Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An Integrated Framework for Retrospective Well Operational Analysis

Daniel Centurion Barrionuevo

daniecba@stud.ntnu.no

Norwegian University of Science and Technology

Alexey Pavlov

Norwegian University of Science and Technology

Research Article

Keywords:

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11763548/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. Daniel Centurion Barrionuevo is employed by Petróleo Brasileiro S.A. (Petrobras), which funded the research and provided the operational data used in this study. Alexey Pavlov declares no competing interests.

An Integrated Framework for Retrospective Well Operational Analysis

Daniel Centurion Barrionuevo[1,2*] and Alexey Pavlov[2]

[1*]Petróleo Brasileiro S.A. (Petrobras), Rio de Janeiro, Brazil.
[2]Norwegian University of Science and Technology, Trondheim, Norway.

*Corresponding author(s). E-mail(s): dcbarrionuevo@petrobras.com;

Abstract

This work proposes a novel approach to well monitoring and investigation in which the object of analysis is a selected time interval containing any operational scenario, such as steady-state operation, well shutdown, or a complex well procedure. The proposed framework combines piecewise-linear segmentation, changepoint detection, and cause–consequence matching, aiming to identify operating regimes and detect anomalies. Detected sensor changes are matched with known operational commands using temporal, topological, and physical compatibility criteria, while unexplained changes and missing expected responses are labeled as potential abnormalities. It was evaluated through field-data case studies involving steady-state production, choke changes, well shutdown and restart procedures, and normal and abnormal ICV opening procedures. The experimental results show that this monitoring approach is feasible, can detect general, previously unknown abnormalities in different scenarios, and can serve as a starting point for broader and more powerful implementations. In addition, the extracted information can support real-time alarm systems, as illustrated by a case study involving simulated subsea pipeline rupture detection. The framework uses clear, transparent rules and a few intuitive inputs for operators, does not depend on large amounts of data, and was tested using real data from Brazilian offshore wells, which facilitates its adoption and builds trust within the industry.

Keywords: Well Monitoring, Retrospective Analysis, Changepoint Detection, Segmentation, Event Matching

List of abbreviations

BIC	Bayesian information criterion
CPOP	Continuous-Piecewise-Linear Pruned Optimal Partitioning
CROPS	Changepoints for a range of penalties
DHSV	Downhole safety valve
ESDV	Emergency shutdown valve
GLV	Gas lift valve
ICV	Interval control valve
IRFL	Iteratively reweighted fused lasso
LightGBM	Light Gradient Boosting Machine
PAB	Production adapter base
PDG	Permanent downhole gauge
PT-A	Annular pressure transmitter
PT-P	Production pressure transmitter
PXO	Pig crossover valve
RSS	Residual sum of squares
TH	Tubing hanger
TPT	Temperature and pressure transmitter
WCT	Wet Christmas tree
XO	Crossover valve

1 Introduction

Monitoring of offshore well operations is a critical topic in the petroleum industry. Flow assurance problems such as hydrates, scaling, wax, or asphaltenes can cause partial or total loss of production and demand heavy rig time to safely return to normal conditions, leading to economic losses that can range from tens to hundreds of millions of dollars (Makogon 1996; Venkatesan et al. 2005; Gonzalez et al. 2016). Equipment failure, such as valve malfunction or lack of hydraulic tightness in tubing, casing, packers, pipelines or valves, can also lead to economic losses and safety risks (Torbergsen et al. 2012; Vignes and Aadnøy 2010; Aljaroudi et al. 2015).

As such, interest in real-time well operation monitoring for the detection of abnormal events and faults has increased in recent years. In contrast to reservoir-surveillance analyses, these applications often focus on transient changes that occur or can be detected over time scales of minutes to hours. In Vargas et al. (2019), a time-series dataset entitled the 3W Dataset, containing undesirable events in oil wells, was presented. Since then, many works, such as Marins et al. (2021); Turan and Jäschke (2021); Fernandes et al. (2024); Bayazitova et al. (2024), have explored different machine learning approaches using the 3W Dataset for fault classification and anomaly detection in oil wells. Barrionuevo et al. (2026) discussed the use of dataset generation through multiphase flow simulation, combined with the 3W Dataset, to create a machine learning classifier for rare but dangerous pipeline partial-rupture events leading to oil leakage or seawater influx. Wang et al. (2025) used operational data from 100 oil wells in China equipped with mechanical pumping systems and ensemble learning techniques to detect equipment faults.

Supervised classifiers for predefined fault classes can be highly accurate and useful for fast, real-time fault detection, but they depend on the availability and quality of representative labeled data, which may be limited in many cases. Other works have therefore proposed methods that reduce the dependence on predefined fault classes. One interesting approach was proposed in Aranha et al. (2024), in which a decision diagram identifies the operating state of the well, and an LSTM autoencoder trained on normal data is used to detect additional anomalies not directly identified by the rule-based approach. Another work presented in Sudyoprasert et al. (2026) uses and compares a rule-based inference system and an LSTM model to classify the well status and subsequently to detect generic abnormalities under steady-state conditions. More recently, Lopes et al. (2025) proposed an Open-World Learning strategy that first uses the same autoencoder-based anomaly detection approach from Aranha et al. (2024) to recognize anomalous behavior. The detected anomalies are subsequently passed to one-vs-all binary classifiers assisted by one-class SVMs to determine whether they belong to any of the known anomaly classes. Unknown anomalies are then processed by clustering methods, with a Random Forest model used to estimate the number of potential new classes and clustering algorithms used to group similar anomalies. After operator validation, these patterns can be incorporated as

new anomaly classes, allowing the monitoring system to incorporate new types of events. Aranha et al. (2024) proposed an unsupervised novelty and outlier detection model for real-time offshore well monitoring, with a focus on identifying unexpected closures of downhole safety valves from pressure and temperature measurements.

Other studies have focused on identifying operational changes or extracting relevant intervals before performing event classification. In Hamedi Shokrlu and Bazile (2024), an online machine learning method is trained to detect changes in ESP operation and can therefore provide an early indication of degradation or changes in pump operating conditions. In Lopes et al. (2026), which builds on the work presented in Lopes et al. (2025), a 1D U-Net is used to segment production time series into candidate operational-event windows, which are subsequently analyzed using classifiers for known events and distance-based methods for novelty detection. Using a similar approach, Silva et al. (2026) proposed a rule-based method using numerical derivatives and signal smoothing to segment multivariate well time series into intervals containing relevant operational events, which were subsequently used to support anomaly detection and classification.

These works focus primarily on online monitoring and anomaly detection, often under steady-state or relatively stable operating conditions. However, the workflow of a monitoring engineer is not limited to continuously observing online data, a task that can often be supported by automatic fault-detection systems. Human analysis becomes particularly important when a sequence of consecutive changes across different parts of the production system must be interpreted over periods ranging from minutes to hours, using a selected group of sensor measurements and equipment states. This reasoning is based not only on the most recent measurements, but on the complete sequence of operational changes and on the overall evolution of each variable. New measurements can also help reinterpret and improve the understanding of earlier data. This type of integrated retrospective analysis has received comparatively less attention in automatic well-monitoring tools.

In this work, we propose a framework for the retrospective analysis and interpretation of petroleum production time series. Rather than continuously monitoring individual signals for predefined anomalies, the framework aims to examine how a well's operational state evolves over a selected time interval, following a workflow similar to that used by a production monitoring engineer during an operational investigation. The engineer selects a well, a time interval, and a set of relevant sensor measurements, which are then segmented into interpretable operating regimes. Changes observed across multiple measurements are associated with known operational events using temporal, topological, and physical compatibility criteria. The resulting representation provides a structured, compact description of the operational history of the well over the analyzed period, highlighting associations between operational events and transient behavior, unexplained changes, and expected but missing sensor responses as points for further investigation. The extracted information can also support other real-time fault-detection systems.

Outside the context of upstream well monitoring, other works have applied offline changepoint detection for process monitoring and operating-state identification, although in different contexts and with different purposes and approaches. Duarte and Saraiva (2003) proposed a retrospective changepoint analysis for detecting changes in process and quality variables in chemical processes. In Gedda et al. (2022), different changepoint methods were applied to segment signals from a multiphase flow facility into operating modes and generate labels in an unsupervised way. In Guo et al. (2019), nonparametric changepoint detection methods were combined with decision rules to generate alarms in automotive manufacturing processes. Hu et al. (2018) followed an integrated monitoring philosophy with similarities to the present work, describing a framework that integrates process and alarm data for process monitoring and anomaly detection, while also using causality and process connectivity for root-cause analysis. However, none of these works follow the same approach proposed in this work.

1.1 Contributions

Our contributions are as follows:

- **Conceptual**: A novel well monitoring and investigation approach in which the object of analysis is an entire selected time range containing potentially different operational scenarios, such as steady-state operation, a well shutdown, or a complex well procedure consisting of sequential steps.
- **Methodological**: An integrated framework that combines:

  - A piecewise-linear segmentation pipeline adapted to real well pressure data, consisting of a main segmentation algorithm, a resolution-aware step to remove spurious changepoints caused by negligible signal variations, a discontinuity-detection step, and a shape-aware step to avoid spurious changepoints in naturally curved time series;
  - A cause–consequence matching method that associates operational events with sensor changepoints using temporal, topological, and physical compatibility criteria, allowing unmatched changepoints and events with expected but missing responses to be identified as alarms.
- **Experimental**: Tests using data from Brazilian offshore subsea wells demonstrating the feasibility of the proposed conceptual approach and methodological framework.

2 Problem Delimitation

This section introduces the system under study by describing the main components of a typical subsea well in the Brazilian pre-salt fields. It then discusses the challenges involved in automating the retrospective analysis of operational time series, establishing the scope of the present work and defining the problem to be addressed. In order to justify the methodology presented in further sections, a detailed view of a subsea well is shown to illustrate the system complexity and its inter-dependencies, although the case studies shown in Section 6 involve only sub-regions and simpler scenarios.

2.1 System Configuration

Petroleum production systems can have very different configurations. The study presented in this work is based on a typical offshore well from the Brazilian Pre-Salt basin, connected to an FPSO (Floating Production Storage and Offloading) through a production pipeline, a service pipeline, and an umbilical. The production pipeline transports the produced fluids, while the service pipeline can be used for well interventions with diesel or other fluids, or for gas-lift injection. The umbilical connects electrical, hydraulic, and chemical lines to the well.

These connections are made through a Wet Christmas Tree (WCT) system mounted on a Production Adapter Base (PAB), where the Tubing Hanger (TH) supports the production tubing. Production fluids flow from the reservoir perforations through the production tubing to the WCT, then through the production pipeline to the FPSO.

Figure 1 shows a possible oil well configuration from the Pre-Salt Basin. This particular case has three different production zones, all isolated from each other by packers. At each zone, there is an Interval Control Valve (ICV) which can partially or totally restrain production (Joubran 2018). At some point along the tubing, there may be a gas-lift valve connecting the annulus to the production tubing. Near the top of the tubing, there is a Downhole Safety Valve (DHSV), which is a safety valve designed to close if there is a loss of communication between the platform and the well.

The Wet Christmas Tree (WCT) has many different valves (Agostini et al. 2017). For normal operation, the most important ones are:

- **Master 1 (M1)**: connects the tubing and the WCT production bore,
- **Master 2 (M2)**: connects the annulus and the WCT annular bore,
- **Crossover (XO)**: connects the production and annular bores,
- **Wing 1 (W1)**: connects the production line and the production bore,
- **Wing 2 (W2)**: connects the service line and the annular bore,
- **Pig Crossover (PXO)**: connects the production and service lines through the pig loop,
- **Chemical Injection Valves (IQ1 and IQ6)**: some WCTs have these valves to connect chemical injection lines from the umbilical into the WCT production and annular bores.
- **Swab Valves (S1 and S2)**: not shown in Figure 1, these valves can only be accessed by a drilling rig during a workover intervention.

At the topside, there are usually shutdown valves right at the arrival of the production and service lines (ESDV and ESDV-S). After that, there are choke valves which regulate the production flow rate (Production Choke) or the gas-lift flow rate (Gas-Lift Choke). There are also manual valves which connect the diesel pump to both lines, designed for diesel circulation or bullheading operations.

Along this system, there are multiple pressure and temperature sensors. For this work, the most important ones are:

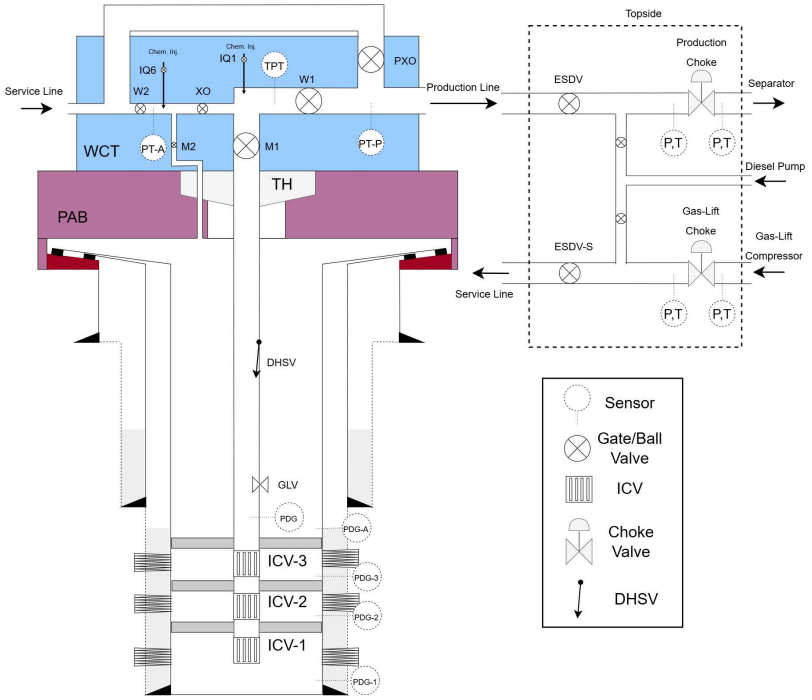

Fig. 1 Typical offshore subsea well scheme from Brazilian pre-salt fields. This particular example has three ICVs (Interval Control Valves) separated by packers, each one controlling the inflow from a reservoir zone. At the subsea bedline, a Wet Christmas Tree (WCT) serves as a manifold to connect the production tubing, well annulus, production line, service line, and chemical injection lines. This equipment has different valves—M1, W1, M2, W2, XO, PXO, IQ1, and IQ6—for different purposes. At the topside, there is also an on-off ESDV (Emergency Shutdown Valve) and a choke for flow control. This system also has different pressure and temperature sensors positioned at different points. Downhole, different PDG sensors measure the pressure at various locations—upstream of each ICV, downstream of the ICVs, or at the upper annulus. The WCT also has sensors for different chambers—TPT for the main production bore, PT-A for the annulus bore, and PT-P upstream of W1, monitoring the initial section of the production flowline.

- **PDG (Permanent Downhole Gauge)**: pressure and temperature sensor installed at the lower parts of the tubing, usually next to the reservoir to capture its dynamics. In this schematic, there are also measurements in the annulus, right above the upper packer (PDG-A), and for each zone (PDG-1, PDG-2, and PDG-3), both upstream and downstream of each ICV.
- **TPT**: pressure (P-TPT) and temperature (T-TPT) sensor located at the production bore of the WCT, between the M1, W1, and XO valves.
- **PT-P**: pressure sensor located downstream of the W1 valve.
- **PT-A**: pressure sensor located at the annular bore, between the W2, M2, and XO valves.
- **PT-CKP**: pressure and temperature sensors located upstream and downstream of the production choke valve.

There are also pressure sensors for each WCT subsea valve (M1, M2, W1, W2, XO, PXO, IQ1, IQ6). These valves are hydraulic gate valves that open when exposed to high pressure. Their operation is straightforward: when pressure is applied to the actuator, it pushes the gate, opening the flow path. When the chamber is depressurized to the seabed hydrostatic pressure, a spring returns the valve to its original position, closing the path. It is possible to infer the valve status by analyzing its activation pressure.

The ICVs also have pressure sensors, yet they operate in different ways, depending on their model. The most common models in Brazilian offshore wells have a mechanism which incrementally opens or closes the valve by applying pressure pulses in specific hydraulic lines dedicated to each ICV, while the entire group of ICVs shares a common pressure input that performs the opposite action as a one-time opening or closing command - this architecture is referred to as "N+1" by some sources (Garcia and Mathew 2016; Rodrigues et al. 2017). Therefore, for the system schematic in Figure 1, there are four hydraulic lines, of which three are dedicated to the incremental opening or closure of each ICV and the remaining one is a common input for common total closure and opening.

Although these wells are designed to remain in normal steady-state operation or to be closed, there are also special procedures that occasionally occur due to routine operations, production issues, or safety protocols. These procedures may only involve the platform operations personnel or may involve a workover by an oil rig, and they represent crucial moments for monitoring teams. A simple well restart or a planned shut-in is an example of such procedures, as they involve a sequence of operational steps. Any well that is going to start production for the first time must be commissioned, and a well that is going to be abandoned, permanently or temporarily, also needs to be decommissioned. Safety routines such as DHSV, WCT, or leak tests, as well as ICV movement tests, also demand periodic procedures. Finally, flow assurance problems—such as hydrates in production or gas-lift lines, scaling, or wax in the tubing—or any kind of subsea equipment failure can also lead to complex operational procedures.

Another important factor is that each well has a specific equipment configuration, reservoir and fluid conditions, and fault history, thus requiring specific operational procedures. The single case shown in Figure 1, for example, could differ in many ways: a different number and type of ICVs, the type of DHSV, the fluid in the annulus, or the type (or lack) of gas-lift valve (GLV). The subsea system could also have tiebacks connected to manifolds or to other wells. And, of course, there are well schematics that are entirely different from the one shown in Figure 1. In summary, there is great variability in well configurations and procedural requirements. For this reason alone, implementing a general algorithm or method capable of detecting anomalies or failures in well operations under any operating state is an extremely challenging task.

2.2 Problem Statement and Scope

The objective is to demonstrate the feasibility of a system capable of analyzing multiple selected sensor measurements from a well over a selected time range under any scenario. In its ideal form, this analysis should be able to cross-reference process variables, such as pressure and temperature at different positions, with operationally controlled variables, such as valve openings or lift-gas, diesel, or any other fluid injection, to check whether the changes are compatible with each other. If any change exhibits abnormal behavior and cannot be associated with other variables, it should result in an alarm. The system should also separate and classify the time range into different periods, providing a summary with crucial information about the well's activity over the analyzed time range. Finally, this system must be adaptable to different well configurations, equipment, and operational sequences.

To provide a manageable initial step in this complicated task, the system developed in this work was designed specifically for pressure measurements and valve openings. A flowchart representing the proposed solution is shown in Figure 2. It was validated using real data from Brazilian pre-salt wells with configurations similar to the one presented in Section 2.1. Furthermore, the analyzed cases focused on the main flow path of the well, that is, the path from the reservoir completion, passing through the ICVs, production tubing, and flowline, and finally reaching the production choke valve. To further simplify the analysis, all studied cases involved naturally flowing wells, i.e., without gas-lift injection.

The developed system can still be adapted to other well configurations. The segmentation method primarily depends on three input parameters related to the well dynamics, while the matching algorithm depends on additional parameters, the topology (which can also be easily modeled), and the physical knowledge of the system.

3 Preliminaries about Changepoint Detection Theory

The changepoint detection problem consists of identifying significant changes in time-series data, such as abrupt level shifts, slope changes, variance changes, or other variations in the behavior of a signal (Truong et al. 2020). Online changepoint detection techniques are designed to use only current and past observations for real-time detection, while offline methods analyze an entire time series after the data have been collected. The offline formulation starts with a signal vector $\mathbf{y} \in \mathbb{R}^n$:

$$\mathbf{y} = (y_0, y_1, \ldots, y_t, \ldots, y_{n-1}), \qquad t \in \{0, 1, \ldots, n-1\}. \tag{1}$$

where t denotes the time index of the element y_t, and n is the total number of elements in $\mathbf{y}$.

The formulation can assume that the number of changepoints is either known or unknown (Truong et al. 2020). In this work, since there is no prior knowledge of the number of changepoints in the signal

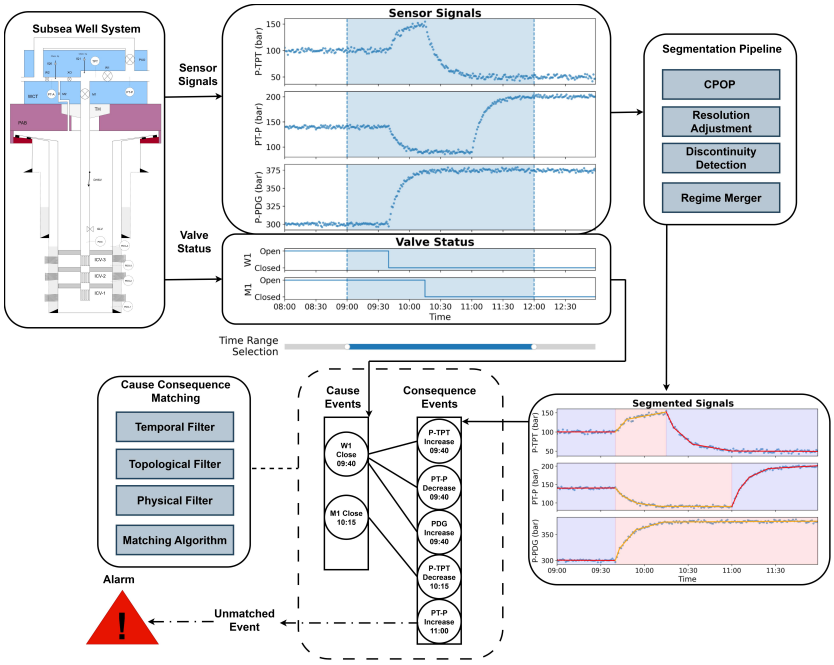

Fig. 2 Complete segmentation and matching framework flowchart. First, a subsea well system of interest is selected, with its sensors, valves, and other equipment. A time range of interest is also selected. A segmentation pipeline, consisting of the main segmentation algorithm (CPOP) and post-processing steps designed to deal with measurement resolution, discontinuities, and non-linear natural regimes, is applied to sensor measurement time series to return a piecewise-linear segmentation and the relevant changepoints. The valve changes are grouped as a set of cause events, while the sensor changepoints obtained from the segmentation pipeline are grouped as consequence events. A cause-consequence matching methodology is applied to these two sets, where first three different filtering steps—temporal, topological, and physical—are run to uncover the valid candidate matches. After that, a final matching algorithm selects the valid cause-consequence matches. The final result is a summary of all the relevant changes and how they are connected to each other. Unmatched events, depending on the system configuration, trigger an alarm for further investigation.

to be analyzed, we adopt the unknown-number approach. A changepoint set is defined as

$$\begin{aligned} \mathcal{T} &= \{\tau_1, \tau_2, \ldots, \tau_K\}, \\ 0 = \tau_0 < \tau_1 < \tau_2 < \tau_3 < \ldots < \tau_K < \tau_{K+1} &= n-1, \end{aligned} \tag{2}$$

where K is the number of changepoints, $\tau_i \in \{1, \ldots, n-2\}$, $i = 1, \ldots, K$, and τ_0 and τ_{K+1} are dummy indexes used to fit the endpoints $t = 0$ and $t = n-1$ in the mathematical formulation. The optimal changepoint set $\mathcal{T}^*$ is defined as

$$\mathcal{T}^* = \arg\min_{\mathcal{T}} \left\{ C\left(y_0 : y_{\tau_1}\right) + \sum_{i=1}^{K} C\left(y_{\tau_i+1} : y_{\tau_{i+1}}\right) + \lambda h(\mathcal{T}) \right\}. \tag{3}$$

where

- $C(\cdot)$ is a cost function, such as the residual sum of squares or negative log-likelihood, that represents the goodness of fit within each segment,
- $y_{\tau_i+1} : y_{\tau_{i+1}}$ denotes the segment containing the values y_t for $\tau_i + 1 \leq t \leq \tau_{i+1}$, while the first segment contains the observations from $t = 0$ to $t = \tau_1$,
- $h(\cdot)$ is a complexity function, usually related to the number of changepoints, used for regularization,
- λ is a penalty parameter that controls the trade-off between goodness of fit and model complexity.

This problem can be solved using different methodologies (Truong et al. 2020). In this work, a parametric approach is adopted to obtain a segmented representation $\mathbf{x} \in \mathbb{R}^n$, which serves as an approximation of the original signal $\mathbf{y}$. The Residual Sum of Squares (RSS) was selected as the cost function, while the ℓ_0 penalty, which is essentially the number of changepoints $|\mathcal{T}| = K$, was chosen for regularization.

Additionally, a method that provides a continuous piecewise-linear segmentation of the raw signal was desired for the subsequent post-processing steps. Continuity is a requirement for the resolution adjustment and regime merging methods described in Sections 4.1 and 4.3. Although discontinuities will be considered, they will be identified by fixed rules, as described in Section 4.2; in that way, it is possible to introduce discontinuities only in segments with predetermined characteristics. The method also needed to be computationally efficient, as it must be applied to multiple signals and potentially long time ranges. Finally, a globally optimal solution is preferred over a potentially suboptimal greedy solution.

Given these requirements, Continuous-Piecewise-Linear Pruned Optimal Partitioning (CPOP), presented in Fearnhead et al. (2019), was chosen as the core segmentation algorithm of the framework. CPOP is based on optimal partitioning, a dynamic programming approach that searches for the changepoint configuration minimizing the complete penalized cost, thereby providing a globally optimal solution rather than relying on greedy decisions.

Standard optimal partitioning assumes that segment costs are independent, which is not valid when continuity is enforced because neighboring segments share the same fitted value x_τ at each changepoint τ. CPOP solves this by jointly optimizing the changepoint set $\mathcal{T}^*$ and the fitted values at the changepoints and at the endpoints $t = 0$ and $t = n - 1$. This allows continuity to be enforced while retaining global optimality for the ℓ_0-penalized RSS objective. In addition, pruning strategies are used to reduce the computational burden and improve the efficiency of the dynamic programming procedure.

The resulting continuous piecewise-linear segmentation $\mathbf{x} \in \mathbb{R}^n$ approximates the original signal $\mathbf{y}$, with each x_t in a segment between changepoints τ_i and τ_{i+1} defined as:

$$x_t = \phi_i(t - \tau_i) + \theta_i, \qquad \tau_i \leq t \leq \tau_{i+1}. \tag{4}$$

where $\theta_i = x_{\tau_i}$ is the fitted value at τ_i and ϕ_i is the slope of the segment between changepoints τ_i and τ_{i+1}:

$$\phi_i = \left(\frac{\theta_{i+1} - \theta_i}{\tau_{i+1} - \tau_i} \right). \tag{5}$$

For this work, the R package *cpop*, presented in Fearnhead and Grose (2022), was used for the data analysis. The main algorithm solves the problem:

$$\min_{\mathcal{T}, \boldsymbol{\theta}} \left\{ \sum_{t=0}^{n-1} \frac{(y_t - x_t)^2}{\sigma_t^2} + \lambda K \right\}, \qquad K = |\mathcal{T}|. \tag{6}$$

where σ_t is the standard deviation of the observation noise at time t and $\boldsymbol{\theta} \in \mathbb{R}^{K+2}$ is the vector of fitted values at the changepoints and endpoints:

$$\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_K, \theta_{K+1}) = (x_0, x_{\tau_1}, \ldots, x_{\tau_K}, x_{n-1}). \tag{7}$$

The noise is estimated following Fearnhead et al. (2019) as a single value for the entire time series, using the scaled median absolute deviation of second differences:

$$\sigma_{\text{noise}} = \frac{1.4826}{\sqrt{6}} \, \text{MAD}(\Delta^2 y), \tag{8}$$

where MAD is the unscaled median absolute deviation and $\Delta^2 y_t = y_{t+2} - 2y_{t+1} + y_t$.

In order to find the best penalty parameter λ, a set of possible solutions was first obtained using the CROPS method (Haynes et al. 2017), which identifies the distinct optimal changepoint solutions over a specified penalty interval, avoiding the need to evaluate a dense grid of penalty values individually. The solutions found by CROPS were then compared using the Bayesian Information Criterion (BIC) (Schwarz 1978). For this problem, the BIC was formulated as in Fearnhead and Grose (2022):

$$\text{BIC} = n \log \left(\frac{\sum_{t=0}^{n-1} (y_t - x_t)^2}{n} \right) + 2K \log(n). \tag{9}$$

The CROPS method requires a penalty interval for λ. If the noise variance is estimated accurately, the first term of the CPOP objective from Equation 6 is approximately equivalent, up to an additive constant, to the first term of the BIC criterion from Equation 9. In that case, setting the penalty $\lambda = 2 \log n$ makes both objective functions approximately equivalent. A CROPS range between $\log n$ and $5 \log n$ was chosen in order to account for noise estimation and other uncertainties.

4 Segmentation Pipeline Methodology

This section describes the pipeline implemented to perform piecewise segmentation of well production data. When applying CPOP together with the CROPS method described in Section 3, several issues were identified that could lead to segmentation overfitting and an excessive number of spurious changepoints. These issues were mainly related to sensor resolution, sporadic discontinuities, and the non-linear nature of transient phases. To address these problems, post-processing steps were designed to adjust the segmentation, remove spurious changepoints, and reduce overfitting. It is also important to note that the methodologies developed in this section were primarily designed for pressure measurements. Although they can also be applied to temperature signals, some of the post-processing steps may be less suitable depending on the underlying temperature dynamics.

The segmentation pipeline flowchart is shown in Figure 3. First, the initial piecewise-linear segmentation is performed using CPOP combined with the CROPS method, returning an initial segmentation $\mathbf{x}$ and a set of changepoints $\mathcal{T}$.

Next, a resolution-adjustment post-processing step, described in Section 4.1, is applied to the initial segmentation, returning the final segmentation $\mathbf{x}_{\text{final}}$. The purpose of this step is to avoid over-segmentation caused by minor, negligible changes.

A discontinuity-detection step, described in Section 4.2, is then performed to classify short segments found by CPOP as abrupt level changes. The segmentation itself is not changed, but the last changepoint of the short segment is removed from further analysis.

Finally, a regime-merging heuristic, described in Section 4.3, is used to define the relevant changepoints $\mathcal{T}_{\text{final}}$ by merging consecutive segments that are interpreted as belonging to the same underlying regime, based on the slope evolution of $\mathbf{x}_{\text{final}}$. This step also does not change the segmentation; it only removes irrelevant changepoints from further analysis.

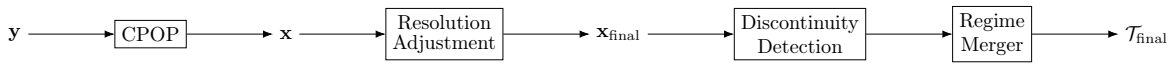

Fig. 3 Overview of the segmentation pipeline. As described in Section 3, the raw signal $\mathbf{y}$ is processed by the CPOP algorithm. To optimize the regularization hyperparameter λ, the CROPS method is applied to return all possible solutions for a chosen range of λ values, and the segmentation $\mathbf{x}$ with the minimum BIC (Equation 9) is chosen. After that, a resolution adjustment post-processing step is applied to the segmentation $\mathbf{x}$ in order to merge segments generated by operationally irrelevant changes. This step returns a final segmentation $\mathbf{x}_{\text{final}}$. The changepoint set from this segmentation is then processed by two further steps. The first step, Discontinuity Detection, detects discontinuities and removes the extra changepoints generated to represent level jumps with linear segments. The second one—Regime Merger—identifies sequences of segments representing non-linear natural pressure behavior and removes the inner changepoints. The final changepoint set $\mathcal{T}_{\text{final}}$ contains all changes to be further analyzed.

This pipeline also uses the following user-defined inputs for each sensor:

1. $\delta_{\min}$: minimum resolution, defined as the smallest individual sensor variation considered operationally relevant. This value is related to the minimum change that should be considered in the absence of noise.
2. $m_{\max}$: maximum discontinuity size, defined as the maximum segment length for which a segment can be considered a dynamical discontinuity. This value should be chosen based on the characteristic timescales of the well dynamics and the sampling frequency.
3. ϕ_{lim}: slope limit, used to classify constant regimes. This value should be related to the maximum continuous rate of change considered compatible with steady-state behavior.

4.1 Resolution Adjustment Post-Processing

Because some sensors can have very high resolution, small but real oscillations originating from the reservoir or production system may lead to changepoints that are not operationally relevant. A PDG sensor, for example, can exhibit a very stable overall trend under steady-state conditions while still capturing small process variations, such as minor disturbances in separator pressure.

To remove changepoints that do not represent operationally meaningful changes, a post-processing method, shown in Algorithm 1, was applied. This method uses only one user-defined input, the minimum resolution $\delta_{\min}$.

For each internal changepoint τ_i, an interpolated value z_{τ_i} is calculated by connecting the fitted values at the previous and next changepoints with a straight line. This interpolation step is illustrated in Figure 4. The deviation

$$d_\tau = |z_\tau - x_\tau| \tag{10}$$

is then calculated for each internal changepoint. After all values are evaluated, the changepoint τ^* with the smallest deviation is selected. If this deviation is smaller than the minimum resolution,

$$d_{\tau^*} < \delta_{\min}, \tag{11}$$

the changepoint is removed.

The fitted segmentation $\mathbf{x}$ is then recalculated using the remaining changepoints. For this purpose, a linear-spline least-squares fit with fixed knots (detailed in Appendix A) is applied to the raw data $\mathbf{y}$ to generate a new resolution-adjusted segmentation $\mathbf{x}$. The same process is repeated until no changepoint satisfies the removal criterion or until only the endpoints τ_0 and τ_{K+1} remain.

Algorithm 1 Resolution-adjustment post-processing

Require: Ordered changepoints $\mathcal{B} = (\tau_0, \ldots, \tau_{K+1})$, original signal $\mathbf{y}$, fitted signal $\mathbf{x}$, resolution threshold $\delta_{\min} > 0$
Ensure: Adjusted changepoints $\mathcal{B}$ and fitted signal $\mathbf{x}$

while $|\mathcal{B}| > 2$ **do**
  for $j = 1, \ldots, |\mathcal{B}| - 2$ **do**
    $(\tau_-, \tau, \tau_+) \leftarrow (\mathcal{B}_{j-1}, \mathcal{B}_j, \mathcal{B}_{j+1})$
    $d_\tau \leftarrow \left| x_\tau - x_{\tau_-} - (\tau - \tau_-) \dfrac{x_{\tau_+} - x_{\tau_-}}{\tau_+ - \tau_-} \right|$
  end for
  $\tau^* \leftarrow \underset{\tau \in \mathcal{B} \setminus \{\tau_0, \tau_{K+1}\}}{\operatorname{argmin}} d_\tau$
  if $d_{\tau^*} \geq \delta_{\min}$ **then**
    break
  end if
  $\mathcal{B} \leftarrow \mathcal{B} \setminus \{\tau^*\}$
  Refit $\mathbf{x}$ using $\mathcal{B}$ and $\mathbf{y}$ ▷ Continuous piecewise-linear fit
end while
return $\mathcal{B}$, $\mathbf{x}$

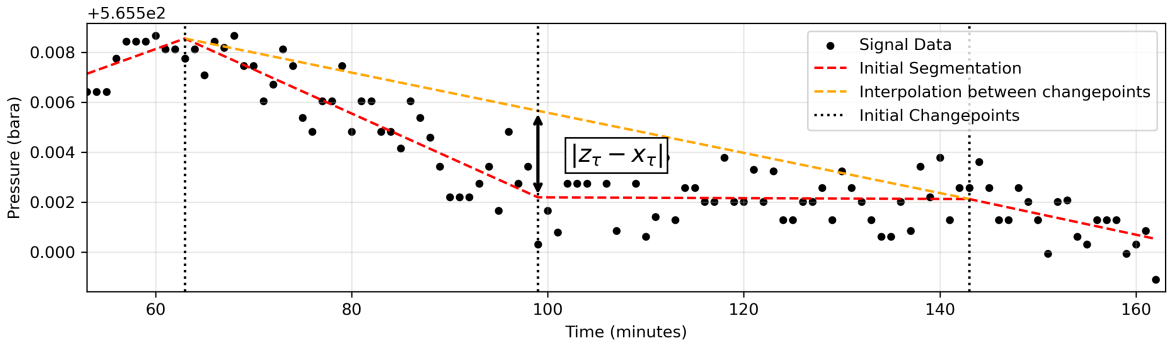

Fig. 4 Resolution-adjustment post-processing interpolation step illustration. Two linear segments, one between changepoints τ_{i-1} and τ_i, and another between changepoints τ_i and τ_{i+1}, are compared to a single linear segment between changepoints τ_{i-1} and τ_{i+1}. The difference $|z_\tau - x_\tau|$ is calculated for all pairs of consecutive segments using this same step.

Figure 5 shows an example where this post-processing works well. The total variation is very small: the range of the signal over 12 hours is no greater than 0.05 bar, and a band with a width of 0.02 bar around the midline contains most of the data points. Even so, because of the high measurement resolution, the signal can exhibit small-scale changes that are statistically significant. The CPOP algorithm initially fits multiple segments (solid red line) that do not represent operationally significant changes. After resolution adjustment, all changepoints are removed, and the signal is represented by a single line (yellow dashed line).

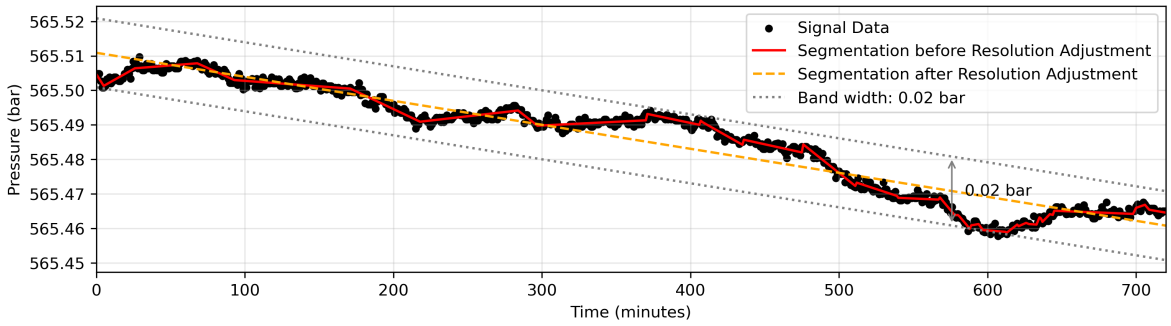

Fig. 5 Resolution-adjustment post-processing application example. The original segmentation, shown by the red splines, covers all minor disturbances detected by the sensor. A band with a range of 0.02 bar can cover almost all raw signal points. After applying the resolution-adjustment method, the resulting segmentation, consisting of a single linear spline shown as a dashed yellow line, implies that no relevant sensor changes occurred during the selected time period.

4.2 Discontinuity detection

Although continuity is a desirable property of the piecewise-linear segmentation method for this problem, very fast transient behavior following valve operations can sometimes lead to an additional changepoint. This can occur, for example, when a low-pressure volume is rapidly connected to another volume at a higher pressure. The resulting rapid pressure equalization produces a sharp transition to a new pressure level, which CPOP may represent as a single short segment followed by a longer segment with a slower trend, separated by an additional changepoint.

To avoid this additional changepoint, a simple procedure was developed. A segment with slope ϕ_i, between changepoints τ_i and τ_{i+1}, is considered a jump if:

1. The segment contains at most $m_{\max}$ sampled points:

$$1 \leq \tau_{i+1} - \tau_i \leq m_{\max} - 1, \tag{12}$$

2. Its slope lies outside the range defined by the slopes of the neighboring segments:

$$\phi_i < \min(\phi_{i-1}, \phi_{i+1}) \quad \text{or} \quad \phi_i > \max(\phi_{i-1}, \phi_{i+1}). \tag{13}$$

The first criterion is related to short-term changes, while the second is related to dynamical discontinuities. If both criteria are satisfied, changepoint τ_{i+1} is removed from the final changepoint set and, consequently, excluded from the cause–consequence matching analysis presented in Section 5 or any other use. If two consecutive segments meet these criteria, only the first will be considered a discontinuity. This was implemented to avoid missing changepoints for successive changes.

Figure 6 shows an example of how the discontinuity detection operates. In the example, two abrupt changes occurred in the pressure signal, which was modeled by CPOP as a short segment of two point samples, leading to two changepoints for each occurrence, while in reality just one change happened. The left chart shows the changepoint separation before the discontinuity detection, while the right shows the separation after the method is applied and the extra changepoints are removed.

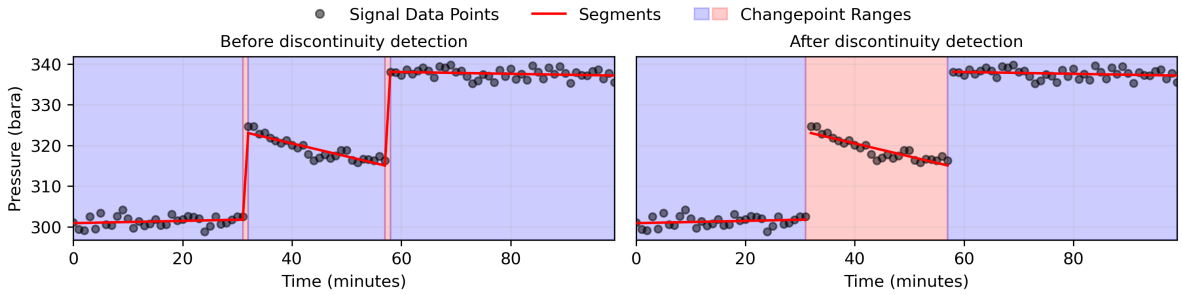

Fig. 6 Discontinuity detection application. In the left chart, it is possible to see two level jumps represented by single segments connecting only two points. Each of these jumps also results in two changepoints, resulting in an extra changepoint to describe a single event. These segments are short, and their slope values are unbounded by the neighboring slopes, which fits the criteria for identifying them as discontinuities. The discontinuity detection step removes the last changepoint from each segment, resulting in the right chart. In this work, discontinuities are graphically represented by simply not having a segment connecting the neighboring segments.

4.3 Regime Merging

A major limitation with piecewise-linear segmentation is that well data are not always linear. Pressure measurements usually present curved behavior during transient periods. Yet, when no disturbance is applied to a well system, pressure is usually expected to remain constant and stable, or to be in a transient phase evolving monotonically towards a new equilibrium, with the rate of change progressively decreasing. This leads to two main transient scenarios: an increasing trend at a decelerating rate (e.g., tubing pressure under shut-in conditions), or a decreasing trend at a decelerating rate (e.g., pressure in the flowline during cooldown).

When a transient phase is approximated by piecewise-linear segments, it can be represented by consecutive slopes with the same sign and progressively decreasing magnitude. Thus, the changepoints between these segments do not necessarily represent real physical changes, but rather the natural evolution of the rate of change during the transient.

To represent these dynamics, a simple heuristic based on sequential slope values was implemented. Let ϕ_{i-1} and ϕ_i be the slopes of two adjacent segments separated by changepoint τ_i, and let $\phi_{\text{lim}} > 0$ represent a user-defined threshold for the absolute slope of a constant regime. The changepoint τ_i is removed and both segments are considered part of the same regime if any of the conditions specified in Table 1 holds true.

In this work, a regime is defined as a maximal sequence of adjacent segments whose consecutive slopes satisfy the same merging condition. Therefore, if segments $i, \ldots, j$ are merged, the resulting regime covers the interval from τ_i to τ_{j+1}. Each regime is classified as constant, decelerating decreasing, or decelerating increasing according to the conditions defined in Table 1.

Since transient phases tend toward stability when no further perturbation occurs, increasing or decreasing regimes followed by constant regimes are also merged, suppressing the changepoint between them. Figure 7 shows a visual example of how this method works.

Table 1 Regime classification and merging conditions for adjacent linear segments.

Regime	**Merging condition**
Constant	$\lvert\phi_{i-1}\rvert < \phi_{\text{lim}}$ and $\lvert\phi_i\rvert < \phi_{\text{lim}}$
Decelerating Decreasing	$\phi_{i-1} < -\phi_{\text{lim}}$ and $\phi_{i-1} < \phi_i < \phi_{\text{lim}}$
Decelerating Increasing	$\phi_{i-1} > \phi_{\text{lim}}$ and $\phi_{i-1} > \phi_i > -\phi_{\text{lim}}$

Regarding the applicability of this heuristic to operational changes, most of the valves described in Section 2.1 are fast-opening valves and can therefore be approximated as discrete events if the sampling period is longer than the duration of the valve movement, which justifies the non-monotonic slope-change rule. However, if the valve opening time is longer than the sampling period, an initial acceleration could appear in the sampled signal. This means that a higher sampling frequency could lead to spurious changepoints. For this work, a 1-minute sampling interval performed well on the pressure signals studied.

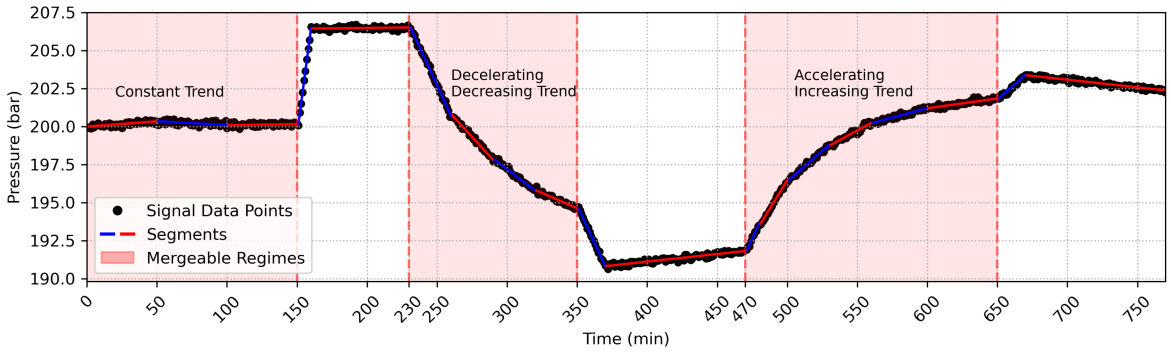

Fig. 7 Regime merging illustration. The first shaded area (between 0 and 150 minutes) contains three segments with a small slope. These three segments are considered to be the same constant steady-state event. The second shaded area (between 230 and 350 minutes) contains four segments in a decelerating decreasing trend, which means that the connected splines have negative slopes with monotonically diminishing absolute values. These four segments are considered to be the same transient pressure decrease event. At the 350-minute mark, the depressurization rate increased, and the data points disconnected from the curvature formed by the preceding segments. The third shaded area (between 470 and 650 minutes), on the other hand, has the opposite behavior, as it merges six segments into a single transient pressure buildup event. It contains six segments in a decelerating increasing trend, which means that the connected splines have positive slopes with monotonically diminishing absolute values. At the 630-minute mark, the pressure buildup rate increased, and the data points disconnected from the curvature formed by the preceding segments.

On the other hand, the choke valve can exhibit a different behavior. A single step change in choke opening is expected to produce a similar transient response. However, a continuous change in choke opening may lead to non-monotonic slope evolution. This should not necessarily be interpreted as a failure of the method, since a continuous choke movement can be viewed as a sequence of successive valve-opening changes, and the presence of multiple relevant changepoints is therefore physically plausible. The treatment to match these changepoints against choke changes is better explained in Section 5.1.1.

This heuristic has several limitations. First, it depends on a user-defined value (ϕ_{lim}). It may also fail to remove spurious changepoints in cases with accelerating trends without additional perturbations, as it does not account for all possible pressure dynamics. Slow-moving valves or high-frequency sampling may lead to spurious changepoints due to accelerating trends at the start of a change. Yet, overall, it obtained good results when applied to pressure measurements from well operational data. In future broader applications, this problem may also arise in situations such as continuous changes in gas-lift rate or pump power. However, because these are causal actions that can also be detected through monitored signals, the resulting changepoints would not necessarily remain unexplained. Nevertheless, future work should aim to provide a more general and application-agnostic method for regime merging.

5 Matching Cause and Consequence Changepoints

In the previous section, a segmentation pipeline was implemented to process measurements, especially pressure measurements, from a subsea petroleum well. This pipeline returns a segmentation $\mathbf{x}_{\text{final}}$ and a set of changepoints $\mathcal{T}_{\text{final}}$ for the selected sensors, which can be any of those described in Section 2.1. However, these results alone, without additional well state information such as valve status and choke opening values, do not provide sufficient information to understand the ongoing evolution of each signal.

As described in Section 2.1, a subsea well contains multiple downhole and subsea valves, as well as pressure sensors for each valve hydraulic activation line. By analyzing the hydraulic pressure level, one can infer whether a valve is open or closed. On the other hand, the choke valve already provides an opening signal, which indicates its opening percentage from 0 to 100%. Depending on the operational state of the well, opening or closing a valve may produce observable changes in one or more sensor measurements, or it may have no measurable effect.

In order to obtain useful information from this entire set of data, a matching methodology was developed, using the segmentation pipeline's changepoint results together with valve and equipment statuses as inputs. First, operational commands, such as valve manipulations, are grouped as cause

events, and sensor changepoints as consequence events. Three filtering steps (temporal, topological, and physical) select which cause–consequence pairs are feasible. Finally, a matching algorithm selects the final matches, providing an interpretation of the selected time range and highlighting unmatched changepoints and cause events with missing expected responses as points for investigation.

5.1 Definitions

The matching algorithm basically connects two types of events: cause and consequence. Before presenting the algorithm, a brief definition containing the crucial information of each kind of event is presented.

5.1.1 Cause Events

A cause event is an operational action associated with a specific piece of equipment that may cause a response in one or more pressure and temperature measurements. In this work, cause events are represented as discrete events:

$$e = (v, t, a) \tag{14}$$

where v is the valve or equipment involved, t is the event time, and a is the action performed (for valves, essentially opening or closing).

As this work focused only on valves, two main types of actuators were considered: fast-moving valves and the choke valve. The first type includes the WCT valves (for example, M1 and W1), the surface ESDV, and the ICVs. Although the ICVs have intermediate opening positions, they are also operated in discrete, incremental steps.

The status of fast-moving, on-off subsea valves is simply given by their hydraulic actuation pressure. If the pressure is below a certain threshold, the valve is considered closed, while if it is above the threshold, it is considered open. When closed, the valve pressure corresponds approximately to the environmental pressure (in the subsea scenario, the seawater pressure at that depth). When open, an actuation pressure, defined by the design of each system, is added to this pressure. Intermediate pressure values are not expected; therefore, a threshold at the midpoint of this range is sufficient to distinguish between open and closed states. The same logic applies to ICVs, with the difference that each time the pressure transitions from low to high in a given hydraulic line, the valve moves one incremental step.

On the other hand, the choke valve can be opened continuously, and its opening signal can be quite noisy. A smoothing technique, IRFL (Iteratively Reweighted Fused Lasso), described in Grantham et al. (2026), was applied using first-order differences. This method was selected because it preserves abrupt changes while reducing noise. It avoids the need to manually select a smoothing weight by evaluating a regularization path and selecting the solution according to the BIC criterion. The first-order differences of the smoothed signal were then calculated, and time steps where the absolute difference exceeded a user-defined threshold δ_{choke} were considered cause events, with the sign of the choke change indicating opening or closing action. Because of that, a continuously changing choke valve can produce multiple discrete events.

5.1.2 Consequence Events

A consequence event is a change of regime in a process variable, such as pressure or temperature, which is detected and characterized by the segmentation pipeline described in Section 4. It is represented as:

$$c = (s, t, r) \tag{15}$$

where s is the sensor involved, t is the changepoint time, and r is the regime classification after the change. As already explained, the framework was designated primarily for pressure signals, although temperature signals were also used in Section 6.5.

5.2 Relationship Filtering

At first, each candidate cause–consequence pair is evaluated against three filtering criteria: temporal compatibility, hydraulic connectivity (topological criterion), and consistency with the expected sensor response (physical criterion).

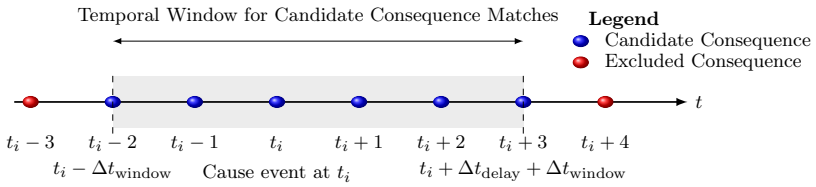

Fig. 8 Illustration of the temporal compatibility filter, using $\Delta t_{\text{window}} = 2$ and $\Delta t_{\text{delay}} = 1$ for a specific sensor-valve pair. Only changepoints detected within the gray-colored window can be considered candidate matches for the cause event happening at t_i.

5.2.1 Temporal Filter

The **temporal** filter restricts candidate pairs according to their temporal separation. A consequence changepoint (sensor response) τ_i^r is considered a candidate match for a cause event occurring at τ_j^c if

$$-\Delta t_{\text{window}} \leq \tau_i^r - \tau_j^c \leq \Delta t_{\text{delay}} + \Delta t_{\text{window}}. \tag{16}$$

The temporal parameters Δt_{window} and Δt_{delay} are user-defined engineering tolerances, which may be assigned for each specific valve-sensor pair. Δt_{window} accommodates uncertainty in changepoint localization by the segmentation pipeline (which means that a consequence changepoint could appear before its cause), while Δt_{delay} allows for a delay between an operational command and the onset of an observable sensor response. Figure 8 illustrates the filter logic.

5.2.2 Topological Filter

The **topological** filter evaluates whether a sensor is hydraulically connected to either side of the valve associated with a cause event. The system is represented as an undirected graph, in which nodes represent hydraulic regions and edges represent valves connecting those regions. For each event, the graph is constructed using the valve configuration immediately before the event, with closed valves excluded. The actuated valve is temporarily removed, and a sensor is considered potentially affected if its node belongs to a connected component containing either endpoint of that valve.

For example, when W1 is closed and no alternative hydraulic path to the topside exists, opening the ESDV or changing the production choke position cannot directly affect the PDG or TPT measurements through the modeled hydraulic connections. The topological filter therefore excludes these candidate pairs. Hydraulic connectivity establishes the possibility of a response, but does not guarantee that an observable response will occur. When a differential pressure measurement is used (as in Section 6.3), both positions are analyzed, and the pair passes the criterion if at least one of the positions can be affected by the valve change. Figure 9 illustrates a topological filter example.

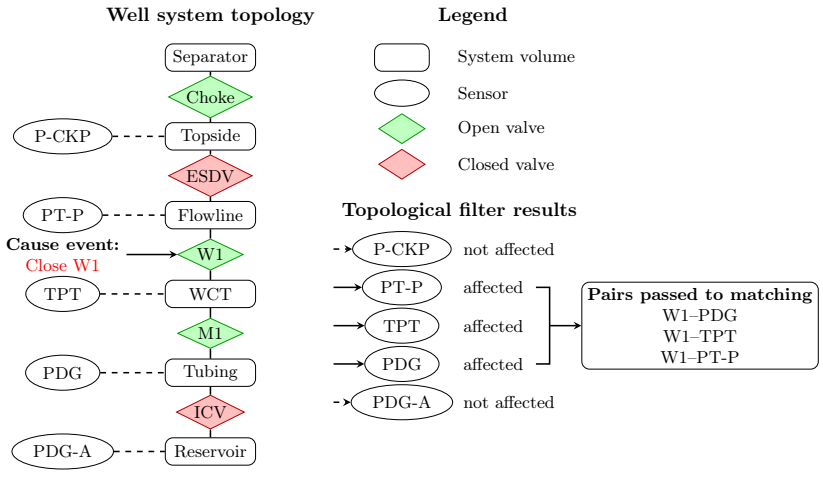

Fig. 9 Illustration of the topological compatibility filter. The well system is represented as a graph in which valves connect adjacent system nodes and sensors are associated with the corresponding topology nodes. For a given cause event, only sensors that can be affected through the current system topology are retained for subsequent matching.

5.2.3 Physical Filter

A **physical** filter is applied to check if the sensor response follows the expected pattern for the cause. This filter is optional in the current framework, as deriving all possible responses for every combination of causes and initial states can be a very difficult task for a rule-based algorithm. If no physical criterion is established for a cause–consequence pair in a given state, the pair can still be considered a valid candidate if it fulfills the temporal and topological criteria.

However, the matching algorithm will prefer cause–consequence pairs with known physical compatibility. Another important detail is that causes without any candidate consequence match for a specific sensor will only generate alarms if a response was expected; otherwise, no alarm is generated.

For this work, a simple rule-based algorithm was employed, with rules derived for the most important scenarios. It was based on three general scenarios:

- **Flow restriction change:** If the opening of a valve along a flow path is changed, the expected immediate pressure response of a sensor (a trend or a jump) depends on the flow direction, its position relative to the valve (upstream or downstream), and the valve action (increasing or decreasing the opening). This can be determined using a simple algebraic rule.
- **Hydraulic connection:** If a valve is opened between two volumes and pressure sensors are available both upstream and downstream, an equalization jump is expected in both sensors, i.e., a fast increase in the lower-pressure chamber and a fast decrease in the higher-pressure chamber. An estimate of the final pressure is calculated using the volume of each region. In this way, if the expected pressure change at a sensor is smaller than $\delta_{\min}$, no change is expected.
- **Thermal isolation:** If a valve is closed, resulting in a hydraulically isolated volume with no connection to the reservoir or surface separator, its temperature is compared with the known ambient temperature. In the case of ambient heating, an increasing temperature trend is expected, while in the case of cooling, a decreasing trend is expected.

Although all three rules are used to evaluate physical compatibility, only the flow restriction and the hydraulic connection rules are used to generate alarms from unmatched cause events from expected but missing responses. The thermal isolation rule is excluded from this assessment because a valve closure does not necessarily produce a new detectable changepoint when the well is already shut in and undergoing thermal evolution.

Sometimes there can also be uncertainty without additional information about the well state; for example, an ICV opening change does not necessarily cause a noticeable shift in the downhole pressure, depending on the circumstances, especially in the first (minimum change) and last (valve opening already close to fully open) increments. Another example is that repeated opening and closing of the M1 valve while W1 is closed may produce no further noticeable changes in TPT pressure once the connected regions have reached pressure equilibrium. Because of this complexity, the current rule-based approach only works for a limited set of well states. In future work, a possible extension would be to use a calibrated digital twin system of the well, which could be a multiphase flow simulation model fitted to the well data.

5.3 Matching Algorithm

After the filtering step, an **order-preserving, preference-aware maximum-cardinality, minimum-cost bipartite matching algorithm** is applied for each sensor to assign cause events to consequence changepoints among the valid candidate pairs. This algorithm is explained in Appendix B. Figure 10 illustrates how this matching occurs.

Each consequence can have only one matched cause, while each cause can have multiple matched consequences from different sensors, but only one per sensor. The priority is to first assign the maximum number of cause–consequence matches which are physically expected. After this, maximize the total number of cause–consequence matches, including those not expected, but also not discarded by the physical filter. Finally, the cost function chooses the remaining pairs with the smallest temporal distance, while penalizing the pairs where the consequence was detected before the cause. This cost function for a cause at time τ_i^c and a consequence (or response) at time τ_j^r is:

$$c(\tau_i^c, \tau_j^r) = \begin{cases} 2(\tau_i^c - \tau_j^r) + 0.1, & \tau_j^r - \tau_i^c < 0, \\ \tau_j^r - \tau_i^c, & \tau_j^r - \tau_i^c \geq 0. \end{cases} \tag{17}$$

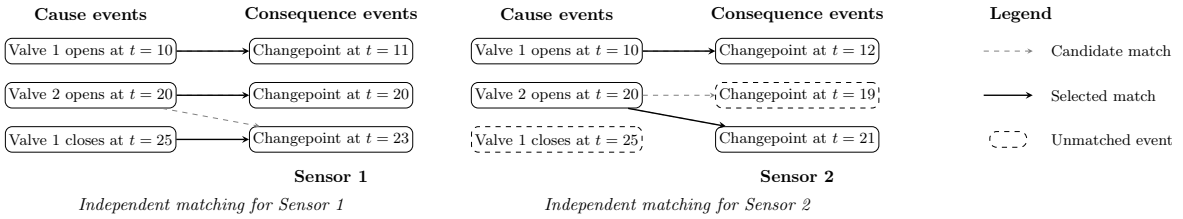

Fig. 10 Illustration of the cause–consequence matching procedure. Matching is performed independently for each sensor. Dashed edges represent candidate pairs obtained after the filtering steps, while full edges represent selected cause–consequence matches.

where the cost is based on temporal distance, and the penalty is doubled if the match associates a cause with an earlier response changepoint. The 0.1 term serves to avoid ties between candidates.

In summary, the assignment works as follows:

1. maximize the number of physically compatible assigned pairs;
2. maximize the number of assigned pairs (maximum cardinality);
3. minimize the total temporal cost (minimum cost)

A consequence changepoint generates an alarm when it receives no final cause assignment. A cause event generates a missing-response alarm when a response is expected according to the flow restriction change and hydraulic connection rules and no valid candidate consequence changepoint exists for that sensor.

Finally, because continuous changes in choke opening can generate multiple cause events that cannot all be uniquely associated with consequence changepoints, choke-change causes were also excluded from missing-response alarms, independently of having an expected physical response or not.

6 Case Studies and Applications

This section shows the results of the retrospective well analysis tool, using the segmentation pipeline presented in Section 4 and the cause–consequence matching algorithm presented in Section 5. It was tested in different scenarios to evaluate its ability to interpret operating regimes and identify abnormal or unexplained behavior under diverse conditions. The data were collected at a frequency of 1 sample/minute for all scenarios. All cases mainly focused on the normal production flow path of the well, from the reservoir to the tubing, through the WCT and flowline, and finally to the topside facilities. Because of that, the cases dealt mainly with changes in ICV, M1, W1, ESDV, and the production choke. Other valves presented in Section 2.1 remained closed in all cases. Finally, the last analysis shows how the segmentation pipeline can also be used as a supporting tool for an online monitoring system.

The parameters used for all cases are presented in Table 2; most of the parameter values, such as $\delta_{\min}$, $\phi_{\lim}$, $m_{\max}$ and δ_{choke}, were chosen according to the operational reality of the wells that provided the data. The ranges $\lambda_{\min}$ and $\lambda_{\max}$ were chosen based on the BIC criterion (Equation 9), as described in Section 3. These parameters can, of course, be changed to fit a different production scenario. For example, in a well with instability and heavy pressure oscillations, increasing $\delta_{\min}$ and $\phi_{\lim}$ is an appropriate measure to avoid many spurious alarms.

6.1 Steady-State Production

The first experiment was conducted on the simplest case possible, in which a production well is operating in a stable steady-state regime for 12 hours. This means that there is no valve opening or closing, and the production choke opening remains constant. The main flowpath valves (ICVs, M1, W1 and ESDV) are always open.

Figure 11 shows the results for the pressure and temperature at the TPT, the pressure at the PDG, the pressure at PT-A, the pressure upstream of the choke, and the choke opening. Besides the PT-A pressure, all time series are represented by only a single linear spline, with no changepoint in any of them. This was the expected result, suggesting that the algorithm is likely to generate few or no spurious changepoints during stable steady-state production.

Table 2 Parameters used in the case studies.

Parameter	Value	Description
$\delta_{\min}$	0.2 bar	Minimum operationally relevant signal variation
$\phi_{\lim}$	1/30 bar/min	Constant-regime slope limit
$m_{\max}$	3 samples	Maximum jump-segment length
$\lambda_{\min}$	$\log n$	Lower CPOP penalty
$\lambda_{\max}$	$5 \log n$	Upper CPOP penalty
δ_{choke}	0.2 %	Minimum discrete choke change
Δt_{window}	3 min	Temporal tolerance around the expected response interval
$\Delta t_{\text{delay}}^{(\text{PDG})}$	2 min	Maximum expected response delay for sensor PDG
$\Delta t_{\text{delay}}^{(s)}$	1 min	Maximum expected response delay for all other sensors

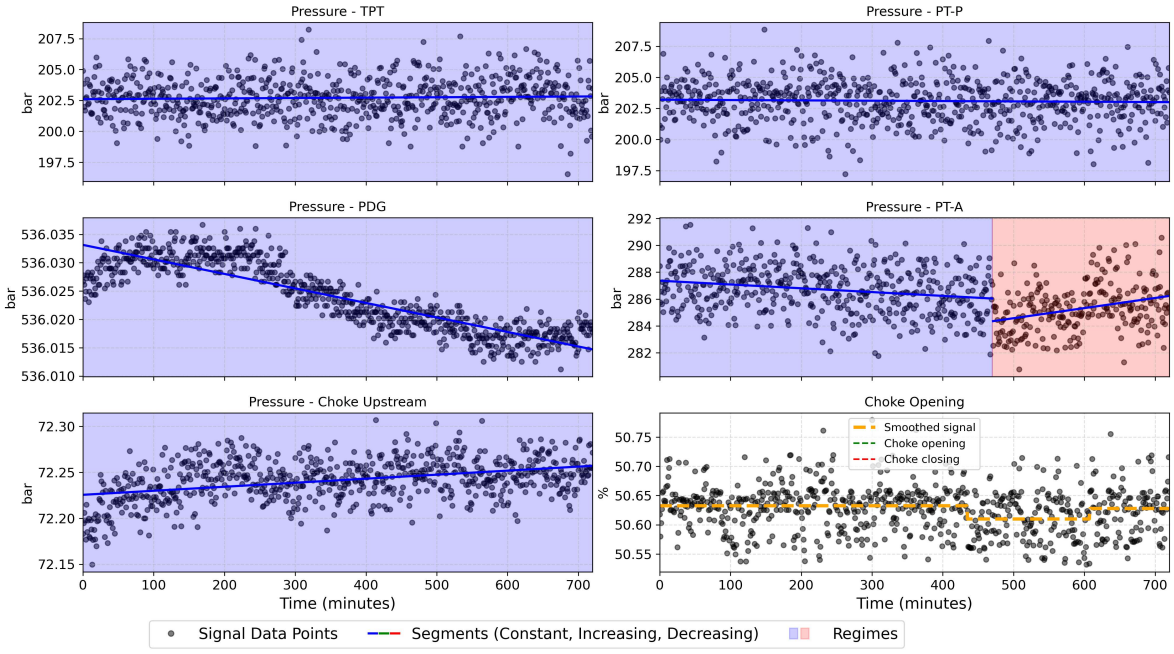

Fig. 11 Steady-state production case. The framework did not detect any changepoints in the pressure sensor signals along the main production flow path (PDG, TPT, PT-P, and Choke Upstream). A single changepoint was detected in the PT-A sensor, possibly caused by the characteristics of its low-volume chamber, where the pressure can be affected by any minor surrounding disturbance.

A single changepoint was detected in the PT-A pressure signal at 470 minutes, represented by a jump followed by a small slope change. The chamber between valves XO, M2, and W2 has a much smaller volume than other chambers (typically 1 against 1000 or 10000) and was filled with a low-compressibility liquid. Consequently, minor disturbances, such as insignificant valve leakage or ambient temperature fluctuations, may cause significant pressure changes. Sensor drift is another possible explanation. This illustrates the difficulty of fully automated interpretation, even in simple scenarios, without physical knowledge of the system.

On the other hand, changes in slope are visually noticeable in the pressure measured at the PDG and, to some extent, upstream of the choke, but the final algorithm returned only a single segment. However, by looking at the signal scale, one can see that, unlike the other sensors, the PDG has very high measurement resolution, with noise around 0.002 bar, which makes even minor fluctuations visible. The same observation applies to the pressure measured upstream of the choke. The resolution-adjustment post-processing step, presented in Section 4.1, transformed the initial segmentation of these signals from multiple splines into a single one.

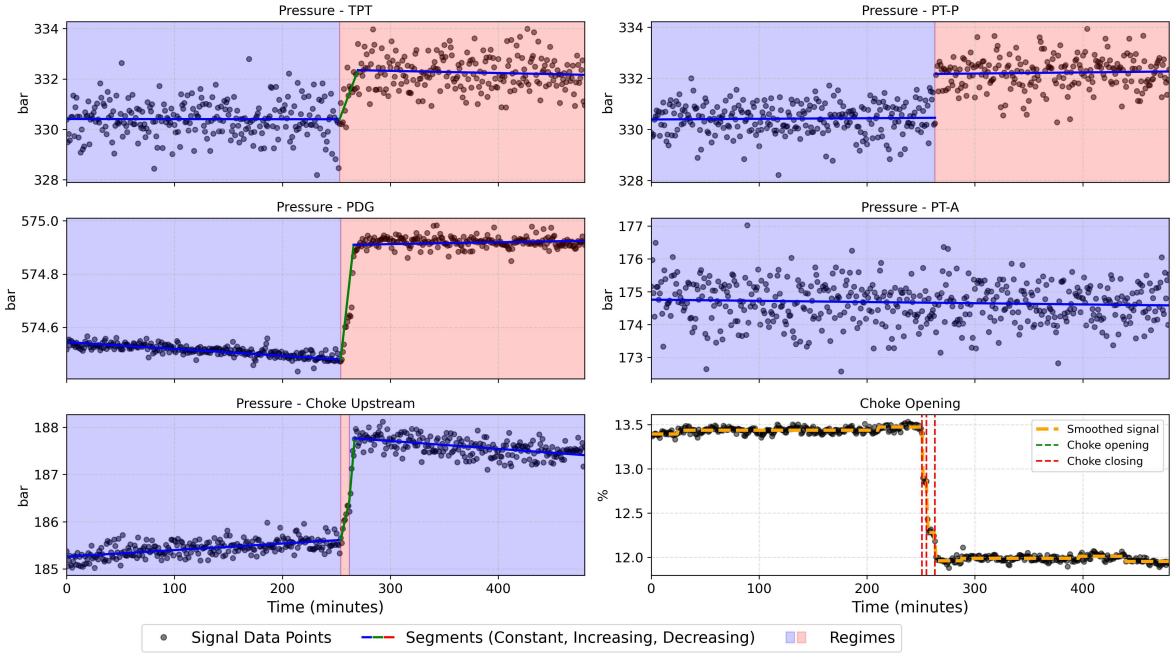

Fig. 12 Choke valve shift case. Changepoints were detected in all pressure sensors along the main production flow path (PDG, TPT, PT-P, and Choke Upstream) around the time the choke valve was manipulated. In the choke opening signal (bottom right chart), three relevant discrete choke changes were also detected, as shown by the red dashed lines representing choke closing events.

6.2 Small Choke Change

The next case, shown in Figure 12, presents one of the simplest operational changes: a small choke shift during steady-state production, while all other valves do not change position. The choke starts to be partially closed, going from a 13.4% opening at $t = 251$ minutes to a 12% opening at $t = 264$ minutes, which can be seen in the bottom-right panel. The methodology described in Section 5.1.1 returned three significant choke changes at $t = 251$, 255 and 263, represented by vertical red dashed lines. The pressure measurements at the TPT, the PDG, and upstream of the choke all presented some increase pattern. The segmentation method was able to correctly detect these changes and differentiate the transient period from the initial and final stable regimes. Because the XO valve was closed during the entire period, there was no change in the PT-A measurements, and the algorithm also correctly did not detect any change.

Table 3 shows the matches obtained. The segmentation had slightly different outcomes for each sensor. The TPT and PDG sensors had the pressure increase represented as a single segment, although the changepoints were at different times (253 and 254 minutes, respectively). The PT-P sensor change was represented by a discontinuity at 263 minutes, while the upstream choke sensor had two changepoints at $t = 254$ and 262 minutes, one representing the start of the increasing trend and the second representing an acceleration of this same trend. These different responses were correctly explained by successive choke changes (three cause events were evaluated at 251, 255, and 263 minutes). However, the continuous nature of the choke change, associated with differences in sensor noise and resolution and with a total analysis time range much greater than the choke change duration, resulted in these different outcomes. This also shows that it is impracticable to expect a sensor changepoint for each consecutive choke change.

6.3 ICV Opening Procedures

In the following examples, two ICV opening procedures are analyzed, consisting of consecutive pressure cycles of the ICV command line. Although from different wells, both cases presented in this section had ICVs operated through discrete opening steps activated by the opening line. Each opening command was represented as a cause event. For both experiments, only the ICV pressure differential (DP-PDG, obtained by subtracting the downstream PDG measurement from the upstream PDG

Table 3 Cause–consequence matching results for the small choke change case.

Cause event		Consequence event	
Operational action	t_{cause} (min)	$t_{\text{consequence}}$ (min)	Sensor response
Matched cause–consequence pairs			
		Choke Upstream	
Decrease choke opening	255	254	Entered an increasing regime
Decrease choke opening	263	262	Accelerated the increasing regime
		PDG	
Decrease choke opening	255	254	Entered an increasing regime
		PT-P	
Decrease choke opening	263	263	Jump from 330.5 bar to 332.2 bar, followed by a stable regime
		TPT	
Decrease choke opening	251	253	Entered an increasing regime

measurement) was used as a sensor, as this is the most sensitive sensor for ICV changes, while other sensors from the WCT and at the topside may not detect some incremental changes. It is also important to note that both cases involved restriction changes, meaning that increasing the ICV opening leads to an expected increase in the flow rate, and consequently in the DP-PDG signal. Therefore, a missing response to an ICV change generates an alarm.

6.3.1 Successful ICV Opening

The first case is a normal ICV opening procedure without any unforeseen complications. Initially, in this case, the well had only one fully closed ICV, while the other ICVs were open. This means that the well was already producing, but with one totally closed ICV. The only commands executed during the procedure were pressurization pulses in this specific ICV command line for incremental opening. The ICV closing function was always at low pressure, and the rest of the subsea valves were not manipulated.

Figure 13 shows the normal ICV opening procedure data. In the first graph, the raw pressure-differential data and its segmentation, together with the detected regimes, are plotted. In the same chart, the ICV opening pressure cycles executed for the opening steps are shown as red vertical dash-dot lines. The last graph shows the choke data and its smoothed signal.

Table 4 presents the matches between cause events (ICV opening commands) and DP-PDG sensor changepoints. Most of the opening steps were matched with a pressure-differential changepoint. There was also no significant change detected in the choke valve opening, which remained fairly constant. Only two opening commands could not be matched to any sensor changepoint. Neither was associated with an operational problem in this case; therefore, the resulting alarms were false positives. The first command may produce no effective change in opening, while an additional opening near the end of the procedure may have little effect on flow, or the ICV may already be fully open. In either case, the pressure response may be too small to detect or simply be absent. This means that additional information about ICV status and opening values would have avoided these alarms.

6.3.2 Unsuccessful ICV Opening

The second case presents an ICV opening procedure during which operational difficulties were observed. Figure 14 presents the results following the same format as the previous figure, while Table 5 presents the matches between cause events (ICV opening steps) and DP-PDG sensor changepoints, but also the unmatched causes and consequence events.

As in the previous case, the well was already producing from other zones with one totally closed ICV, and then pressure pulses were applied to incrementally open this ICV. Although the subsea valves were also not manipulated, small changes were detected in the choke valve.

Several opening commands were applied to the ICV command line without producing any change in the pressure differential. Of the nine opening commands, only two were matched with DP-PDG

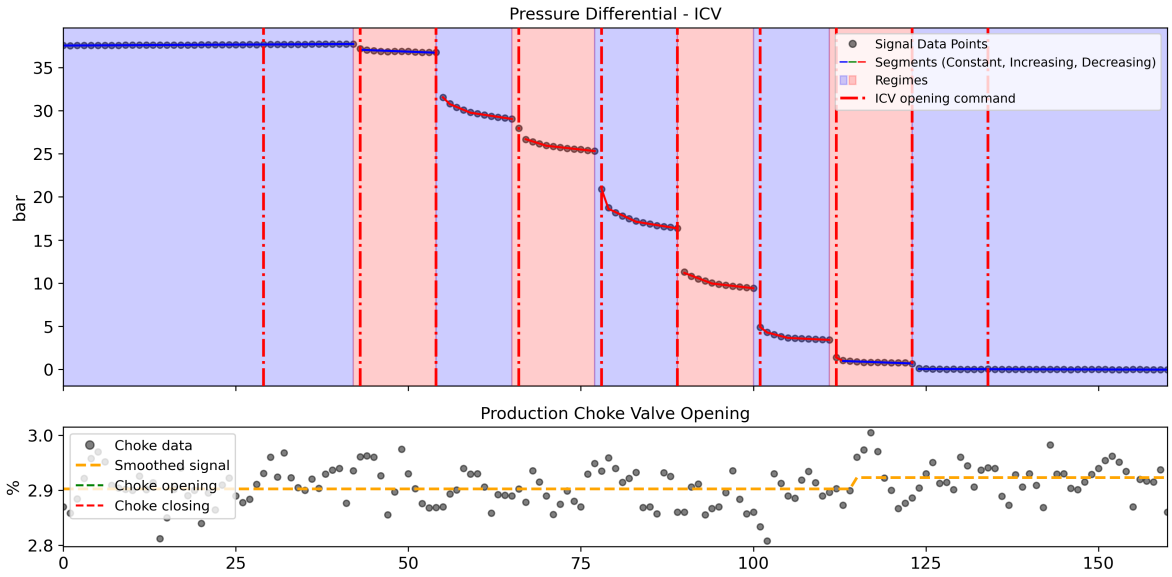

Fig. 13 ICV successful opening procedure case. In the upper graph, each vertical red dash-dot line shows a pressurization pulse to incrementally open the ICV. As can be seen, most of the pulses resulted in compatible pressure differential changes, which were also detected by the segmentation algorithm. The bottom chart shows the choke analysis, which returned no significant changes.

Table 4 Cause–consequence matching results for the successful ICV opening procedure case.

Cause event		**Consequence event**	
Operational action	t_{cause} (min)	$t_{\text{consequence}}$ (min)	Sensor response
Matched cause–consequence pairs			
			DP-PDG
Open ICV valve	43	42	Jump from 37.7 bar to 37.0 bar, followed by a stable regime
Open ICV valve	54	54	Jump from 36.7 bar to 31.5 bar, followed by a decreasing regime
Open ICV valve	66	65	Jump from 29.0 bar to 26.7 bar, followed by a decreasing regime
Open ICV valve	78	77	Jump from 25.3 bar to 20.9 bar, followed by a decreasing regime
Open ICV valve	89	89	Jump from 16.3 bar to 11.2 bar, followed by a decreasing regime
Open ICV valve	101	100	Jump from 9.4 bar to 4.9 bar, followed by a decreasing regime
Open ICV valve	112	111	Jump from 3.4 bar to 1.4 bar, followed by a decreasing regime
Open ICV valve	123	123	Jump from 0.7 bar to 0.1 bar, followed by a stable regime
Unmatched cause events			
Open ICV valve	29	–	No response from DP-PDG
Open ICV valve	134	–	No response from DP-PDG

changepoints, while the other seven were unmatched. Even with additional ICV information, most of the commands without a response would still remain unmatched.

There were also DP-PDG changepoints which could not be matched:

- A decreasing regime at 54 minutes, which happened 8 minutes after the first command and could not be matched to any cause due to the temporal filtering criterion. This was probably a delayed response of the opening mechanism after the first pressure cycle.
- An increase at 136 minutes, which did not match with any command and could indicate an unintended valve movement to its last position.

The differential pressure also changed at the 157-minute mark, but this change was matched to the choke shift represented in the bottom chart. The same behavior occurred at the 195-minute mark.

The comparison between the two cases, one involving a normal procedure without operational issues and the other involving a malfunctioning ICV, demonstrates that the proposed anomaly-detection framework can identify abnormal operational behavior when a system fault or technical issue occurs, while triggering zero (if full information about the ICV status is given) or few spurious but explainable alarms during normal operation.

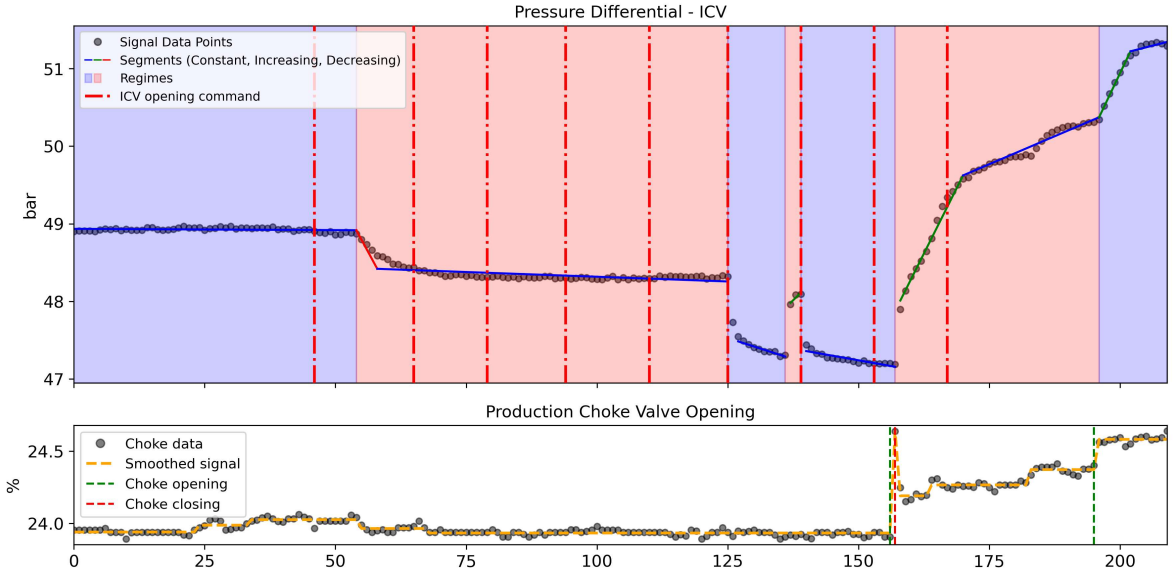

Fig. 14 ICV opening procedure with abnormalities. In this case, as can be seen in the upper chart, many unmatched cause and consequence events can be observed. Initially, the ICV pulse produces no response, which is only observed after almost 10 minutes, possibly indicating a delayed response. After that, multiple opening commands did not cause any expected change in the pressure differential measurement. At 136 minutes, an increasing response, not compatible with the ICV opening process, is also detected as an unmatched event. Finally, choke movements are detected, as seen in the bottom chart, and associated with the last changes observed in the differential pressure signal.

Table 5 Cause–consequence matching results for the abnormal ICV opening procedure case.

Cause event		**Consequence event**	
Operational action	t_{cause} (min)	$t_{\text{consequence}}$ (min)	Sensor response
Matched cause–consequence pairs			
		DP-PDG	
Open ICV valve	125	125	Jump from 48.3 bar to 47.5 bar, followed by a stable regime
Open ICV valve	139	139	Jump from 48.1 bar to 47.4 bar, followed by a stable regime
Increase choke opening	156	157	Jump from 47.2 bar to 48.0 bar, followed by an increasing regime
Increase choke opening	195	196	Entered an increasing regime
Unmatched consequence events			
		DP-PDG	
–	–	54	Entered a decreasing regime
–	–	136	Jump from 47.3 bar to 48.0 bar, followed by an increasing regime
Unmatched cause events			
Open ICV valve	46	–	No response from DP-PDG
Open ICV valve	65	–	No response from DP-PDG
Open ICV valve	79	–	No response from DP-PDG
Open ICV valve	94	–	No response from DP-PDG
Open ICV valve	110	–	No response from DP-PDG
Open ICV valve	153	–	No response from DP-PDG
Open ICV valve	167	–	No response from DP-PDG

6.4 Shutdown, Valve Tests and Well Restart

This case involves a shutdown followed by different valve maneuvers and, finally, a well restart, with the segmentation pipeline applied to the PDG, TPT, and PT-P pressure measurements. The well is shut-in first by closing the ESDV at 80 minutes, then the W1 at 81 minutes and then M1 at 82 minutes. The M1 valve is subsequently reopened at 104 minutes, after which the W1 valve is moved to different positions at different times while the ESDV remains closed. The well restart begins after the 300-minute mark, with the production choke being gradually opened. This is a complex case in which multiple valves (ESDV, M1, W1, and the production choke) are consecutively operated, while three pressure sensors located at different positions are analyzed.

The results can be seen in detail in Table 6, which shows the matched and unmatched causes and consequences, and in Figure 15, which visually presents the complete procedure. Out of 42 changepoints obtained from the final sets of all three sensors, only one was not matched. This was a PDG changepoint at the 396-minute mark, related to a small pressure recovery followed by stabilization. The last cause

Table 6 Cause–consequence matching results for the shutdown, test valves and restart case.

Cause event		Consequence event	
Operational action	t_{cause} (min)	$t_{\text{consequence}}$ (min)	Sensor response
Matched cause–consequence pairs			
			PDG
Close ESDV valve	80	83	Jump from 564.7 bar to 574.0 bar, followed by an increasing regime
Open W1 valve	107	107	Jump from 575.1 bar to 574.0 bar, followed by an increasing regime
Open W1 valve	154	153	Jump from 575.4 bar to 575.1 bar, followed by an increasing regime
Open W1 valve	219	218	Jump from 575.5 bar to 575.0 bar, followed by an increasing regime
Open W1 valve	287	286	Jump from 575.6 bar to 575.1 bar, followed by an increasing regime
Open ESDV valve	294	293	Jump from 575.6 bar to 574.0 bar, followed by an increasing regime
Increase choke opening	306	306	Jump from 575.6 bar to 575.1 bar, followed by an increasing regime
Increase choke opening	307	309	Changed from an increasing to a decreasing regime
Decrease choke opening	313	312	Changed from a decreasing to an increasing regime
Increase choke opening	372	372	Jump from 575.5 bar to 575.2 bar, followed by a stable regime
Increase choke opening	379	379	Jump from 575.1 bar to 574.4 bar, followed by a decreasing regime
Increase choke opening	386	387	Accelerated the decreasing regime
Increase choke opening	387	388	Jump from 573.8 bar to 572.9 bar, followed by a decreasing regime
Increase choke opening	393	393	Accelerated the decreasing regime
Increase choke opening	438	438	Entered a decreasing regime
Increase choke opening	439	439	Jump from 571.3 bar to 569.5 bar, followed by a decreasing regime
Increase choke opening	447	447	Jump from 568.2 bar to 566.9 bar, followed by a decreasing regime
Increase choke opening	451	451	Accelerated the decreasing regime
Increase choke opening	453	452	Jump from 565.6 bar to 563.1 bar, followed by a decreasing regime
Decrease choke opening	454	455	Changed from a decreasing to an increasing regime
Increase choke opening	459	459	Changed from an increasing to a decreasing regime
Increase choke opening	465	465	Jump from 562.8 bar to 560.2 bar, followed by a decreasing regime
			PT-P
Close ESDV valve	80	81	Jump from 301.3 bar to 323.0 bar, followed by a decreasing regime
Open W1 valve	107	107	Jump from 315.1 bar to 337.9 bar, followed by a stable regime
Close W1 valve	173	176	Entered a decreasing regime
Open W1 valve	219	218	Jump from 325.6 bar to 336.1 bar, followed by a decreasing regime
Open W1 valve	287	286	Jump from 320.2 bar to 336.1 bar, followed by a stable regime
Increase choke opening	372	376	Entered a decreasing regime
Increase choke opening	438	438	Entered a decreasing regime
Increase choke opening	447	447	Entered a decreasing regime
Decrease choke opening	454	455	Jump from 282.2 bar to 290.6 bar, followed by a decreasing regime
Increase choke opening	464	464	Accelerated the decreasing regime
			TPT
Close ESDV valve	80	81	Jump from 301.3 bar to 337.0 bar, followed by a stable regime
Close M1 valve	82	83	Jump from 337.0 bar to 309.1 bar, followed by a decreasing regime
Open M1 valve	104	104	Jump from 253.4 bar to 337.5 bar, followed by a stable regime
Increase choke opening	378	377	Entered a decreasing regime
Increase choke opening	438	438	Entered a decreasing regime
Increase choke opening	446	446	Accelerated the decreasing regime
Increase choke opening	451	452	Jump from 295.8 bar to 280.8 bar, followed by an increasing regime
Increase choke opening	453	456	Changed from an increasing to a decreasing regime
Increase choke opening	464	464	Accelerated the decreasing regime
Unmatched consequence events			
			PDG
–	–	396	Changed from a decreasing to an increasing regime

event before this changepoint was a choke opening, which was correctly assigned to an acceleration in the decreasing regime. This was a very brief episode, possibly associated with delayed dynamics, and the resulting changepoint might have been avoided by increasing the values of $\delta_{\min}$ and $\phi_{\lim}$.

A careful inspection of the cause–consequence matches in Table 6 shows that all matches were reasonable, indicating that the filters performed well in avoiding impossible associations. Some sensor responses, however, could be better explained by a combination of causes. For example, all initial sensor changepoints were related to the ESDV closure due to the physical compatibility preference over unknown responses. Yet, the initial response from PT-P is in fact related to both ESDV and W1 closure, as the ESDV closure caused the fast initial positive pressure jump, while the subsequent W1 closure was responsible for the decreasing regime due to the isolation of the flowline, which was exposed to a colder environment. This shows that a more accurate analysis should take into account the joint effects of multiple changes happening close to each other.

Nevertheless, this example shows that the framework can consistently be used for the analysis of complex operational procedures. Most of the relevant changes were detected and matched, and a structured summary of the entire period was generated.

6.5 Subsea Pipeline Partial Rupture Detection

In this last experiment, the segmentation pipeline alone was tested as a supporting tool for another real-time monitoring system. The data shown were entirely generated by a multiphase flow simulation and were previously presented in Barrionuevo et al. (2026), where a machine learning classifier was trained to detect leaks and seawater influx through subsea pipelines connected to oil wells. In all

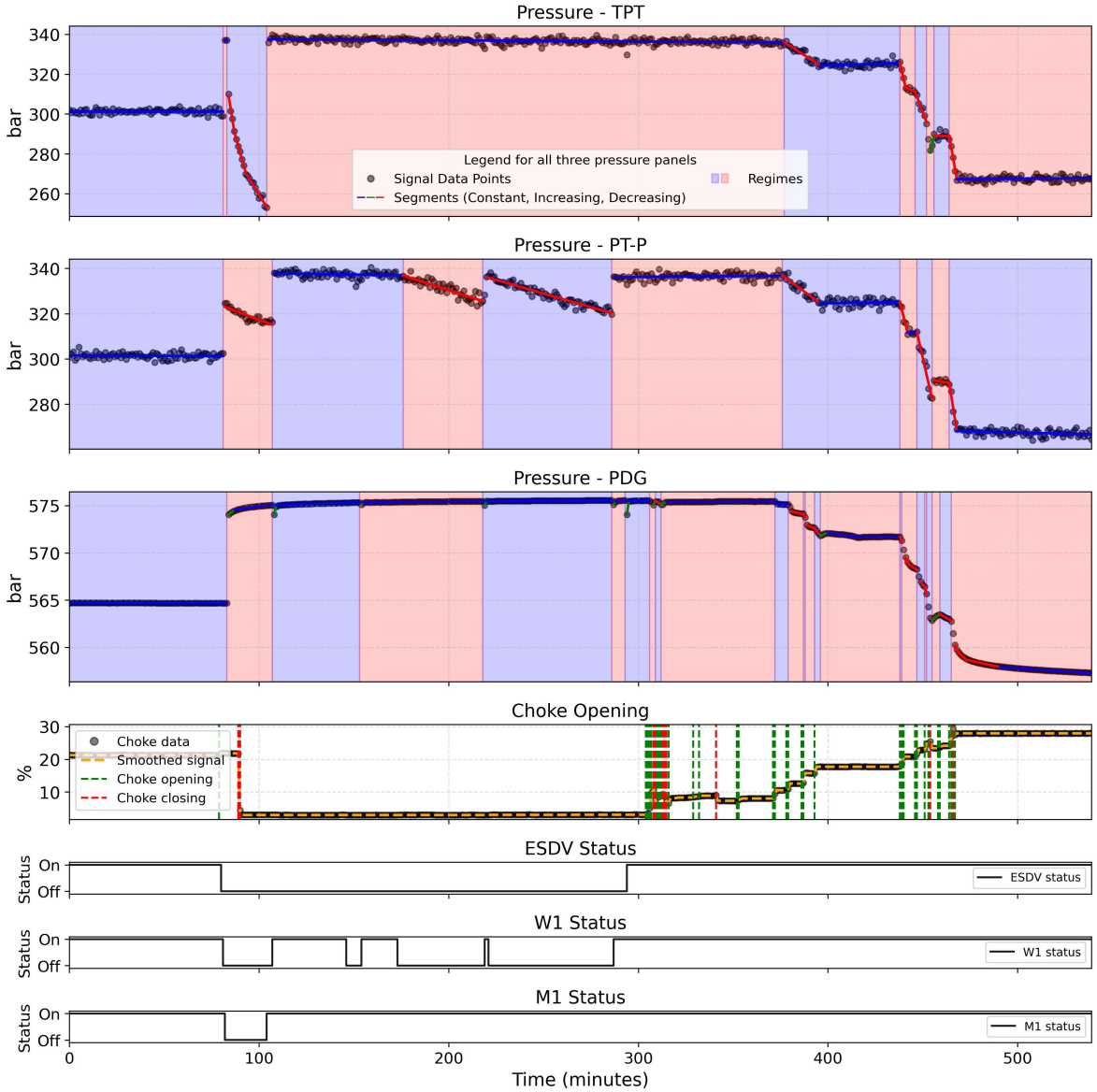

Fig. 15 Shutdown, valve tests, and well restart case. This case shows a shutdown, followed by multiple valve manipulations while the well is shut in, and finally a well restart. This case shows the main purpose of the developed tool, in which multiple operational commands and signals are cross-compared, a task that can be difficult and time-consuming for a production monitoring engineer.

cases, there were no valve changes, which means that any sensor changepoint is unmatched. This classifier used as input a relative deviation feature that compared the expected baseline values of four different sensor measurements — P-TPT (pressure at TPT), T-TPT (temperature at TPT), pressure at choke upstream, and temperature at choke upstream — with their current readings. This baseline comparison approach was adopted for several reasons discussed in Barrionuevo et al. (2026), including level-shift anomaly behavior, uncertain rupture formation velocity, rupture formation during shut-offs and ramp-ups, and extended detection windows. However, it also introduced a major weakness in the proposed methodology: the absence of a clear approach for determining the reference baseline value of each variable. Because of this issue, different approaches were tested in the field, such as manual input or a delayed moving window.

The segmentation pipeline proposed in this work, however, could provide a more suitable approach, as it can automatically identify different regimes within a time-series interval, thereby facilitating direct comparisons between them. In a scenario involving rupture formation during shut-offs and ramp-ups,

for example, steady-state regimes under the same choke conditions, separated by a transient phase, could be detected and compared.

For this purpose, the same LightGBM classifiers for leak and influx detection developed in Barrionuevo et al. (2026) were evaluated using two methodologies: the original approach — with a fixed reference value calculated using a 30-minute window for each sensor and a 10-minute moving average applied to the classifier output — and a new approach based on the segmentation framework proposed in this work.

In the new methodology used in this analysis, the segmentation pipeline is applied to each of the four input sensors. All changepoints from the four sensors are combined into a single set to divide the total time range into periods. A period that intersects at least three sensor segments with absolute slope $|\phi| < \phi_{\text{lim}}$ and contains no segment with slope $|\phi| > 3\phi_{\text{lim}}$ is treated as a baseline period. Its final x values are then used as baseline values and compared with the final values of subsequent periods until a new baseline period appears, which is also used as a new reference.

Note that this comparison between the base method from Barrionuevo et al. (2026), which uses a fixed time range reference, and the proposed pipeline segmentation method is not completely equivalent because:

- The base method already uses a fixed 30-minute window to calculate the reference value, while the proposed method receives no prior information about the data;
- The base method operates as an online system and is evaluated at every data point, while the segmentation approach compares only changepoint values between baseline periods;
- The segmentation method also depends on the correct localization of the changepoints relative to the fault occurrence, which means that it must correctly detect the moment of rupture formation;
- It is also important to note that, in these rupture simulations, some cases can present a fast initial transient period before stabilization, which can sometimes lead to false alarms or difficulties in detecting the correct baseline value. Because of that, the segmentation method could also detect spurious changepoints during transient periods before rupture formation.

This means that the segmentation approach is expected to produce worse results. What matters in this case is whether the classification performance remains good. In that case, a lower but still acceptable performance, combined with automatic reference-value estimation, may represent a reasonable trade-off.

Because the rupture data were entirely generated by simulation, noise was added to each sensor in order to make the data closer to reality. The noise was obtained from the steady-state case shown in Section 6.1 and added to each data point of each sensor as a random Gaussian value with zero mean. In this way, the steady-state case study serves as a complementary experiment to this one, showing that, under stable conditions, few spurious alarms are triggered.

Finally, although parts of the framework were not developed for temperature measurements (such as the regime-merging procedure and the matching heuristics), for this specific study, which aimed to analyze whether the methodology can identify stable regimes before and after an anomaly and estimate normal baseline values, these functionalities are less important. Similar parameter values from Table 2 were used for the temperature sensors: $\delta_{\text{min}} = 0.2$ °C and $\phi_{\text{lim}} = 1/30$ °C/min.

The results obtained for the test dataset using both methodologies can be seen in Table 7, where three different time ranges were tested: one window of 60 minutes, from $t = 10$ to 70 min; another of 60 minutes, from $t = 30$ to 90 min; and finally one of 120 minutes, from $t = 0$ to 120 min. The rupture formation always occurs at 60 minutes, which means that these studies analyzed detection rates at 10, 30, and 60 minutes after the fault. Surprisingly, for both classifiers, the detection rate increased in the first time range, which probably occurred because of the delaying effect caused by the 10-minute moving average applied to the classifier probability output. For the leak classifier, the segmentation approach increased the recall by approximately 12 percentage points in the first time range, while decreasing it by approximately 12–14 percentage points in the two longer ranges, still maintaining a performance close to 80%. For the influx classifier, the number of detected cases increased for the two 60-minute time ranges, while showing a slight decrease of almost 4 percentage points for the 120-minute window. Both approaches also presented very low rates of incorrect alarms and alarms before rupture.

The performance of both classifiers did not degrade too much even for longer time ranges, suggesting that the segmentation pipeline can be used as a baseline estimator for online monitoring systems.

Table 7 Post-rupture influx and leak predictions and pre-rupture alarms obtained using the known reference and segmentation approaches for the indicated analysis windows. Values are event counts followed by percentages of all n events in each actual class. Events with both post-rupture alarms, no post-rupture alarm, or no available evaluation are included in n but omitted from the displayed post-rupture outcomes. Pre-rupture counts represent events with at least one alarm from either detector before rupture.

		$t_0 = 10$ min, $\Delta t = 60$ min		$t_0 = 30$ min, $\Delta t = 60$ min		$t_0 = 0$min, $\Delta t = 120$ min	
Actual	Outcome	Known Reference	Segmentation	Known Reference	Segmentation	Known Reference	Segmentation
Influx ($n = 336$)	Post-rupture: Influx	5 (1.5%)	47 (14.0%)	271 (80.7%)	278 (82.7%)	318 (94.6%)	301 (89.6%)
	Post-rupture: Leak	0 (0.0%)	4 (1.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Pre-rupture: Any alarm	4 (1.2%)	7 (2.1%)	1 (0.3%)	2 (0.6%)	5 (1.5%)	5 (1.5%)
Leak ($n = 618$)	Post-rupture: Influx	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.2%)	0 (0.0%)	3 (0.5%)
	Post-rupture: Leak	362 (58.6%)	462 (74.8%)	567 (91.7%)	495 (80.1%)	578 (93.5%)	496 (80.3%)
	Pre-rupture: Any alarm	0 (0.0%)	1 (0.2%)	0 (0.0%)	4 (0.6%)	1 (0.2%)	4 (0.6%)

Table 8 Summary of the case studies used to demonstrate the proposed framework.

Case	**Purpose**	**Expected behavior**	**Observed result**
Steady-State Production	Check for spurious changepoints in baseline, steady-state regimes.	No regime changes.	The framework did not return any spurious changepoints for the most relevant well-production sensors (PDG, TPT, PT-P, and topside pressure). The PT-A pressure sensor returned one changepoint; however, this may be related to a minor disturbance in the very small annular bore chamber within the WCT.
Small Choke Change	Detect a small operational change.	Changes only in affected sensors.	Changes detected in TPT, PDG, and choke pressure; none in PT-A. The matching results for the choke opening varied across sensors, but all detected changes were correctly matched to at least one choke change.
Normal ICV opening	Evaluate normal command–response matching.	Expected commands followed by pressure changes.	Most ICV opening commands were correctly matched with pressure-differential responses. Two unmatched commands generated false-positive alarms that could be avoided with additional well-specific information.
Abnormal ICV opening	Identify abnormal command–response behavior.	Missing, delayed, or unexplained responses.	Several unmatched causes and consequences were identified in a case involving operational difficulties during ICV opening, demonstrating the framework's effectiveness in detecting abnormalities.
Shutdown, Valve Tests and Restart Case	Analyze how the system performs under multiple consecutive operational changes and with multiple sensors.	All relevant changes in the three pressure sensors detected and matched to an operational change.	Almost all regime changes were correctly identified and assigned to a cause. Only one PDG changepoint associated with a delayed reservoir response was unmatched out of a total of 42 sensor changepoints. The matches were analyzed and no incoherent or impossible relationship was detected, although some responses would be better explained by considering a sequence of cause events rather than a single cause.
Subsea Pipeline Partial Rupture	Assess if the framework can detect separate regimes before and after an event and estimate the correct reference value for each sensor.	Leak and influx detection performance slightly affected compared to base case using known baseline value.	For longer periods after the fault, the leak classifier lost performance but still maintained good results (recall around 80%), while the influx classifier maintained most of its performance. For shorter periods after the fault, the performance of both classifiers actually improved. The number of spurious alarms increased slightly but remained very low.

7 Conclusions

A retrospective well operational analysis tool, consisting of a segmentation pipeline combined with a cause–consequence matching algorithm, was proposed for the monitoring and interpretation of well operational data. A piecewise-linear segmentation pipeline built around CPOP was developed to account for operationally significant signal variations and merge consecutive linear segments representing naturally curved transient regimes. A cause–consequence matching method was subsequently incorporated to associate operational commands with sensor responses using temporal, topological, and physical compatibility criteria. The final tool uses a small number of simple, operator-intuitive user inputs, and its formulation is based on clear, transparent rules, which facilitates its adoption.

The main outcomes of the case studies are summarized in Table 8. Overall, the case studies demonstrate the applicability of the proposed framework under different operating conditions, ranging from stable production and small operational changes to complex valve procedures and fault scenarios. The results show that the segmentation pipeline can identify relevant regime changes while avoiding unnecessary changepoints in stable conditions, and that the cause–consequence matching approach can highlight missing, delayed, or unexplained sensor responses during operational procedures.

Future work should address current limitations, improve existing functionalities, and explore further applications. The main points for improvement are:

- **Regime merging:** the current proposal relies heavily on heuristics and uses a user-defined value. Future work should investigate more general and less parameter-dependent methods.
- **Cause–consequence matching:** the cases presented in this work considered limited subsystems, mainly the production path from the reservoir to the production choke. The framework could be expanded to the entire well system, including the annulus and service lines. For the annulus, for example, cause–consequence matching should also consider effects due to heat exchange with the tubing or due to pressurization of the tubing (ballooning effect);
- **Physical matching:** the current implementation relies on physics-based rules, which can be difficult to define for all possible operational configurations and scenarios. Future work could investigate the use of digital twins to automatically estimate the expected physical response to operational actions and support the cause–consequence matching process;
- **Broader application:** future work could evaluate the framework across additional wells, operational configurations, and production scenarios to further assess its applicability and generalization.
- **Integration with AI systems:** As was shown, regime properties and cause–consequence matches could be used as inputs to specialized monitoring, alarm, and decision-support systems. Furthermore, the framework seems especially suitable for integration with AI reasoning models because it transforms multiple raw time series into structured summaries of a selected time interval. A system that could further interpret well activity, make suggestions, and generate reports would be a very useful tool for monitoring.

A Least Squares with fixed knots

Let the set of known changepoints be

$$\tau_1, \tau_2, \ldots, \tau_K,$$

with the endpoint knots defined as

$$\tau_0 = 0, \qquad \tau_{K+1} = n - 1.$$

The objective is to estimate the fitted values at the knots,

$$\hat{\mathbf{x}} = [x_0, x_{\tau_1}, x_{\tau_2}, x_{\tau_3}, \ldots, x_{\tau_K}, x_{n-1}]^T$$

which will generate the following function for each timestep t between changepoints τ_i and τ_{i+1}:

$$x_t = \frac{\tau_{i+1} - t}{\tau_{i+1} - \tau_i} x_i + \frac{t - \tau_i}{\tau_{i+1} - \tau_i} x_{i+1} \tag{18}$$

by minimizing the least squares term:

$$\min_{x_0, x_{\tau_1}, .., x_{n-1}} \sum_{i=0}^{K} \sum_{t=\tau_i}^{\tau_{i+1}} \left(y_t - \frac{\tau_{i+1} - t}{\tau_{i+1} - \tau_i} x_i - \frac{t - \tau_i}{\tau_{i+1} - \tau_i} x_{i+1} \right)^2 \tag{19}$$

It is possible to rewrite the function x_t as:

$$x_t = \sum_{i=0}^{K+1} \beta_{t,i} x_i, \tag{20}$$

where each $\beta_{t,i}$ is a function in t related to each knot i. For the first knot $i = 0$:

$$\beta_{t,i} = \begin{cases} \frac{\tau_1 - t}{\tau_1}, & 0 \leq t \leq \tau_1 \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

For the middle knots $i = 1, 2, \quad , K$:

$$\beta_{t,i} = \begin{cases} \frac{t-\tau_{i-1}}{\tau_i - \tau_{i-1}}, & \tau_{i-1} \leq t \leq \tau_i \\ \frac{\tau_{i+1}-t}{\tau_{i+1}-\tau_i}, & \tau_i \leq t \leq \tau_{i+1} \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

For the last knot $i = K + 1$:

$$\beta_{t,i} = \begin{cases} \frac{t-\tau_K}{n-1-\tau_K}, & \tau_K \leq t \leq n-1 \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The following relationship for any $\beta_{t,i}$ also exists:

$$\beta_{t,i} + \beta_{t,i+1} = 1, \quad \text{for} \quad \tau_i \leq t \leq \tau_{i+1} \tag{24}$$

By applying these transformations, the least squares matrix solution becomes:

$$\hat{\boldsymbol{x}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{y} \tag{25}$$

where the $\mathbf{B} \in \mathbb{R}^{n\times(K+2)}$ matrix is equal to:

$$\mathbf{B} = \begin{bmatrix} \beta_{0,0} & 1-\beta_{0,0} & 0 & \cdots & 0 \\ \beta_{1,0} & 1-\beta_{1,0} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \beta_{\tau_1-1,0} & 1-\beta_{\tau_1-1,0} & 0 & \cdots & 0 \\ 0 & \beta_{\tau_1,1} & 1-\beta_{\tau_1,1} & \cdots & 0 \\ 0 & \beta_{\tau_1+1,1} & 1-\beta_{\tau_1+1,1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta_{n-1,K} & 1-\beta_{n-1,K} \end{bmatrix}. \tag{26}$$

Additionally, because the $\mathbf{B}^T\mathbf{B}$ matrix is always tridiagonal, the solution can be obtained by Thomas algorithm in linear time O(n)(Ford 2015).

B Order Preserving Bipartite Matching Algorithm

The classical bipartite matching problem considers a graph $G = (A, B, E)$ (Korte and Vygen 2018), where:

- A and B are two disjoint sets of vertices with cardinalities $|A| = n$ and $|B| = m$,
- E is a set of edges, where each edge necessarily connects a vertex of A to a vertex of B

A matching M is a subset of E ($M \subseteq E$) in which no two edges share a common vertex. A vertex with no incident edge in M is defined as unmatched. Figure 16 illustrates an example of a graph G with two vertex sets A and B, an edge set E, and a matching M.

Usual problems arising from this formulation are:

- **Maximum cardinality matching problem**: find the maximum size matching $K = |M|$, i.e., the matching with the most possible number of edges,
- **Minimum weight matching problem**: minimize the matching cost $C = \sum_{(i,j)\in M} c_{ij}$, where c_{ij} is a given cost for the edge e_{ij} connecting a_i and b_j.
- **Maximum cardinality, minimum cost matching problem**: combination of both problems, where first all possible maximum cardinality solutions are found, and from these the one with minimum cost is selected.

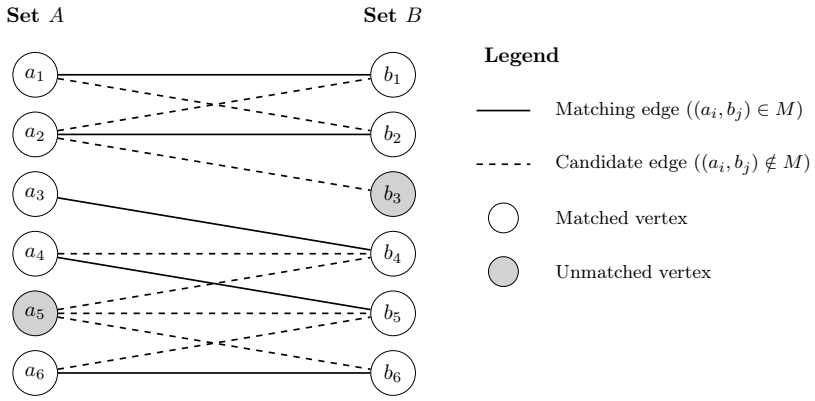

Fig. 16 Illustration of a bipartite graph $G = (A, B, E)$, where A is the left vertex set, B is the right vertex set, and E is the set of edges connecting vertices from the two sets. A matching $M \subseteq E$ is shown. Edges belonging to the matching are represented by solid thick lines, while candidate edges not included in M are represented by dashed lines. Unmatched vertices are shown in gray.

In this current work, however, the matching should be **order-preserving**, which means it should have no crossing edges in M. In mathematical terms, for every pair of edges e_{ij} and $e_{i'j'}$ in M, $i < i'$ and $j < j'$ (or $i > i'$ and $j > j'$) must be true at the same time. The matching shown in Figure 16, for example, is order-preserving.

Another difference is that some edges in E are preferred over others. In the proposed problem, these correspond to physically compatible cause–consequence pairs. For each edge e_{ij}, a binary indicator p_{ij} is assigned, equal to 1 if the edge is preferred and 0 otherwise.

In order to perform order-preserving matching with the following objectives, in order of priority:

- Maximize the number of preferred matched edges,
- Maximize the total number of matched edges,
- Minimize the total cost of the selected edges,

a dynamic programming algorithm was implemented, adapting the ordered bipartite matching approach proposed by Kim et al. (2011), as shown in Algorithm 2. The algorithm defines a $(n+1) \times (m+1)$ matrix D, where each entry $D[i, j] = (P_{ij}, K_{ij}, C_{ij})$ is a tuple. Here, P_{ij} represents the maximum number of preferred matched edges using only the first i vertices of A and the first j vertices of B. Similarly, K_{ij} represents the maximum total number of matched edges among matchings with this maximum preference count, while C_{ij} represents the minimum total cost among matchings with the maximum preference count and cardinality.

During the calculation of the D matrix, three options are evaluated at each step: (1) skip a_i, (2) skip b_j, or (3) match a_i with b_j, if $(a_i, b_j) \in E$. The candidate values of $D[i, j]$ corresponding to these three options are calculated, and the algorithm selects the best one, prioritizing maximum cardinality K_{ij}, followed by minimum total cost C_{ij} as a tie-breaker. After calculating the D matrix, a backtracking procedure finds the optimal matching by starting at the last entry $D[n, m]$ and following the best decisions recorded in a separate matrix P.

Acknowledgements. The authors thank Petróleo Brasileiro S.A. (Petrobras) for providing the operational well data used in this study.

Statements and Declarations

Funding. This research was funded by Petróleo Brasileiro S.A. (Petrobras) and supported by the BRU21 research and innovation program at the Norwegian University of Science and Technology (NTNU), Digital and Automation Solutions for the Offshore Energy Industry.

Author contributions. Daniel Centurion Barrionuevo: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Validation, Writing – original draft, Writing – review and editing.

Alexey Pavlov: Supervision, Writing – review and editing.

Algorithm 2 Order-preserving bipartite matching

Require: Ordered vertex sets $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_m)$, edge set E, costs c_{ij}, preferences p_{ij}
Ensure: Matching $M \subseteq E$
The function $\textsc{best}(D_1, D_2, D_3)$ *selects the lexicographic maximum of* $(P_k, K_k, -C_k)$*, where* $D_k = (P_k, K_k, C_k)$*, breaking ties by the smallest* k*.*

$D[i, 0] \leftarrow (0, 0, 0)$ for $i = 0, \ldots, n$
$D[0, j] \leftarrow (0, 0, 0)$ for $j = 0, \ldots, m$
for $i = 1, \ldots, n$ **do**
  for $j = 1, \ldots, m$ **do**
    $D_1 \leftarrow D[i-1, j]$
    $D_2 \leftarrow D[i, j-1]$
    if $(a_i, b_j) \in E$ **then**
      $D_3 \leftarrow D[i-1, j-1] + (p_{ij}, 1, c_{ij})$
    else
      $D_3 \leftarrow (-\infty, -\infty, \infty)$
    end if
    $k^* \leftarrow \textsc{best}(D_1, D_2, D_3)$
    $D[i, j] \leftarrow D_{k^*}$
    $Q[i, j] \leftarrow k^*$
  end for
end for
$M \leftarrow \emptyset$
$i \leftarrow n$, $j \leftarrow m$
while $i > 0$ **and** $j > 0$ **do**
  if $Q[i, j] = 1$ **then**
    $i \leftarrow i - 1$
  else if $Q[i, j] = 2$ **then**
    $j \leftarrow j - 1$
  else
    $M \leftarrow M \cup \{(a_i, b_j)\}$
    $i \leftarrow i - 1$, $j \leftarrow j - 1$
  end if
end while
return M

Competing interests. Daniel Centurion Barrionuevo is employed by Petróleo Brasileiro S.A. (Petrobras), which funded the research and provided the operational data used in this study. Alexey Pavlov declares no competing interests.

Data availability. The operational well data analyzed in this study are available from the corresponding author upon reasonable request.

Use of generative AI and AI-assisted technologies. During the preparation of this work, the authors used ChatGPT (OpenAI) to assist with English-language editing and code development. All AI-assisted code was reviewed and validated by the authors, who take full responsibility for the methodology, results, and content of the publication.

References

Aljaroudi, A., Khan, F., Akinturk, A., Haddara, M., Thodi, P.: Risk assessment of offshore crude oil pipeline failure. Journal of Loss Prevention in the Process Industries **37**, 101–109 (2015) https://doi.org/10.1016/j.jlp.2015.07.004

Aranha, P.E., Lopes, L.G.O., Paranhos Sobrinho, E.S., Oliveira, I.M.N., Araújo, J.P.N., Santos, B.B., Lima Junior, E.T., Silva, T.B., Vieira, T.M.A., Lira, W.W.M., Policarpo, N.A., Sampaio, M.A.: A system to detect oilwell anomalies using deep learning and decision diagram dual approach. SPE Journal **29**(03), 1540–1553 (2024) https://doi.org/10.2118/218017-PA https://onepetro.org/SJ/article-pdf/29/03/1540/3376130/spe-218017-pa.pdf

Aranha, P.E., Policarpo, N.A., Sampaio, M.A.: Unsupervised machine learning model for predicting anomalies in subsurface safety valves and application in offshore wells during oil production. Journal of Petroleum Exploration and Production Technology **14**(2), 567–581 (2024) https://doi.org/10.1007/s13202-023-01720-4

Agostini, C., Yamamoto, M., Miura, K.: The application and functionalities of a wet christmas tree applied in santos basin pre-salt cluster. Conference Proceedings The Japan Society of Naval Architects and Ocean Engineers **25**, 77–80 (2017) https://doi.org/10.14856/conf.25.0_77

Bayazitova, G., Anastasiadou, M., dos Santos, V.D.: Oil and gas flow anomaly detection on offshore naturally flowing wells using deep neural networks. Geoenergy Science and Engineering **242**, 213240 (2024) https://doi.org/10.1016/j.geoen.2024.213240

Barrionuevo, D.C., Oliveira, G.H., Faller, A.C., Vieira, S.C., Assmann, B.W., Araújo, J.C.D., Pavlov, A.: Detection of oil leak and seawater influx through partial ruptures in subsea petroleum production pipelines. Journal of the Brazilian Society of Mechanical Sciences and Engineering **48**(7), 504 (2026) https://doi.org/10.1007/s40430-026-06488-2

Duarte, B.P.M., Saraiva, P.M.: Change point detection for quality monitoring of chemical processes. In: Kraslawski, A., Turunen, I. (eds.) European Symposium on Computer Aided Process Engineering-13. Computer Aided Chemical Engineering, vol. 14, pp. 401–406. Elsevier, Amsterdam (2003). https://doi.org/10.1016/S1570-7946(03)80148-7

Fearnhead, P., Grose, D.: cpop: Detecting changes in piecewise-linear signals (2022). https://arxiv.org/abs/2208.11009

Fernandes, W., Komati, K.S., Souza Gazolli, K.: Anomaly detection in oil-producing wells: a comparative study of one-class classifiers in a multivariate time series dataset. Journal of Petroleum Exploration and Production Technology **14**(1), 343–363 (2024) https://doi.org/10.1007/s13202-023-01710-6

Fearnhead, P., Maidstone, R., Letchford, A.: Detecting changes in slope with an l0 penalty. Journal of Computational and Graphical Statistics **28**(2), 265–275 (2019) https://doi.org/10.1080/10618600.2018.1512868 https://doi.org/10.1080/10618600.2018.1512868

Ford, W.: Algorithms. In: Numerical Linear Algebra with Applications, pp. 163–179. Academic Press, London (2015). Chap. 9. https://doi.org/10.1016/B978-0-12-394435-1.00009-0 . https://doi.org/10.1016/B978-0-12-394435-1.00009-0

Gedda, R., Beilina, L., Tan, R.: Change point detection for process data analytics applied to a multiphase flow facility. CMES - Computer Modeling in Engineering and Sciences **134**(3), 1737–1759 (2022) https://doi.org/10.32604/cmes.2022.019764

Guo, S., Guo, W.G., Abolhassani, A., Kalamdani, R., Puchala, S., Januszczak, A., Jalluri, C.: Manufacturing process monitoring with nonparametric change-point detection in automotive industry. Journal of Manufacturing Science and Engineering **141**(7), 071013 (2019) https://doi.org/10.1115/1.4043732

Gonzalez, D., Gonzalez, F., Pietrobon, M., Haghshenas, M., Shurn, M., Mees, A., Stewart, C., Ogugbue, C., Duvivier, G.: Strategies to monitor and mitigate asphaltene issues in the production system of a gulf of mexico deepwater subsea development. OTC Offshore Technology Conference, vol. Offshore Technology Conference, pp. 011–011002 (2016). https://doi.org/10.4043/27123-MS . https://doi.org/10.4043/27123-MS

Garcia, E., Mathew, M.: Delivering selective interval control: Electro-hydraulic intelligent completion enhances reservoir management for multilateral well. SPE/IADC Middle East Drilling Technology Conference and Exhibition, vol. SPE/IADC Middle East Drilling Technology Conference and Exhibition, pp. 023–008001 (2016). https://doi.org/10.2118/178168-MS . https://doi.org/10.2118/178168-MS

Grantham, M., Shi, X., Clarke, B.: Changepoint Detection As Model Selection: A General Framework (2026). https://arxiv.org/abs/2601.22481

Haynes, K., Eckley, I.A., Fearnhead, P.: Computationally efficient changepoint detection for a range of penalties. Journal of Computational and Graphical Statistics **26**(1), 134–143 (2017) https://doi.org/10.1080/10618600.2015.1116445 https://doi.org/10.1080/10618600.2015.1116445

Hamedi Shokrlu, Y., Bazile, J.: Improving esp production for unconventional wells through real-time machine learning based changepoint detection. SPE Western Regional Meeting, vol. SPE Western Regional Meeting, pp. 011–004001 (2024). https://doi.org/10.2118/218873-MS . https://doi.org/10.2118/218873-MS

Hu, W., Shah, S.L., Chen, T.: Framework for a smart data analytics platform towards process monitoring and alarm management. Computers & Chemical Engineering **114**, 225–244 (2018) https://doi.org/10.1016/j.compchemeng.2017.10.010 . FOCAPO/CPC 2017

Joubran, J.: Intelligent completions: Design and reliability of interval control valves in the past, present, and future. OTC Offshore Technology Conference, vol. Offshore Technology Conference, pp. 041–047001 (2018). https://doi.org/10.4043/28917-MS . https://doi.org/10.4043/28917-MS

Kim, J., Peng, Y., Ivezic, N., Shin, J.: An optimization approach for semantic-based xml schema matching. International Journal of Trade, Economics and Finance **2**(1), 78–84 (2011)

Korte, B., Vygen, J.: Combinatorial Optimization: Theory and Algorithms, 6th edn. Algorithms and Combinatorics, vol. 21. Springer, Berlin, Heidelberg (2018). https://doi.org/10.1007/978-3-662-56039-6

Lopes, L.G.O., Almeida Vieira, T.M., Aranha, P.E., Lima Junior, E.T., Lira, W.W.M.: Detection and classification of anomalies in oil well production using open-world learning. Engineering Applications of Artificial Intelligence **159**, 111514 (2025) https://doi.org/10.1016/j.engappai.2025.111514

Lopes, L.G.O., Vieira, T.M.A., Oliveira, I.M.N., Silva, A.C.A., Lima Junior, E.T., Lira, W.W.M., Vargas, R.E.V., Aranha, P.E.: Real-time recognition of operational events in oil wells using u-net segmentation and distance-based classification. OTC Offshore Technology Conference, vol. Offshore Technology Conference, pp. 041–056001 (2026). https://doi.org/10.4043/36929-MS . https://doi.org/10.4043/36929-MS

Makogon, Y.F.: Formation of hydrates in shut-down pipelines in offshore conditions. OTC Offshore Technology Conference, vol. Offshore Technology Conference, p. 8235 (1996). https://doi.org/10.4043/8235-MS . https://doi.org/10.4043/8235-MS

Marins, M.A., Barros, B.D., Santos, I.H., Barrionuevo, D.C., Vargas, R.E.V., de M. Prego, T., de Lima, A.A., de Campos, M.L.R., da Silva, E.A.B., Netto, S.L.: Fault detection and classification in oil wells and production/service lines using random forest. Journal of Petroleum Science and Engineering **197**, 107879 (2021) https://doi.org/10.1016/j.petrol.2020.107879

Rodrigues, D.N., Minassa, L.B., Nunes, R.S., Costa, L.C.P.: Intelligent completions as a key technology in brazilian presalt. Offshore Technology Conference Brasil, vol. OTC Brasil, pp. 011–007005 (2017). https://doi.org/10.4043/28056-MS . https://doi.org/10.4043/28056-MS

Schwarz, G.: Estimating the dimension of a model. The Annals of Statistics **6**(2), 461–464 (1978) https://doi.org/10.1214/aos/1176344136

Silva, A.C.A.d., Oliveira, I.d.M.N., Lopes, L.G.O., Lima Junior, E.T.d., Vieira, T.M.d.A., Aranha, P.E.: Automatic detection and classification of relevant events in oil wells using numerical derivatives and rule-based segmentation. Brazilian Journal of Petroleum and Gas **20**(1), 33–40 (2026) https://doi.org/10.5419/bjpg2026-0004

Sudyoprasert, K., Pang, D., Oentoro, N., Valena, M., Rahimli, J., Ashok, P., Oort, E.v., Dahlem, J.: Machine learning for smarter well integrity management. SPE/ICoTA Well Intervention Conference and Exhibition, vol. SPE/ICoTA Well Intervention Conference and Exhibition, pp. 011–007004 (2026). https://doi.org/10.2118/231400-MS . https://doi.org/10.2118/231400-MS

Torbergsen, H.-E.B., Haga, H.B., Sangesland, S., Aadnøy, B.S., Sæby, J., Johnsen, S., Rausand, M., Lundteigen, M.A.: An introduction to well integrity. Technical Report Revision 0, Norwegian Oil and

Gas Association, Well Integrity Forum (December 2012). https://www.offshorenorge.no/contentassets/c5b69b8714984f50927b6142d9ea60de/introduction-to-well-integrity---04-december-2012.pdf Accessed 2026-08-04

Turan, E.M., Jäschke, J.: Classification of undesirable events in oil well operation. In: 2021 23rd International Conference on Process Control (PC), pp. 157–162 (2021). https://doi.org/10.1109/PC52310.2021.9447527

Truong, C., Oudre, L., Vayatis, N.: Selective review of offline change point detection methods. Signal Process. **167**(C) (2020) https://doi.org/10.1016/j.sigpro.2019.107299

Vignes, B., Aadnøy, B.S.: Well-integrity issues offshore norway. SPE Production & Operations **25**(2), 145–150 (2010) https://doi.org/10.2118/112535-PA

Vargas, R.E.V., Munaro, C.J., Ciarelli, P.M., Medeiros, A.G., Amaral, B.G., Barrionuevo, D.C., Araújo, J.C.D., Ribeiro, J.L., Magalhães, L.P.: A realistic and public dataset with rare undesirable real events in oil wells. Journal of Petroleum Science and Engineering **181**, 106223 (2019) https://doi.org/10.1016/j.petrol.2019.106223

Venkatesan, R., Nagarajan, N.R., Paso, K., Yi, Y.-B., Sastry, A.M., Fogler, H.S.: The strength of paraffin gels formed under static and flow conditions. Chemical Engineering Science **60**(13), 3587–3598 (2005) https://doi.org/10.1016/j.ces.2005.02.045

Wang, M., Su, X., Song, H., Wang, Y., Yang, X.: Enhancing predictive maintenance strategies for oil and gas equipment through ensemble learning modeling. Journal of Petroleum Exploration and Production Technology **15**(3), 46 (2025) https://doi.org/10.1007/s13202-025-01931-x