Data-driven Detection and Identification of Undesirable Events in Subsea Oil Wells

Chrisander Brønstad
Dept. of Science and Industry Systems,
University of South-Eastern Norway
Kongsberg, Norway
email: c.bronstad94@gmail.com

Sergio L. Netto
Dept. of Electrical Engineering Program,
Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
email: sergioln@smt.ufrj.br

Antonio L. L. Ramos
Dept. of Science and Industry Systems,
University of South-Eastern Norway
Kongsberg, Norway
email: antonior@usn.no

***Abstract*—Condition-Based Monitoring (CBM) systems have grown in popularity in recent years owing to innovations in areas, such as sensor-technology, communication systems, and computing. That has fostered the development of more efficient systems to monitor, analyze, and identify failures in industrial plants, production lines, and machinery. Gas and oil industries lose billions of dollars yearly related to abnormal events and systems failures. Thus, Abnormal Event Management (AEM), which aims at early detection and identification of these events, has become their number one priority so that preventive actions can be taken timely. This work addresses the issue of detection and classification of faults in offshore oil wells. The aim is to create a CBM system based on the random forest classifier to support decision-making. The events used in this work are part of the 3W database developed by Petrobras, Brazil, one of the world's largest oil producer. Seven events categorized as faulty events are considered, as well as several instances labeled as normal operation. We conducted two experiments related to two different classification scenarios. The proposed systems achieved an overall accuracy of 90%, indicating that the system is not only able to detect faulty events but can also anticipate incoming failures successfully.**

Keywords - condition-based monitoring; machine learning; data-driven detection and classification; random forest classifier.

I. Introduction

Recent advances in sensor technologies, communication systems for data acquisition, storage, and computational capability gave birth to an era of massive automatic data gathering, storage, and processing. This has resulted in a paradigm shift bringing new opportunities for developing innovative solutions and systems for a wide range of applications. As an example of that, Germany launched recently a project known as "Industry 4.0" to revitalize the industry based on such systems, commonly referred to as smart technology. Today, Condition-Based Monitoring (CBM) applies state-of-art technology and is often related to solutions such as Cyber-Physical Systems (CPS) and Internet of Things (IoT) [1], [2].

Systems for CBM are being widely adopted to monitor and evaluate the condition of processes, machinery, and components of interest. The goal is to anticipate and detect incoming failures so that preventive actions can be taken minimize downtime and guarantee a stable production. Such solutions often involve data-driven analysis, where acoustic and vibration signatures, current, and temperature are examples of features that are monitored to evaluate the condition of bearings, motors, and other machinery [3]. In the oil and gas industry, keeping a stable production is particularly important. This is because undesirable abnormal events can cause production losses for days and even weeks, not to mention potential disasters with catastrophic consequences for the environment. It is estimated that the oil and gas industries lose 20 billion dollars every year due to abnormal events. Thus, they have rated abnormal event management (AEM) as their number one problem that needs to be addressed. Similarly to CBM systems, AEM addresses fault detection and diagnosis, and has as main objective timely detection, diagnose, and correction of abnormal conditions or faults in a process [4].

Researchers and engineers have been studying and proposing the application of detection and classification algorithms in the oil and gas industry at the different stages, from as early as drilling and construction stages to production and operation phases of the oil well and its subsystems. An example is the work of Ahmadi et al. [5] that investigates the issue of early detection of flow influx during drilling. The authors present an approach to determining underlying reservoir models from noisy pressure data with the use of Random Forest (RF), Support Vector Machine (SVM), Linear Regression (LR), and Probabilistic Neural Networks (PNN) as classifiers for well-testing model classification. Another example is the work of Tang and S. Zhang, F. Zhang, and Venugopal [6], that presents a method of applying statistical features on real-time drilling data to automatically detect flow influx during drilling. The authors report a reliable performance and claim to be able to predict undesirable flow influx trends on average 10 minutes before detection.

Other examples of detection algorithms in later stages include the work of Liu et al. [7] which discusses an approach for semi-supervised classification to detect failures in artificial lift systems. Artificial lift systems are techniques to enhance oil production by increasing the pressure within the reservoir, which directly lifts fluids to the surface. The authors present a framework that combines features with Decision Trees (DT), SVM, and Bayes Net to enable learning and separation of failures from normal patterns based on noisy and poorly labeled multivariate time series. In addition, Liu, Li, and Xu [8] present an integrated model for the detection and location of leakages in pipelines. The authors investigate two modules:

 ISBN: 978-1-61208-918-8

one that can detect larger leakages and another one for micro-leakages. In [9], the authors present an approach to predict valve failures in gas compressors from oil fields, with the use of sensor data from multiple sensors. The authors' approach consisted of the use of feature extraction and selection, combined with DT. In [10], the authors present data-driven models to predict failure rates and their influencing factors of equipment based on data from six Norwegian oil and gas facilities.

This work presents a proper CBM system using random forest classifier to identify and detect undesirable events in subsea oil wells. The aim is to create a CBM system based on the random forest classifier to support decision-making. Specifically, we propose two classification scenarios that have shown promising results under testing. The rest of this paper is organized as follows. Section II introduces background knowledge about sub-sea oil wells and the eight different types of faults characterized as undesirable abnormal events in oil wells. Section III presents the test results related to model performance. Section IV and Section V provide an in-depth analysis, summarize the achieved results of the system and propose ideas for future work.

II. Data Analysis

This section provides an overview of offshore oil wells, followed by relevant details regarding the sensors used in this work to detect undesirable events during oil production and a description of the 3W dataset collected for the purpose of this study. It also gives a general description of the eight fault types contemplated in the 3W dataset.

A. Offshore Oil Wells

An oil well is a boring in the Earth, inshore or offshore, build using traditional drilling, and designed with the finality of extracting petroleum oil hydrocarbons from underground reservoirs. Usually, associated petroleum gas is also released in the process. The term "oil well" usually refers to a complex system consisting of several subsystems: a production tubing, which is the main path for the well fluid; a wellhead to ensure structural safety during drilling and production; and a "Christmas tree" installed on the top of the wellhead giving access to the production tubing. The 'Christmas tree" controls the production with several valves and sensors that can be accessed from the surface. Figure 1 illustrates a typical offshore oil well set-up.

The communication link between the surface and the oil well on the seabed is referred to as an "umbilical". An "umbilical" is an electro-hydraulic unit responsible for transmission of electrical signals and hydraulic power. It is connected to the Christmas tree and to the surface control system, i.e., a nearby production platform [12].

B. 3W Dataset

The 3W dataset is a public dataset released by Petrobras, the Brazilian state oil company [11]. The dataset consists of real, simulated, and hand-drawn data of oil wells sensor data

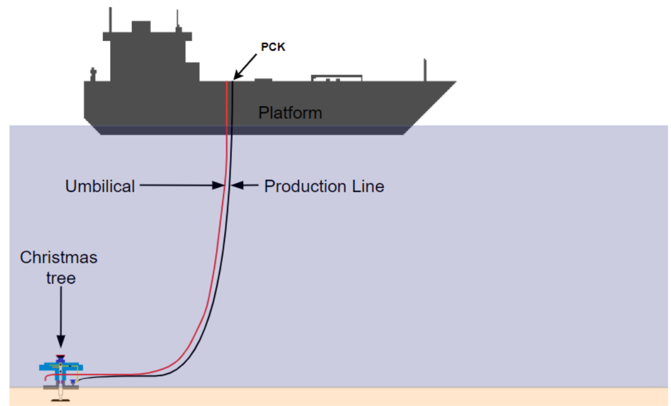

Fig. 1: Simplified schematic of a typical offshore naturally flowing well based on [11].

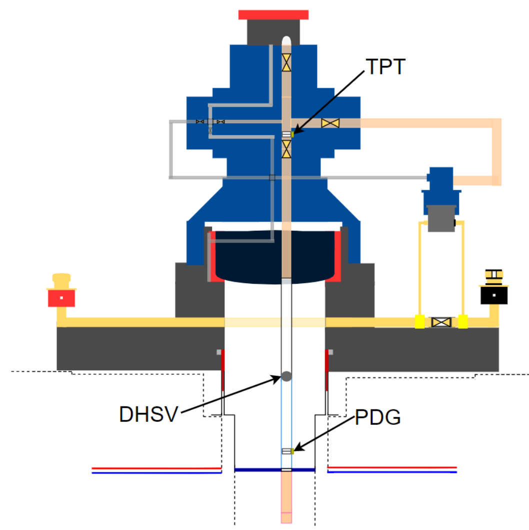

Fig. 2: Simplified schematic of a typical subsea Christmas tree based on [11].

gathered during operation. The data is acquired during oil well normal operation and featuring occurrences of undesired or abnormal events in the oil. This is shown through the closure mechanism of the Downhole Safety Valve (DHSV) and sensor readings extracted from five monitored variables: Pressure at the Permanent Downhole Gauge (PDG); Pressure at the Temperature and Pressure Transducer (TPT); Temperature at the TPT; Pressure upstream of the Production Choke Valve (PCK); and the Temperature downstream of the PCK. The PCK and DHSV and their impact are explained further on. Figure 1 illustrates Simplified schematic of a typical offshore oil well, whereas Figure 2 shows a simplified schematic of a typical subsea Christmastree [11].

The 3W dataset was build targeting eight types of undesirable events in oil wells. Factor such as Water, sediment, natural gas, and flow rate are found to be correlated to the undesirable events under consideration. As mentioned earlier, there are real, simulated, and hand-drawn undesirable events in

 ISBN: 978-1-61208-918-8

the dataset, where all real instances have been extracted from the plant information system of Petrobras. Every undesirable event in the dataset is a sequence of observations with three states, namely *normal*, *faulty transient*, and *faulty steady state*. A normal state is characterized by the absence of any evidence of abnormal behavior, whereas in the faulty transient state the dynamics caused by undesirable events are ongoing. When these dynamics cease, the faulty steady state period begins. These states are defined to allow early detection of a given failure event. Properties considered when dining the events are pressure in Pascal $[\mathrm{Pa}]$, volume flow in standard cubic meters per second $[\mathrm{sm}^3/\mathrm{s}]$, and temperature in degree Celsius $[^\circ\mathrm{C}]$.

C. Fault Description

The reference in [11] defines eight types of fault as follows:

Class 1 - Abrupt Increase of Basic Sediment & Water: Basic Sediment and Water (BSW) is defined as the ratio between the water and sediment flow rate and the liquid flow rate, both measured under Normal Temperature and Pressure (NTP). During the life cycle of a well, its BSW is expected to increase due to increased water production. However, a sudden increase of BSW can lead to several problems related to flow assurance, lower oil production, and incrustation.

Class 2 - Spurious Closure of DHSV: The Downhole Safety Valve (DHSV) is placed in the production tubing, where its purpose is to ensure the closing of the oil well. It provides safety by shutting off the well in situations in which the production unit and well are physically disconnected or in the event of an emergency or catastrophic failure of surface equipment. However, the closing mechanism will eventually fail in a spurious manner. This kind of failure is problematic because there are often no indications of the failure on the surface, which causes production losses and additional cost.

Class 3 - Severe Slugging: This type of undesirable event occurs frequently at irregular intervals, on mature oil fields. Severe slugging takes place when “slugs” of liquid separate bubbles of gas through the pipeline. In the 3W dataset, it is considered a critical type of instability and can result in stress or even damage to equipment in the well and/or the industrial plant.

Class 4 - Flow Instability: During flow instability, there is a periodical change of pressure but with acceptable amplitudes. Flow instability is not necessarily equal to slugging, what separates those two anomalies is the lack of periodicity. Though flow instability can result in slugging. As instability can progress to severe slugging, its prognosis avoids all the negative aspects associated with this more severe anomaly.

Class 5 - Rapid Productivity Loss: There are several factors that can change the productivity of a naturally flowing well, the factors consist of the diameter of the production line, percentage between water and basic sediment, static pressure of the reservoir, and the viscosity of the produced fluid. When any of these factors are changed to the extent that the system’s energy is not sufficient enough to overcome the losses, the flow of the well will slow down or even stop, which causes productivity loss.

Class 6 - Quick Restriction in PCK: Production Choke (PCK) is a control valve located at the beginning of the production unit. It is responsible for well control and can restrict, control, and regulate the flow. The choke can be controlled from the surface and when operated manually problems may occur.

Class 7 - Scaling in PCK: Inorganic deposits will occur during production. Therefore, it is important to monitor the Production Choke since it significantly reduces oil and gas production. If detected, losses of oil and gas production can be avoided. Thus, detecting it a early stage is favorable, so actions can be taken.

Class 8 - Hydrate in Production Line: This undesirable event occurs when water and natural gas form a crystalline compound, which happens under extreme pressure and temperature conditions. This crystalline compound resembles ice and when it is formed in production lines it can stop production for days and weeks. This is one of the biggest problems in the oil industry. Thus, avoiding this is desirable.

D. Data Review and Challenges

Despite its great technical value, the 3W dataset includes many of missing and frozen variables and unlabeled observations. In this case, a ’variable’ refers to the monitored operational settings and sensor readings. Furthermore, an ’instance’ refers to a recorded event of one of the eight fault types in the 3W dataset, while an ’observation’ is a sample from an instance, showing the true label, timestamp, operational settings, and sensor readings. These definitions are used in the following subsections, which review challenges related to the 3W dataset.

1) Unlabeled Observations: An observation is considered unlabeled when there is no label of the fault type for a given sample of an instance. A total of 5,130 (0.01% of all 50,913,215 observations of all 15,872 variables of all 1,984 instances) observations are considered unlabeled in the 3W dataset.

2) Missing and Frozen Variables: A variable is considered missing when all observations of that particular variable in an instance have a missing value. 4,947 (31.17% of all 15,872 variables of all 1,984 instances) variables are considered missing in the 3W dataset. In the case of frozen variables, they are considered frozen when all observations of that particular variable in an instance have the same constant value. 1,535 (9.67% of all 15,872 variables of all 1,984 instances). variables in the 3W dataset is considered frozen.

3) Hand-drawn Instances: A challenge of this dataset is related to hand-drawn instances because the behavior varies a lot compared to real instances. The hand-drawn instances are too artificial and are quite distinct from the real ones. Therefore, the subsequent analysis omits fault type seven (Scaling in PCK) since 10 of all 14 instances are hand-drawn.

III. Experiments and Results

This section presents a comprehensive analysis of system performance. The analysis has been conducted by approaching this as a binary classification problem under two classification scenarios. In this work, these classification scenarios are characterized as *Fault versus Normal Operation* and *Fault versus Not Fault*.

In the scenario of fault versus normal operation, each binary classifier is designed to discriminate and identify between normal operation and a single specific class, categorized as a fault event. Thus, seven individual uncorrelated classifiers are required (one for each event representing an anomaly). Furthermore, all samples from normal operation (event/class 0) are combined to a single unique class, with the samples that show normal dynamic behavior preceding the given fault class (initial normal). For the classification scenario of fault versus not fault, each binary classifier discriminates and identifies between a single specific fault against everything that does not belong to that specific fault. This is done by designing seven uncorrelated classifiers, one for each fault, and combining the remaining events into a single unique class.

The experiments above have been conducted to achieve a greater understanding of the system and to acquire more information about each fault class. Besides, this continues and contributes to the work of Marins et al. [13], which have shown that a CBM system can be used to classify the faults, with binary and multiclass classifiers. This work introduces a new classification scenario (fault versus not fault). The samples that belong to the faulty transient and faulty steady states for each fault are combined to one unique class in the two classification scenarios described above. In this work, the model performance is assessed in detail: firstly, by measuring the overall model performance; secondly, by measuring the capability to discriminate between normal operation and each transitional state; and lastly, by individually measuring the simulated and real instances of the latter two metrics.

A. Experiment 1: Fault versus Normal Operation

This section reviews the experiment characterized as *fault versus normal operation*. In this experiment, each classifier is fit on data from normal operation (class 0) and their respective fault. The only change in any hyperparameter relates to the subsampling factor. In this case, fault events with less than 20 real instances apply a subsampling factor of 1 for real instances. This is with the intent to balance the distribution ratio, between simulated and real instances. Moreover, the window sizes related to the feature extraction of each classifier consist of 900, 400, 1200 samples for classes 1, 2, and 8, respectively; and, 300 samples for the remaining classes. These specific window sizes represent the best performance for their given classifier.

The test results of the 3W dataset for the transitional states and overall accuracy can be seen in Table I. In this table, 'Transitional ACC' denotes the overall accuracy for the transitional states, i.e., initial normal (normal operation preceding an anomaly event), faulty transient state, and faulty steady-state. Additionally, 'Overall ACC' represents the accuracy for all transitional states combined with samples from class 0. The reason for showing the overall and transitional accuracy separately is because approximately 30% of all instances from the complete 3W dataset belongs to Class 0. Thus, the overall accuracy is not sufficient alone and may misrepresent the performance of the classifier. More details can be seen in Table II, which shows real and simulated test results for class 0 and each transitional state.

TABLE I: TEST RESULTS OF CLASSIFICATION SCENARIO FAULT VERSUS NORMAL OPERATION.

Fault	Window Size	Type	Transitional ACC	Overall ACC
Class 1	900	Real	0.214	0.989
		Simulated	0.999	0.999
Class 2	400	Real	0.986	0.999
		Simulated	0.996	0.996
Class 3	300	Real	0.998	0.999
		Simulated	0.998	0.998
Class 4	300	Real	0.967	0.986
		Simulated	-	-
Class 5	300	Real	0.888	0.992
		Simulated	0.993	0.993
Class 6	300	Real	0.972	0.999
		Simulated	0.951	0.951
Class 8	1200	Real	0.892	0.999
		Simulated	0.956	0.956

TABLE II: TEST RESULTS OF CLASSIFICATION SCENARIO FAULT VERSUS NORMAL OPERATION.

Fault	Window Size	Type	Normal (Class 0)	Initial Normal	Transient State	Steady State
Class 1	900	Real	1.000	0.779	0.098	0.011
		Simulated	-	0.999	0.999	0.999
Class 2	400	Real	0.999	0.952	0.996	1.000
		Simulated	-	1.000	0.970	1.000
Class 3	300	Real	0.999	-	-	0.998
		Simulated	-	-	-	0.998
Class 4	300	Real	0.990	-	-	0.967
		Simulated	-	-	-	-
Class 5	300	Real	0.999	0.514	0.904	-
		Simulated	-	0.134	0.999	1.000
Class 6	300	Real	0.999	0.981	0.882	1.000
		Simulated	-	0.876	0.955	0.956
Class 8	1200	Real	0.999	0.000	1.000	1.000
		Simulated	-	0.258	0.996	0.999

The imbalance of 3W dataset reflects on the performance of each classifier. This is evident in the case of classes 1 and 8, where both classifiers seem to struggle with real data. These two classes have less than 10 real instances combined, in contrast to classes 2, 3, and 4, which have 22, 34, and 344 instances, respectively. One can also observe that class 5 has difficulties correctly classifying samples of the initial state. This applies to the case of both real and simulated instances. On the other hand, the classifier achieves satisfactory accuracy in correctly classifying samples belonging to normal operation and the transient state. In this case, the results show an accuracy of over 90% for the latter two events.

1) System Efficiency and Reliability Evaluation: This section reviews the system in a real-world scenario as a CBM system. To be a reliable CBM system, the system must have the capability to detect any event as soon as possible. To assess how reliable and efficient the system is to anticipate incoming failures, three time-intervals have been applied. These time intervals (in seconds) are defined as following: how fast each classifier is to detect the transient state (time of detection); how many consecutive correct predictions are made after the time of detection; and how long time there is to take action before the incoming failure occurs, respectively denoted as t_1, t_2, and t_3. Table III shows the average values of the time-intervals t_1, t_2, and t_3 for each classifier and their respective faults, where each number designated in parenthesis is the percentage of the corresponding time-interval concerning the transient state. Classes 0, 3, and 4 are not included as their transient phase is absent in the 3W dataset.

TABLE III: EFFICIENCY AND RELIABILITY ANALYSIS OF FAULT VERSUS NORMAL OPERATION.

Fault	Type	t1 [s]	t2 [s]	t3 [s]
Class 1	Real	2463.0 (11.73%)	2710.0 (12.90%)	18530.0 (88.26%)
Class 2	Real	15.5 (0.33%)	4700.1 (99.67%)	4700.1 (99.67%)
Class 5	Real	564.2 (1.06%)	41879.0 (78.87%)	52533.7 (98.93%)
Class 6	Real	73.5 (11.87%)	546.0 (88.20%)	546.0 (88.20%)
Class 8	Real	1.0 (0.00%)	20078.0 (100%)	20078.0 (100%)

B. Experiment 2: Fault versus Not Fault

In the the classification scenario fault versus not fault, each classifier is fit on data from all classes. The same settings from the latter classification scenario apply in this case, with regards to hyperparameter selection and training routine. As for the training routine, the subsampling factor had to be increased, since each classifier is fit on data from every class. Each classifier under training applied a subsampling factor of 100 for instances that did not belong to their given class (Not Fault). Also, real instances belonging to their given class were not subsampled, but simulated instances applied a subsampling factor of 10.

The test results of this classification scenario of the 3W dataset can be seen in Tables IV and V, where empty entries indicate the absence of data for that given fault type. The classification method shows satisfactory results with the classifiers of classes 2, 3, and 4. These classes are correctly classified with an average accuracy of over 90% of Class 0 and all transitional states when it comes to real instances. Simulated instances of these classes indicate to be harder to predict, where the accuracy drops as low as 80% for the transient state and 93% for the steady-state for classes 2 and 3, respectively.

1) System Efficiency and Reliability Evaluation: The assessment of each classifier based on the time-intervals t_1, t_2, and t_3 can be seen in Table VI, where the time-intervals are given in seconds and the designated numbers in parenthesis are the percentage of the corresponding time-interval concerning the total transient state. Besides Class 5, the system shows similar results when compared to Experiment 1. In particular, Class 5 performs poorly in the transitional states t_1 and t_2 due to the appearance of inconsistent classification behavior.

IV. Discussion

The results achieved from Experiment 1 and Experiment 2 show that the system is capable of correctly classifying faults but also capable of predicting incoming faults from their transient state. These incoming faults are often predicted in an early stage, such that it is possible to take necessary preventive action. The amount of data at disposal for each class tends to reflect the system performance for the given classifier. This applies to both classification methods, and in particular, to real instances of classes 1 and 8. Neither of the two classification methods can correctly classify any sample related to the initial normal state of Class 8. Classes 2, 3, and 4 achieves great accuracy for their respective transitional states and Class 0, for both classification scenarios. On the other hand, the fault versus normal classification scenario accomplishes better results for classes 5 and 6 compared to its counterpart, fault versus not fault.

Comparing these results with the work of Marins et al. [13], it is noticeable that their binary classification method achieved better results when classifying the initial normal state for each class on real instances. However, when comparing their multiclass classification method, the method proposed in this work for each class exhibits higher accuracy in average on real instances over all transitional states for most classes. This can be seen in the results of the classification scenario 'fault versus normal'.

TABLE IV: TEST RESULTS OF CLASSIFICATION SCENARIO FAULT VERSUS NOT FAULT.

Fault	Window Size	Type	Transitional ACC	Overall ACC
Class 1	900	Real	0.213	0.992
		Simulated	0.999	0.999
Class 2	400	Real	0.991	0.999
		Simulated	0.862	0.997
Class 3	300	Real	0.995	0.980
		Simulated	0.936	0.990
Class 4	300	Real	0.977	0.985
		Simulated	-	1.000
Class 5	300	Real	0.373	0.966
		Simulated	0.993	0.996
Class 6	300	Real	0.868	0.999
		Simulated	0.925	0.987
Class 8	1200	Real	0.892	0.999
		Simulated	0.972	0.982

TABLE V: TEST RESULTS OF CLASSIFICATION SCENARIO FAULT VERSUS NOT FAULT.

Fault	Window Size	Type	Not Fault (Class 0)	Initial Normal	Transient State	Steady State
Class 1	900	Real	1.000	0.779	0.098	0.001
		Simulated	0.999	0.999	0.999	0.999
Class 2	400	Real	0.999	0.970	0.998	1.000
		Simulated	0.999	1.000	0.821	0.848
Class 3	300	Real	0.979	-	-	0.995
		Simulated	0.998	-	-	0.936
Class 4	300	Real	0.987	-	-	0.977
		Simulated	1.000	-	-	-
Class 5	300	Real	0.999	0.396	0.372	-
		Simulated	0.997	0.103	0.999	0.999
Class 6	300	Real	1.000	0.997	0.384	0.150
		Simulated	0.999	0.813	0.936	0.930
Class 8	1200	Real	0.999	0.000	1.000	1.000
		Simulated	0.983	0.939	0.961	0.999

TABLE VI: EFFICIENCY AND RELIABILITY ANALYSIS OF FAULT VERSUS NOT FAULT.

Fault	Type	t1 [s]	t2 [s]	t3 [s]
Class 1	Real	2463.0 (11.73%)	2707.0 (12.90%)	18530.0 (88.26%)
Class 2	Real	9.3 (0.19%)	4706.4 (99.81%)	4706.4 (99.81%)
Class 5	Real	1.0 (0.00%)	822.3 (1.55%)	53097.3 (99.9%)
Class 6	Real	318.5 (51.5%)	237.5 (38.4%)	300.5(48.54%)
Class 8	Real	1.0 (0.00%)	20078.0 (100%)	20078.0 (100%)

A. Inconsistency

A limitations of this system is tied to the observed fluctuations in the classifications of Classes 4 and 5. To mitigate that, we suggest the use of a simple filter to smoothen these fluctuations that occur during inconsistent classifications, referred to as “time-consistency filter”. The filter strides over the system classifications using a window and removes the class with the fewest output classifications in that window. Figure 3 shows an example of a real instance of Class 4, along with the system classifications and the outputs of the time-consistency filter. Here, the purple marker shows the output of the filter. In this case, the window size of the time-consistency filter is 120 samples. The filter is unable to prevent all classification oscillations, but it does smoothen out the majority of them. The filtering process may be more efficient if the window size is increased, but this also increases the delay, which is undesirable.

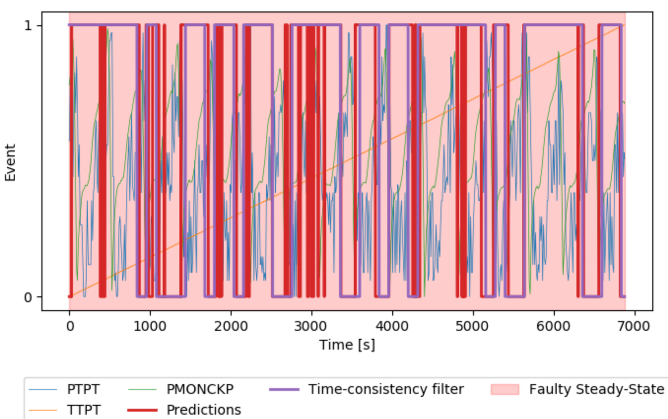

Fig. 3: Example of a real instance of Class 4 along with the inconsistent system classifications and with the time-consistency filter. Event values ’1’ and ’0’ denote normal and faulty states.

V. Conclusion and Future Work

This work develops a CBM system to detect and identify real abnormal events in offshore oil wells. The CBM system is tested in two different classification scenarios. Either case includes preprocessing of raw-sensor data, feature extraction, dimensionality reduction, and classification using the random forest algorithm. However, analysis of the results indicates that inconsistent classifications may occur for Classes 4 and 5. These inconsistent classifications occur predominantly during faulty states. That means that the system would still be able to detect the fault, which is a positive feature.

As future work, we would suggest further investigation of additional features, e.g, other second-order features that have not been considered at this stage of this work. It is also interesting to explore other promising machine learning classification algorithms such as the XGBoost algorithm, which attempts to exploit the advantages of Random Forest and gradient boosting, and the Light Gradient Boosting Machine (LGBM), which has a similar architecture to the Random Forest and the XGBoosts algorithms.

Acknowledgment

This study was financed in part by the Norwegian Agency for International Cooperation and Quality Enhancement in Higher Education (Diku), project number UTF-2018-CAPES-Diku/10002, and in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES), finance code n° 23038.018065/2018-17.

References

[1] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, “Industry 4.0,” *Business & Information Systems Engineering*, vol. 6, pp. 239–242, 08 2014.

[2] J. Lee, B. Bagheri, and H.-A. Kao, “A cyber-physical systems architecture for industry 4.0-based manufacturing systems,” *SME Manufacturing Letters*, vol. 3, 12 2014.

[3] K. K. Agrawal, G. N. Pandey, and K. Chandrasekaran, “Analysis of the condition based monitoring system for heavy industrial machineries,” in *2013 IEEE International Conference on Computational Intelligence and Computing Research*, 2013, pp. 1–4.

[4] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, “A review of process fault detection and diagnosis: Part i: Quantitative model-based methods,” *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293 – 311, 2003.

[5] R. Ahmadi, B. Aminshahidy, and J. Shahrabi, “Well-testing model identification using time-series shapelets,” *Journal of Petroleum Science and Engineering*, vol. 149, pp. 292–305, 2017.

[6] H. Tang, S. Zhang, F. Zhang, and S. Venugopal, “Time series data analysis for automatic flow influx detection during drilling,” *Journal of Petroleum Science and Engineering*, vol. 172, pp. 1103 – 1111, 2019.

[7] Y. Liu, K. Yao, S. Liu, C. S. Raghavendra, O. Balogun, and L. Olabinjo, “Semi-supervised failure prediction for oil production wells,” in *2011 IEEE 11th International Conference on Data Mining Workshops*, 2011, pp. 434–441.

[8] C. Liu, Y. Li, and M. Xu, “An integrated detection and location model for leakages in liquid pipelines,” *Journal of Petroleum Science and Engineering*, vol. 175, pp. 852 – 867, 2019.

[9] O. P. Patri, R. Nabor, A. Panangadan, and V. Prasanna, “Predicting compressor valve failures from multi-sensor data,” ser. SPE Western Regional Meeting, vol. All Days, 04 2015, sPE-174044-MS.

[10] L. Xie, S. Håbrekke, Y. Liu, and M. A. Lundteigen, “Operational data-driven prediction for failure rates of equipment in safety instrumented systems: A case study from the oil and gas industry,” *Journal of Loss Prevention in the Process Industries*, vol. 60, pp. 96 – 105, 2019.

[11] R. E. V. Vargas, C. J. Munaro, P. M. Ciarelli, A. G. Medeiros, B. G. do Amaral, D. C. Barrionuevo, J. C. D. de Araújo, J. L. Ribeiro, and L. P. Magalhães, “A realistic and public dataset with rare undesirable real events in oil wells,” *Journal of Petroleum Science and Engineering*, vol. 181, p. 106223, 2019.

[12] S. Laik, *Offshore Petroleum Drilling and Production*. Boca Raton: CRC Press, 2018.

[13] M. A. Marins, B. D. Barros, I. H. Santos, D. C. Barrionuevo, R. E. Vargas, T. de M. Prego, A. A. de Lima, M. L. de Campos, E. A. da Silva, and S. L. Netto, “Fault detection and classification in oil wells and production/service lines using random forest,” *Journal of Petroleum Science and Engineering*, p. 107879, 2020.

 ISBN: 978-1-61208-918-8