp. 1–5, 2019.

[11] R. S. F. Nascimento, R. E. V. Vargas, B. H. Groenner, and dos I. H. F. Santos. Detecção de falhas com stacked autoencoders e técnicas de reconhecimento de padrões em poços de petróleo operados por gas lift. In *Congresso Brasileiro de Automática (CBA)*, volume 2, 2020.

[12] L. A. Aguirre, B. O. Teixeira, B. H. Barbosa, A. F. Teixeira, M. C. Campos, and E. M. Mendes. Development of soft sensors for permanent downhole gauges in deepwater oil wells. *Control Engineering Practice*, vol. 65, pp. 83 – 99, 2017.

[13] C. Lu, Z.-Y. Wang, W.-L. Qin, and J. Ma. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. *Signal Processing*, vol. 130, 2016.

[14] L. Wen, L. Gao, and X. Li. A new deep transfer learning based on sparse auto-encoder for fault diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, n. 1, pp. 136–144, 2019.

[15] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. Classification and regression trees. belmont, ca: Wadsworth. *International Group*, vol. 432, pp. 151–166, 1984.

[16] G. McLachlan. *Discriminant analysis and statistical pattern recognition*. Wiley Series in Probability and Statistics. Wiley, 1992.

[17] C. Cortes and V. Vapnik. Support-vector networks. In *Machine Learning*, pp. 273–297, 1995.

[18] N. Altman. An introduction to kernel and nearest-neighbor nonparametric regression, 1992.

[19] C. Goutte and E. Gaussier. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *Proceedings of the 27th European Conference on Advances in Information Retrieval Research*, ECIR'05, pp. 345–359, Berlin, Heidelberg. Springer-Verlag, 2005.