Petroleum Science 23 (2026) 5636–5647

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Physics-informed machine learning for sustained casing pressure diagnostics and root cause analysis for well integrity

Ahmed Alsubaih [a,*], Evan Wetmore [a], Watheq J. Al-Mudhafar [b], Kamy Sepehrnoori [a], Alberto L. Manriquez [a], Mojdeh Delshad [a]

[a] *Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, 78712, TX, USA*
[b] *Basrah Oil Company, Basrah, 61001, Iraq*

A R T I C L E I N F O

Article history:
Received 20 August 2025
Received in revised form
3 February 2026
Accepted 23 March 2026
Available online 26 March 2026

Edited by Xi Zhang

Keywords:
Sustained casing pressure (SCP)
Well integrity
Physics-informed informed machine learning (PIML)
Annular pressure
SCADA monitoring

A B S T R A C T

Sustained casing pressure (SCP) is a primary indicator of well integrity degradation, arising from compromised barriers such as cement, casing, tubing/packer components, or wellhead seals. Traditional SCP diagnostics—based on manual interpretation of annular pressure trends, bleed-off tests, and operational records—are often time-consuming, analyst-dependent, and difficult to scale across large well populations. This study presents a physics-informed machine learning (PIML) framework that integrates engineering-based physical principles, including fluid compressibility, thermal expansion, and leak-path mechanics, into a machine learning workflow for automated and scalable SCP classification. The framework is applied to field monitoring data from 26 wells, with detailed multi-annulus case studies for wells A5, A8, and A12. The method classifies cycle-level SCP behavior into six diagnostic types—no pressure, thermal pressure, trapped pressure, recharge pressure, constant pressure, and high-rate recharge—and maintains physically plausible and interpretable outputs under noisy or complex pressure signatures. Operationally, the framework supports real-time, SCADA-integrated surveillance to enable early detection of integrity threats, prioritization of higher-risk wells, and proactive intervention planning. The novelty of this work lies in embedding physical constraints directly into the classification logic, producing a robust and interpretable diagnostic capability that advances SCP analysis from a reactive, manual task toward automated well-integrity surveillance applicable to offshore, HPHT, CO_2 sequestration, and hydrogen storage operations.

1. Introduction

Sustained casing pressure (SCP) is a primary well integrity indicator caused by the failure of well components, including tubulars, cement sheaths, completion tools, and the Christmas tree, to maintain fluid containment within the reservoir along the well trajectory (Ibukun et al., 2024). Failure to contain fluids can lead to catastrophic consequences, including uncontrolled flow, environmental contamination, and premature well abandonment (Alsubaih et al., 2025a, b). SCP is defined as annular pressure that rebuilds after bleed-down, indicating active hydraulic communication with a higher-pressure source (Alam and Minor, 2024). Field statistics highlight its prevalence: SCP affects more than 45% of wells in the Gulf of Mexico and approximately 35% of wells in the UK North Sea (Ibukun et al., 2024). Far from being a transient anomaly, SCP often reflects long-term degradation of wellbore barriers or the presence of an active leak path. Wells are constructed with multiple concentric casing strings as shown in Fig. 1, forming distinct annular spaces, A-annulus (between production tubing and production casing), B-annulus (between production and intermediate casing), and C-annulus (between intermediate and surface or conductor casing).

SCP can result from tubing leaks, packer failures, casing breaches, or cement degradation, each producing distinctive pressure signatures and risk profiles (Alam and Minor, 2024; Kinik, 2012). Perhaps most concerning are cement failures, which can occur through micro-annuli formation, chemical degradation, or mechanical debonding, potentially allowing formation fluids to migrate vertically along the wellbore (Rusch et al., 2004). During

* Corresponding author.
E-mail address: aaa8649@my.utexas.edu (A. Alsubaih).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.03.049
1995-8226/

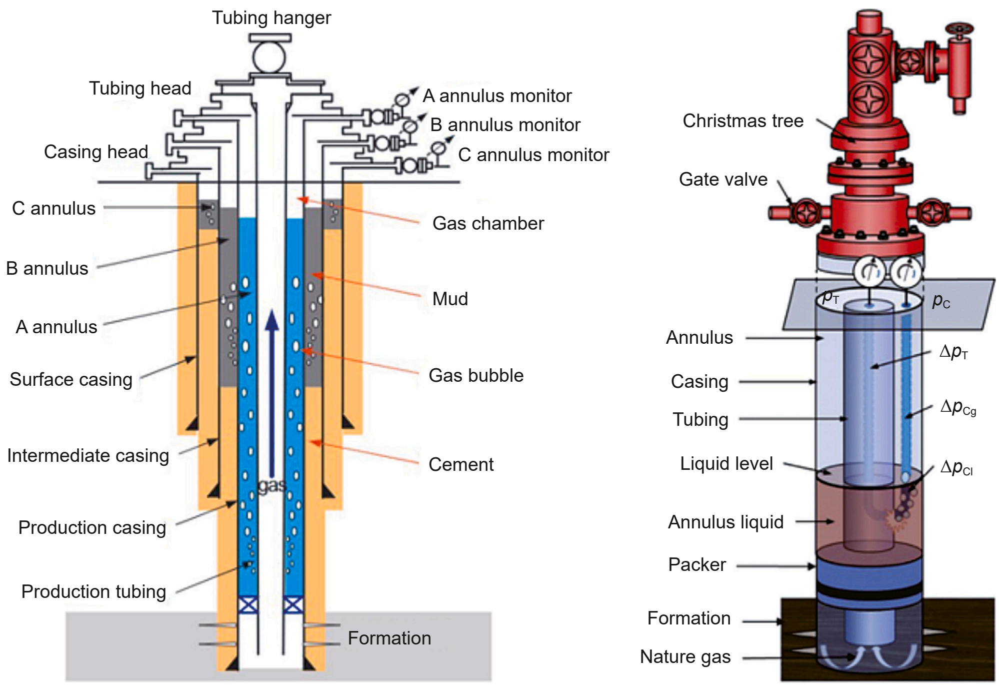

Fig. 1. Combined schematic showing the structure and pressure dynamics in a typical oil or gas well. Left: Vertical cross-section of a well with multiple concentric casing strings and annuli (A, B, C), showing annulus monitors, potential gas migration, and fluid interfaces (modified from Zhu et al., 2012a, b).

normal operations, annular pressures are expected to remain low or stable, with some systems open to atmosphere or sealed with inert fluids. Any abnormal or unplanned pressure increase, especially one that rebuilds after venting, signals integrity loss and should trigger detailed diagnostic investigation. The diagnostic challenge presented by SCP extends far beyond simple pressure measurement and interpretation. The complexity of SCP diagnosis is further compounded by the dynamic nature of well operations and the multitude of variables that can influence annular pressure behavior. Temperature fluctuations during production startup and shutdown can create thermal expansion effects that mimic or mask SCP signatures (Kazemi and Wojtanowicz, 2022). Historically, SCP diagnosis has relied heavily on post-bleed-off diagnostic process of annular pressure trends, bleed-down testing procedures, and the subjective judgment of experienced engineers (Kazemi and Wojtanowicz, 2022). While these traditional approaches have served the industry for decades, they suffer from limitations that become increasingly problematic as well populations grow and operational complexity increases, including interpretation non-uniqueness, inconsistent analyst-to-analyst conclusions, and limited scalability for high-frequency surveillance data.

Recent literature highlights a broader transition from manual, single-sensor interpretation toward integrated diagnostic workflows that combine continuous pressure monitoring with multiphysics measurements (e.g., high-precision temperature, acoustic, and electromagnetic methods) and data-driven analytics to reduce ambiguity in leak identification and localization (Almulhim and Ikhsanov, 2023; Kalwar et al., 2024; Volkov et al., 2019). However, comprehensive reviews of SCP diagnostics emphasize persistent gaps that limit reliable, repeatable diagnosis at scale: continued reliance on qualitative or semi-quantitative interpretation, limited accommodation of dynamic operating conditions (pressure and temperature fluctuations, shut-ins, and interventions) that drive non-stationary annular responses, and practical difficulty integrating heterogeneous monitoring streams due to noise, data volume, and specialized expertise requirements. In parallel, machine learning has been applied successfully to subsurface characterization and property prediction (e.g., lithofacies classification and permeability estimation), demonstrating the potential to scale interpretation across complex and heterogeneous systems (Al-Mudhafar and Wood, 2022; Abbas et al., 2023). Yet ML-only SCP diagnostics can be sensitive to dataset shift across wells and operating regimes and may yield nonphysical labels if not constrained by engineering plausibility, motivating PIML as a practical middle ground between physics-only models and purely data-driven classifiers.

Advances in high-resolution pressure and temperature sensing, coupled with SCADA-enabled data acquisition, now allow continuous monitoring of annular pressure behavior across multiple casing strings. Leveraging these capabilities, this study develops a hybrid PIML framework for automated and scalable SCP diagnosis. The methodology ingests continuous annular pressure–time histories, segments them into build-up and bleed-off cycles, and extracts diagnostic features such as slope matching and pressure matching. A decision tree algorithm, augmented with embedded physical rules governing fluid compressibility, thermal expansion effects, and leak-path mechanics, classifies each cycle into one of six SCP behavior types: no pressure, thermal pressure, trapped pressure, recharge pressure, constant pressure, and high-rate recharge. This integration of domain physics into the machine learning workflow ensures that predictions remain not only data-driven but also physically consistent and interpretable, reducing the risk of misclassification in noisy or complex datasets. Operationally, the approach enables real-time, SCADA-integrated surveillance across large well populations, allowing early identification of integrity threats, prioritization of high-risk wells, and proactive intervention planning. The novelty of this research lies in uniting physical modeling with automated pattern recognition for SCP diagnostics, creating a robust, interpretable, and field-ready tool that shifts well integrity management from reactive troubleshooting toward predictive, preventive decision-making at scale.

2. Barrier elements, barrier envelopes, and root causes of sustained casing pressure

The foundation of effective well integrity management rests on the barrier philosophy outlined in key industry standards (API, 2013; ISO, 2017; NORSOK, 2021). This philosophy provides a systematic framework to ensure containment of wellbore fluids and pressure throughout the well's lifecycle. This is achieved by identifying, designing, installing, verifying, monitoring, and maintaining well barrier elements (WBEs) to impede flow along potential leak paths (Alawad and Mohammad, 2016; Khalifeh and Saasen, 2020; Tayab et al., 2016). A WBE is a single, physical object or component designed, installed, and verified to prevent unintended flow along a specific leak path (Khalifeh and Saasen, 2020).

Each WBE must be fit for purpose, designed to withstand maximum anticipated pressures, temperatures, fluids (including corrosive CO_2 and H_2S), and mechanical loads over its service life (Ahmed and Salehi, 2021; Fanailoo et al., 2017). Their condition and verification status must be continuously tracked within the well integrity management system (WIMS) (Corneliussen et al., 2007; Yakoot et al., 2021). A well barrier envelope is a system of interdependent WBEs working together to form a continuous seal against flow from a pressure source to an undesired destination (Chakravorty, 2021; Choueri et al., 2024). If any critical element within the envelope fails, the entire envelope is considered compromised (Ashena, 2025). The most widely adopted implementation of the barrier philosophy in the oil and gas industry is the two-barrier principle (Torbergsen, 2012; NORSOK, 2021). This means that for any significant pressure source that could potentially drive fluid flow out of the wellbore (particularly wells capable of natural flow or penetrating over-pressured zones) (Davies et al., 2014), there must be two distinct, complete barrier envelopes in place to prevent that flow from reaching the surface, seabed, or other unintended locations (like a shallow aquifer or another formation. As illustrated in Fig. 2, during the drilling phase, the primary barrier typically consists of the hydrostatic column of drilling fluid, which provides well control under normal conditions. The secondary barrier is made up of the previously set casing strings, cement sheath, wellhead seals, and, if required, the blowout preventer (BOP). These two independent envelopes provide layered protection, ensuring that if the mud column is lost or compromised, pressure cannot reach the surface or shallow aquifers. Crucially, the failure of one barrier envelope should not compromise the second barrier envelope (IADC, 2020).

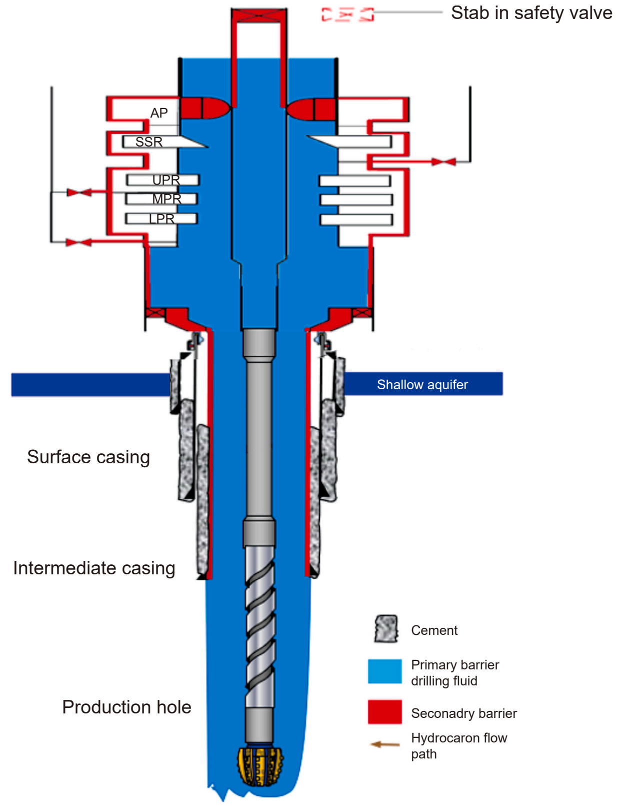

Fig. 2. Barrier schematic during drilling, showing the primary (mud hydrostatic) and secondary (casing–cement–wellhead) barriers.

3. Identifying the root cause of SCP

Traditionally, SCP surveillance relied on periodic pressure readings, bleed-off tests, and visual inspections. While these methods remain foundational, they are limited by intermittent sampling, subjective interpretation, and the inability to forecast evolving pressure trends. These limitations are increasingly critical as modern wells become deeper, more complex, and subject to stricter regulatory scrutiny. It may originate from several mechanisms, ranging from internal tubing leaks to external cement degradation. Common causes include.

- **Thermal expansion (APB)**: Incompressible fluid heating in sealed annuli during startup can cause pressure spikes. If the pressure drops and stabilizes after cooling, it is APB, not SCP. However, unmanaged APB can itself become a failure mode by exceeding casing burst ratings (Ibukun et al., 2024; Kiran et al., 2017).
- **Tubing leaks**: Corrosion, cracks, or failed connections in production tubing allow reservoir fluids to escape into the A-annulus (Dou et al., 2023).
- **Packer failure**: Damaged or degraded packer seals allow fluids from below the packer to communicate with the A-annulus (Zhu et al., 2012a, b).
- **Casing leaks**: Corrosion, mechanical stress, or poor connections can cause failures in the production casing, permitting pressure into the B-annulus.
- **Cement sheath defects**: Voids, channels, micro-annuli, or chemical degradation behind casing can create flow paths for formation fluids into outer annuli (B or C).
- **Wellhead/Hanger seal leaks**: Failed seals at casing or tubing hangers can allow cross-communication between annuli.
- **Operational pressure induction**: Intentional pressure introduced during tests or circulation may mimic SCP but should not rebuild unless a blockage or leak is present.

Fig. 3 illustrates typical bleed-off and build-up behaviors, linking each pattern to its likely root cause and severity. Fig. 4 illustrates the relationship between well failure mechanisms, their causes, and the specific well components or materials affected. Fig. 3 categorizes affected materials into casing/tubing, cement, wellhead and Christmas tree, formation, seals, or multiple elements. This mapping reinforces that SCP is often the consequence of multiple barrier element failures, underscoring the need for integrated monitoring and maintenance strategies.

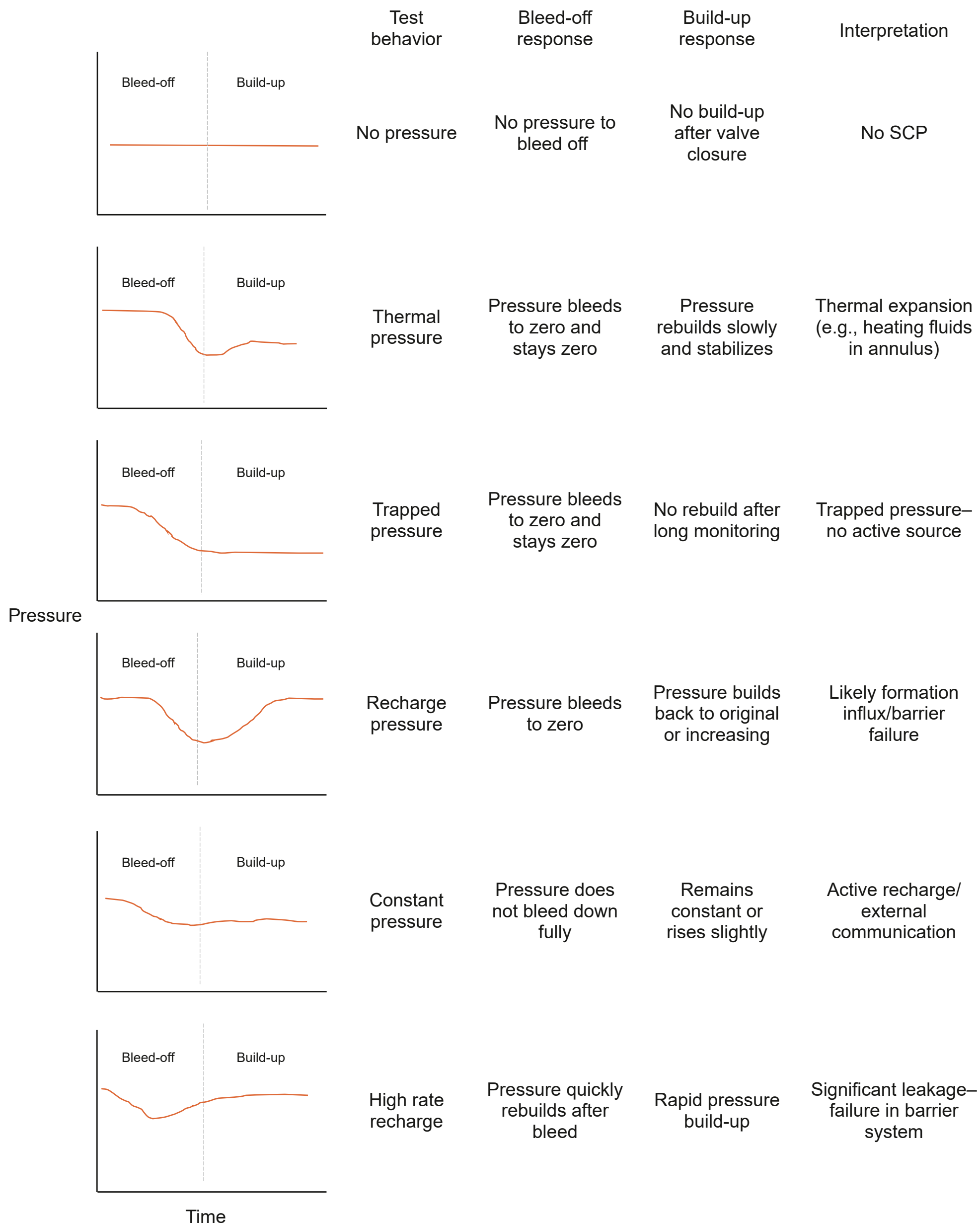

Fig. 3. Pressure diagnostic chart for SCP, illustrating typical bleed-off and build-up pressure responses over time.

4. Data and statistical summary

A comprehensive dataset of SCP diagnostic measurements was compiled across multiple wells, with pressure histories extracted for surface, intermediate, production, and conductor casing strings. The statistical distribution of median pressures per well and casing is shown in Fig. 5, where clear variability is observed among casing types. Production casing (Prod csg) exhibited the highest variability in median pressures across wells, reaching values up to 3484 psi, while conductor and surface casings were consistently lower, with median values of 70 psi and 50 psi, respectively.

The aggregated results for all wells are summarized in Table 1, which reports the count of valid observations, mean, standard deviation, and percentile ranges for each casing type. Conductor casing measurements (Cond csg) show a mean of 74 psi with a relatively narrow spread (standard deviation ~52 psi). Intermediate casing (Interm csg) displays a broader range, with mean 316 psi and standard deviation ~260 psi, reflecting both trapped and recharge pressure signatures. Production casing (Prod csg) has a mean of ~395 psi but with the widest overall range, from 0 up to 3484 psi. In contrast, surface casing (Surf csg) pressures are generally low, averaging 61 psi, though occasional spikes up to 593 psi are recorded.

High-pressure columns labeled as “Pressure” and “Pressure (psi)” correspond to wells with direct annulus monitoring, reporting consistent means of 956–967 psi with narrow percentile bands (P25–P75~863–1057 psi), highlighting stable recharge pressures under controlled conditions. These values indicate that annular systems without effective barrier integrity can sustain long-term pressures near 1000 psi, consistent with findings reported by Xu (2002).

Further, Fig. 6 depicts the mean ± standard deviation of pressure for each casing type across all wells. This highlights the clear

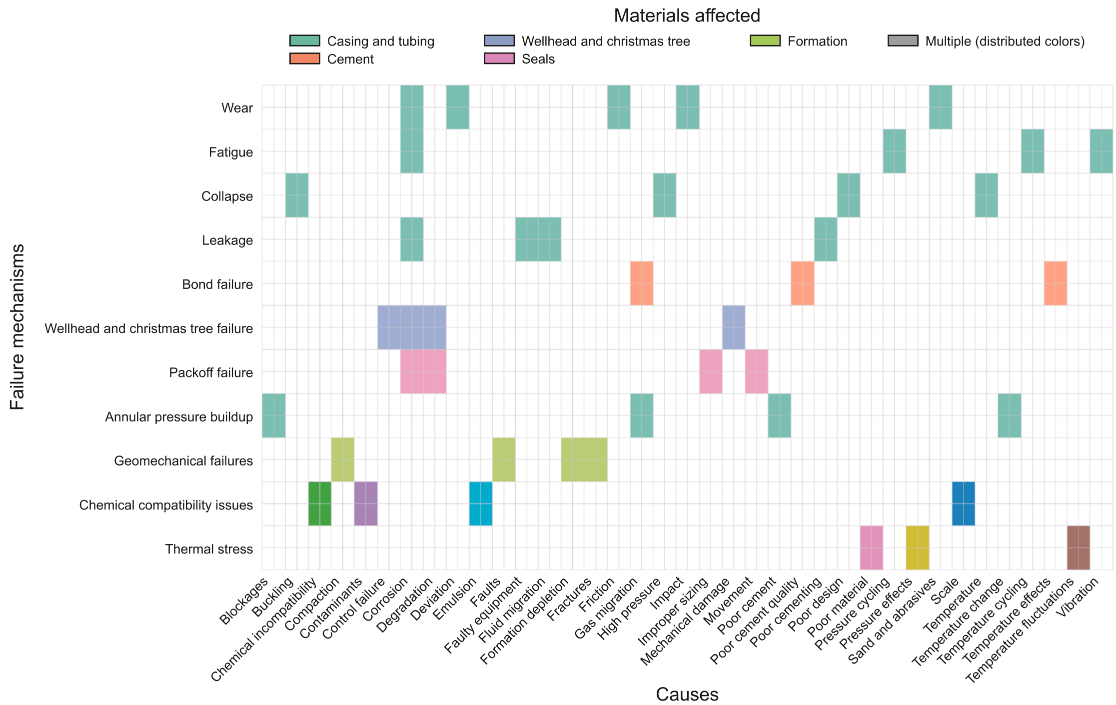

Fig. 4. A conceptual map illustrating the relationship between root causes (e.g., corrosion, geomechanical stress), failure mechanisms (e.g., cement debonding, packer leak), and the affected well barrier elements.

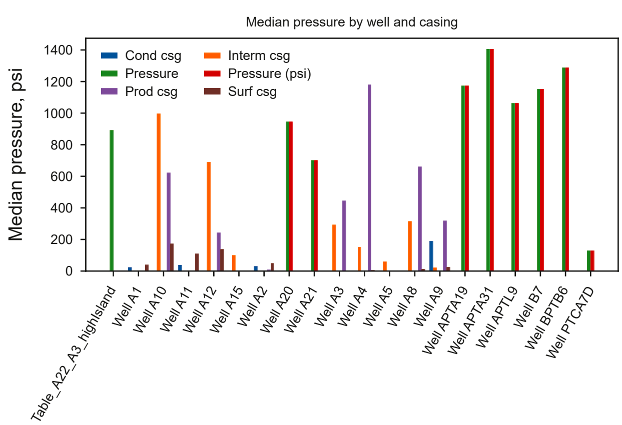

Fig. 5. Median SCP by well and casing string (conductor, surface, intermediate, production).

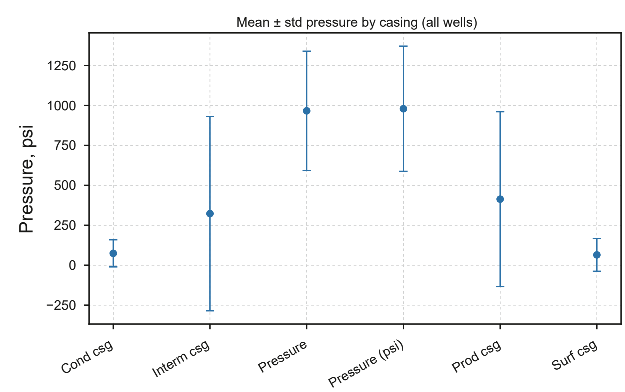

Fig. 6. Mean SCP and variability (± standard deviation) across wells, grouped by casing string.

separation between lower-pressure annuli (surface and conductor) and higher-pressure annuli (intermediate and production casings). Fig. 7 provides a boxplot representation, emphasizing the wide interquartile ranges observed for production and intermediate casing pressures, compared to the narrow spread of conductor and surface casing pressures.

Taken together, the statistical trends (Table 1) and graphical analyses (Figs. 4–7) underscore the strong casing-to-casing variability in SCP behavior. The results demonstrate that outer annuli (surface and conductor) generally retain limited pressure capacity, whereas inner annuli (intermediate and production) frequently exhibit recharge and higher sustained pressure behavior.

Table 1
Statistical descriptive summary of SCP measurements across all wells and casing types. Reported values include count of observations, mean, standard deviation, minimum, selected percentiles (P10, P25, median, P75, P90), maximum, and monitoring duration.

Casing	Depth, ft	Lithology	Mean, psi	Std, psi	Min, psi	P10, psi	P75, psi	P90, psi	Max, psi	Duration, h
Cond csg	582–738	Clays/sands	74	51	0	1	119	13	350	2098
Surf csg	1332–4776	Sands	61	55	0	6	97	140	593	2005
Interm csg	4310–6433	Shales/sands	315	259	0	50	499	696	4060	1995
Prod csg	6433–11196	Sandstone	395	309	0	49	619	781	3484	1995

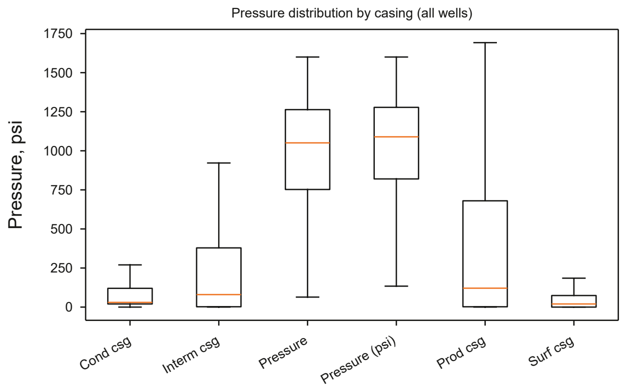

Fig. 7. Distribution of SCP by casing string across the full dataset (box-and-whisker summary).

5. Methodology for SCP diagnosis using machine learning and physics-informed modeling

Building on the monitoring and diagnostic capabilities of machine learning, a hybrid diagnostic methodology was developed to accurately and consistently classify SCP behaviors across diverse well environments. This approach combines the adaptive pattern recognition of machine learning (ML) with physics-informed rules to ensure engineering plausibility. The workflow, illustrated in Fig. 8, automates the interpretation of annular pressure data by segmenting it into meaningful pressure cycles, extracting diagnostic features, and applying a decision tree classifier enhanced with physical domain knowledge. The full dataset comprises 26 wells from Xu (2002) and the classifier is evaluated at the cycle level, where each sample corresponds to a complete build-up/bleed-off pressure cycle extracted from continuous annular pressure time series. Based on the extracted cycle inventory summarized by casing string, the total number of labeled cycles used for model development is $N = 2{,}326$ (Cond csg: 278; Surf csg: 727; Interm csg: 666; Prod csg: 655). To avoid information leakage, data splitting is performed by well (grouped split) such that all cycles from a given well are assigned exclusively to either training or testing. The decision-tree classifier is trained on 80% of wells (21 wells) and evaluated on a held-out set comprising the remaining 20% of wells (5 wells) using a fixed random seed (SEED = 42).

Pressure–time histories from well monitoring systems are first imported into the analysis framework. These datasets include continuous annular pressure measurements from multiple casing strings or annuli. The raw time-series data are then divided into two phases based on slope direction: build-up phases, where pressure increases over time, and bleed-off phases, where pressure decreases following venting or release. This segmentation isolates meaningful patterns and enables the identification of complete pressure cycles, each consisting of one build-up and one bleed-off event.

The local pressure derivative between two consecutive samples is approximated as

$$\left(\frac{\mathrm{d}P}{\mathrm{d}t}\right)_i \approx \frac{P_{i+1} - P_i}{t_{i+1} - t_i}, \tag{1}$$

where P_i is the measured pressure at time t_i. Build-up segments satisfy $(\mathrm{d}P/\mathrm{d}t)_i > 0$, and bleed-off segments satisfy $(\mathrm{d}P/\mathrm{d}t)_i < 0$. For each segment, a representative slope is estimated using a least-squares linear fit of the form

$$P(t) = mt + b, \tag{2}$$

where m is the slope (m_{bu} for build-up, m_{bo} for bleed-off) and b is the intercept. Two key diagnostic parameters are then computed for each complete cycle. The slope-matching metric is defined as

$$\delta_m = \frac{||m_{\mathrm{bu}}| - |m_{\mathrm{bo}}||}{\max\left(|m_{\mathrm{bu}}|, |m_{\mathrm{bo}}|\right)}, \tag{3}$$

and the pressure-matching metric is defined as

$$\delta_P = \frac{|P_{\mathrm{bu,end}} - P_{\mathrm{bo,end}}|}{\max\left(P_{\mathrm{bu,end}}, P_{\mathrm{bo,end}}\right)}, \tag{4}$$

where $P_{\mathrm{bu,end}}$ and $P_{\mathrm{bo,end}}$ are the end pressures of the build-up and bleed-off segments, respectively. To capture nonlinear build-up behavior, the build-up segment is also fitted with an exponential approach-to-stabilization model:

$$P(t) = P_0 + (P_\infty - P_0)[1 - \exp(-\lambda t)], \tag{5}$$

where P_0 is the pressure at the start of the build-up segment, P_∞ is the stabilized (asymptotic) pressure, and λ is the build-up rate constant. The goodness of this fit is quantified by the coefficient of determination

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(P_i - \widehat{P}_i)^2}{\sum_{i=1}^{n}(P_i - \overline{P})^2}, \tag{6}$$

where $\widehat{P}_i$ is the model-predicted pressure, $\overline{P} = \frac{1}{n}\sum_{i=1}^{n}$, P_i is the mean measured pressure over the fitted segment, and n is the number of samples in the fitted segment. These quantitative metrics (slopes, δ_m, δ_P, P_∞, λ, and R^2) ensure repeatability and remove the subjectivity of traditional interpretation.

The machine learning component of the workflow uses a decision tree algorithm trained on historical SCP cases to receive the computed features and predict the most likely SCP behavior. To prevent purely data-driven misclassification, a physics-informed module is integrated into the decision tree logic. This module incorporates domain-specific principles such as thermal expansion effects on incompressible annular fluids, leak flow dynamics through casing breaches, tubing leaks, or micro-annuli, and fluid communication between zones via failed seals or degraded cement.

Based on slope- and pressure-matching results (Eqs. (3)–(4)), combined with the relative magnitudes of build-up and bleed-off pressures and the exponential-fit parameters (Eq. (5) and R^2 from Eq. (6)), each cycle is assigned to one of six diagnostic categories: no pressure, thermal pressure, trapped pressure, recharge pressure, constant pressure, or high-rate recharge. The resulting classifications can be integrated into well integrity management systems to track SCP progression over time, flag wells requiring immediate intervention, and provide predictive warnings for wells with similar operational or structural characteristics. By merging automated analytics with physics-based validation, this methodology delivers accurate, repeatable, and scalable SCP diagnostics that are robust across a wide range of complex well conditions.

In this study, the identification of SCP behavior moves beyond conventional pattern recognition by embedding physical understanding directly into the classification process. The adopted PIML

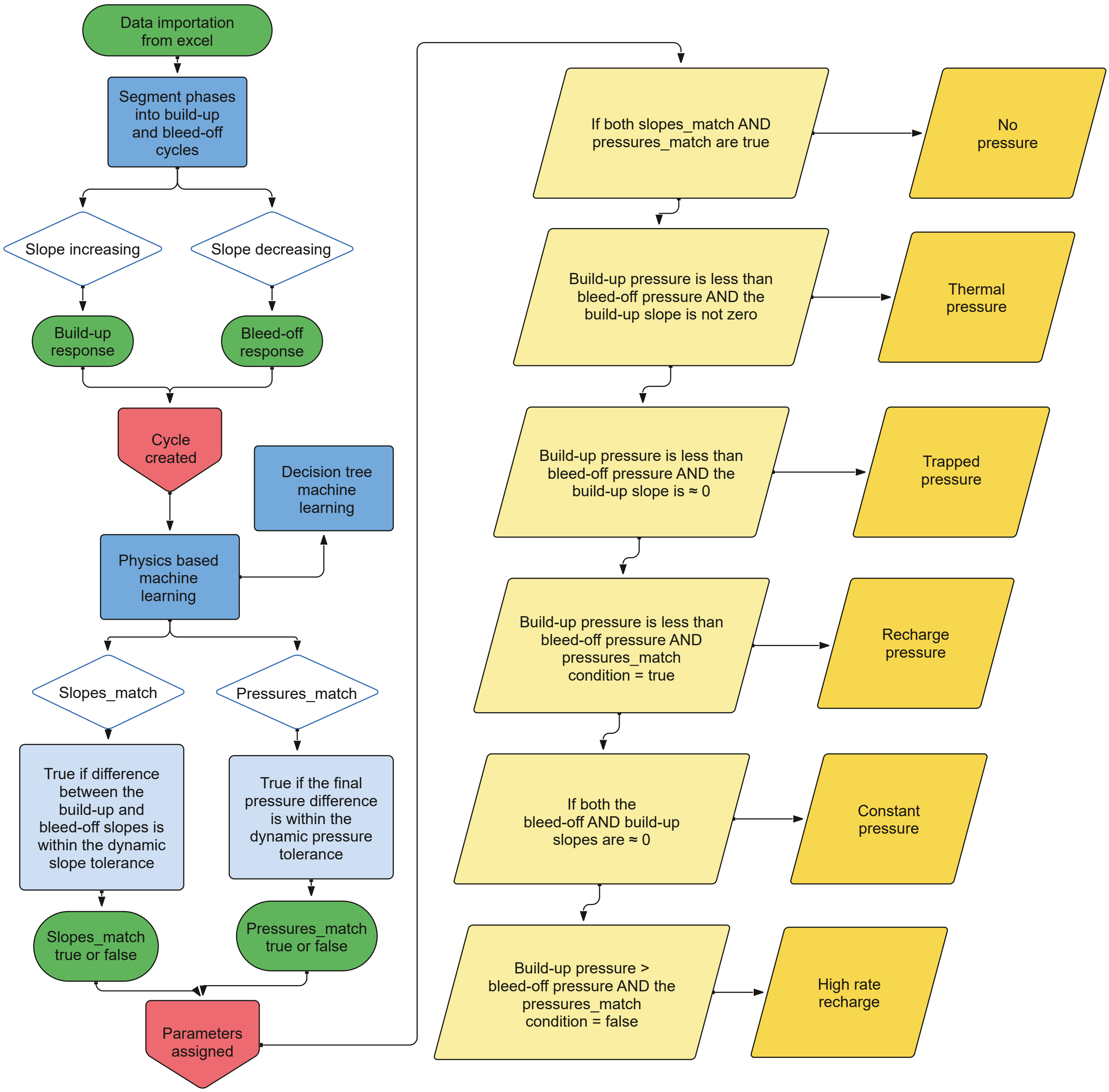

Fig. 8. PIML workflow for SCP diagnosis: cycle segmentation, physics-guided feature extraction, and classification into six SCP behavior types.

algorithm begins by segmenting continuous annular pressure–time histories into distinct build-up and bleed-off cycles, ensuring that each diagnostic event is analyzed in its full temporal context. For each cycle, quantitative features such as the rate of pressure change (slopes from Eq. (2)), residual pressure after venting, and the metrics δ_m and δ_P (Eqs. (3)–(4)), together with exponential build-up parameters (Eq. (5) and R^2 from Eq. (6)), are extracted and evaluated against dynamic tolerance thresholds derived from engineering principles. These thresholds reflect the governing physics of annular fluid compressibility, thermal expansion, and leak-path transmissibility, enabling the algorithm to differentiate between benign thermal effects and critical recharge patterns that signal barrier failure. By combining these physics-based constraints with a decision tree classifier trained on historical SCP cases, the method produces classifications that are not only data-driven but also physically plausible, reducing misclassification risk in noisy or ambiguous datasets and providing more reliable inputs for root cause analysis.

The workflow for physics-informed SCP behavior estimation can be summarized as follows.

1) **Data acquisition:** continuous annular pressure and temperature readings collected from wellhead sensors via SCADA.
2) **Cycle segmentation:** raw time-series data divided into build-up and bleed-off phases using slope direction changes based on Eq. (1).

3) **Feature extraction:** calculation of diagnostic parameters such as slopes, δ_m and δ_P (Eqs. (3)–(4)), and exponential-fit parameters (Eqs. (5)–(6)).
4) **Physics-based filtering:** application of constraints from engineering principles such as SCP behavior (Fig. 3).
5) **Classification via decision tree:** features fed into a trained decision tree algorithm to assign one of six SCP behavior types (no pressure, thermal pressure, trapped pressure, recharge pressure, constant pressure, high-rate recharge).
6) **Root cause linking:** classified behavior mapped to probable barrier failures or operational causes.
7) **Operational output:** results integrated into SCADA dashboards and well integrity management systems for real-time decision-making.

The PIML framework adopted in this study follows the physics-guided feature engineering and post-classification validation strategy within the broader family of physics-informed machine learning approaches. Rather than embedding governing equations directly into a model's loss function, the method leverages engineering principles at multiple stages of the diagnostic process. First, physics-derived features such as slope matching and pressure matching are computed from segmented build-up and bleed-off cycles, encoding known behaviors of annular fluids under compressibility, thermal expansion, and leak-path transmissivity constraints. These features guide the decision tree classifier toward physically plausible patterns. Second, classification outputs are validated against rule-based constraints derived from fluid mechanics and well integrity standards to avoid predictions that violate known operational or material limits. This approach ensures interpretability, maintains compatibility with limited or noisy datasets, and facilitates integration into real-time SCADA-based well integrity management systems, while retaining the core advantage of PIML: embedding domain physics to improve both accuracy and trustworthiness of machine learning outputs. Compared to more computationally intensive PIML variants, such as physics-informed neural networks (PINNs) that embed differential equations into deep learning architectures, this strategy was selected for its operational practicality, lower data and hardware requirements, and greater transparency for engineering decision-making. Similar PIML concepts have been successfully applied to permeability prediction in heterogeneous carbonate reservoirs, demonstrating the value of embedding domain physics within ML frameworks to improve generalization and interpretability (Khassaf et al., 2025).

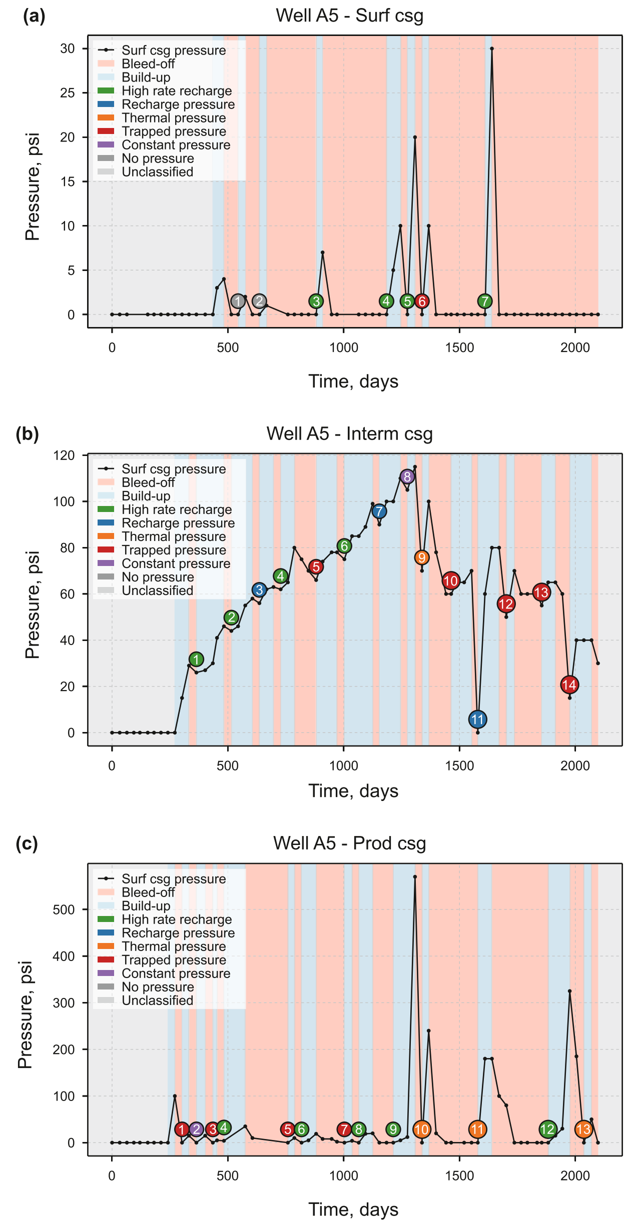

Fig. 9. (**a**) Well A5 surface casing pressure time series with PIML cycle classification. (**b**) Well A5 intermediate casing pressure time series with PIML cycle classification. (**c**) Well A5 production casing pressure time series with PIML cycle classification.

6. Results

The proposed PIML workflow was applied to a series of real-world well datasets to evaluate its capability in accurately diagnosing and classifying SCP behaviors across multiple annuli and well types. The results show strong agreement between model predictions and observed pressure patterns, supporting the robustness and generalizability of the framework. The automated diagnostic classification leveraged slope analysis, pressure matching, and cycle segmentation to detect six distinct SCP conditions, as outlined in the methodology. The PIML framework was applied to 26 wells for training the model but it presented for only three wells (A5, A8, A12) across three annuli, surface casing, intermediate casing, and production casing, to evaluate classification accuracy, curve-fit quality, and diagnostic robustness under varying pressure behaviors.

Surface casing (Fig. 9(a)) exhibited long periods of insignificant pressure, correctly classified as no pressure, followed by multiple significant pressure spikes. These spikes were consistently identified as high-rate recharge events, reflecting rapid annular pressurization likely driven by leak paths with high transmissivity. The algorithm achieved reasonable classification across all cycles, with curve-fit R^2 values exceeding 0.99 in most cases. The lowest R^2 was 0.99 for Cycle 4, maintaining a high visual correlation between predicted and measured trends. Intermediate casing (Fig. 9(b)) displayed a steady increase in pressure from initial conditions to approximately day 1250, followed by a gradual decline to near-baseline values. The PIML model classified all 14 cycles accurately, with classifications dominated by recharge pressure and trapped pressure conditions. These behaviors are consistent with leak sources that recharge annular pressure over time and partially seal during bleed-off. The lowest R^2 was 0.79 for Cycle 1, with an overall mean R^2 of 0.93, indicating strong agreement between modeled and observed pressure-time profiles.

Production casing (Fig. 9(c)) pressure history revealed intermittent high-magnitude spikes, interspersed with extended low-pressure periods. The algorithm identified a mixture of high-rate recharge and recharge pressure events, reflecting both fast and moderate pressure rebuild rates. Lower R^2 values (minimum 0.51 for Cycle 9) were associated with highly irregular or noisy data, but classification accuracy remained above 90% across all cases, demonstrating the model's resilience in handling variable data quality. Overall, the results confirm that the PIML framework successfully distinguishes between SCP types with high accuracy, even in complex or irregular pressure profiles. Embedding physical constraints, such as slope matching and pressure matching, ensures interpretability and mitigates overfitting, enabling reliable application in real-time monitoring environments.

After that, the PIML framework was evaluated on Well A8 to assess performance across surface casing, intermediate casing, and production casing annuli under varying operational and SCP conditions. Surface casing (Fig. 10(a)) for approximately the first 600 days, the surface casing maintained negligible pressure, correctly classified as no pressure. Subsequent events included significant spikes identified as high-rate recharge. One misclassification occurred in Cycle 8, where the alg orithm labeled the event as no pressure despite a large pressure drop—likely due to borderline slope and pressure match values. Overall, classification accuracy remained high, with curve-fit R^2 values exceeding 0.80 for most cycles, and a lowest R^2 of 0.61 for Cycle 6, indicating reduced fit quality during highly irregular transitions.

Intermediate casing (Fig. 10(b)) exhibited complex and highly variable pressure behavior over the monitoring period, with alternating recharge pressure, trapped pressure, and thermal pressure classifications. The algorithm demonstrated strong adaptability to these variations, achieving an average R^2 of 0.93. Lower R^2 values (minimum 0.81) were associated with abrupt pressure shifts, yet classification accuracy was maintained. The model's ability to differentiate recharge-driven trends from trapped or thermal patterns underscores the benefit of physics-informed constraints in improving interpretability. Production casing (Fig. 10(c)) started with a thermal pressure decline, followed by a recovery to approximately 70% of maximum pressure, accompanied by multiple high-rate recharge events. Later fluctuations were predominantly classified as recharge pressure. The average curve-fit R^2 was 0.975, with all but one cycle exceeding 0.90, reflecting excellent alignment between model predictions and observed trends. The consistent classification performance highlights the framework's reliability for higher-pressure annuli. Across all annuli, the PIML framework maintained high classification accuracy, robust handling of variable cycle patterns, and strong curve-fit performance, further validating its suitability for real-time SCP diagnostics in operational environments.

The PIML framework also was applied to Well A12 to evaluate classification accuracy for surface casing, intermediate casing, and production casing annuli. Well A12's surface casing (Fig. 11(a)) exhibits diverse pressure trends, with multiple recharge and high-rate recharge cycles accurately identified. All classifications visually align with observed data, resulting in an average R^2 of 0.95. Notably, Cycle 10 (trapped pressure) and Cycle 4 (thermal pressure) showcase correct categorization despite sharp transitions, confirming the robustness of slope and trend detection. Intermediate casing (Fig. 11(b)) data is rich in pressure variations, with many cycles achieving R^2 values above 0.90. The classification accuracy is visually high, including correct identification of recharge, thermal, and trapped pressure cycles. Cycle 3, labeled as constant pressure, could alternatively be considered recharge pressure given its slight upward trend—highlighting an area where classification thresholds may be fine-tuned. Production

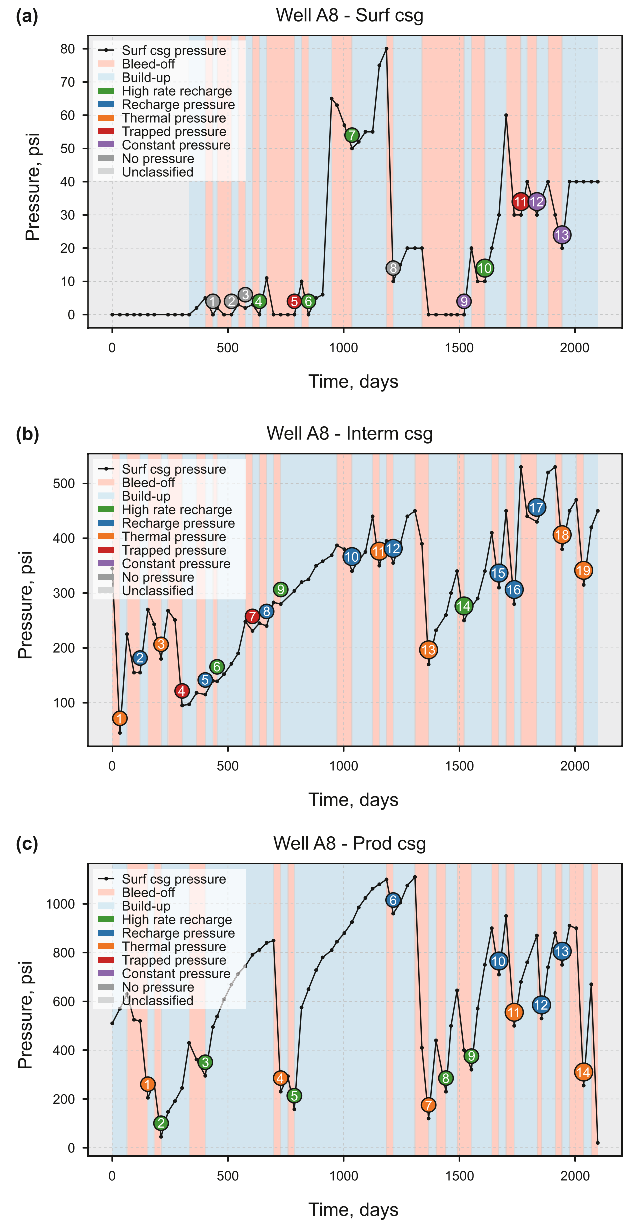

Fig. 10. (**a**) Well A8 surface casing pressure time series with PIML cycle classification. (**b**) Well A8 intermediate casing pressure time series with PIML cycle classification. (**c**) Well A8 production casing pressure time series with PIML cycle classification.

casing (Fig. 11(c)) demonstrates unique diagonal spikes in pressure, all correctly classified as recharge or high-rate recharge. Even with these distinct patterns, curve-fit R^2 values remain high, reinforcing the model's ability to adapt to rapid and sharp pressure changes.

Across all annuli, Well A12 confirms the PIML model's capability to consistently identify recharge, thermal, and trapped pressure events with high visual and statistical accuracy. Occasional borderline cases suggest potential refinement for distinguishing constant from recharge trends.

7. Discussion

Across the 26 wells analyzed, the PIML framework consistently demonstrated high cycle-level classification accuracy and strong curve-fit performance, confirming its robustness under diverse

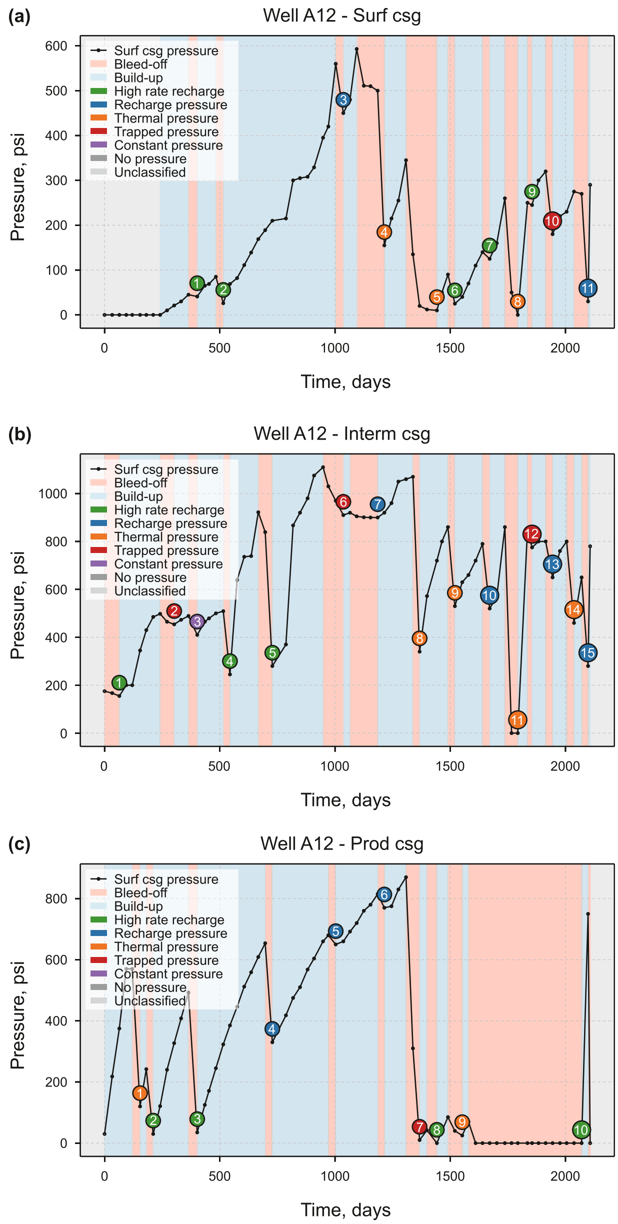

Fig. 11. (**a**) Well A12 surface casing pressure time series with PIML cycle classification. (**b**) Well A12 intermediate casing pressure time series with PIML cycle classification. (**c**) Well A12 production casing pressure time series with PIML cycle classification.

operating conditions. For the detailed case studies, per-annulus accuracies ranged from about 90% to 100%, with average R^2 values typically above 0.90 (Table 2). Surface casing in wells A5 and A12 achieved 100% classification accuracy with average $R^2 \approx 0.95$–0.99, reflecting the relatively simple pressure behavior dominated by long no pressure intervals interrupted by distinct high-rate discharge spikes. Intermediate casings exhibited more complex behavior, with mixed recharge, trapped, and thermal pressure responses; nevertheless, accuracies remained above 90% and average R^2 values were around 0.93, even when multiple SCP mechanisms operated within the same monitoring period. Production casings, which were characterized by intermittent, high-magnitude spikes and noisier signals, still showed accuracies greater than 90% and average R^2 values typically between 0.88 and 0.97, demonstrating that the framework can maintain reliable performance in the most challenging annuli.

A comparison across annuli highlights where the PIML approach provides the greatest diagnostic value. Surface and conductor casings generally show low to moderate pressures and relatively few complex cycles, so both manual and automated methods can classify these events with high confidence. In contrast, intermediate and production casings display a wider range of behaviors—overlapping recharge pressure, trapped pressure, and thermal pressure cycles—which are more prone to conflicting interpretations when evaluated manually. In these annuli, the physics-guided features (slope matching, pressure matching, and exponential build-up parameters) allow the decision tree to separate thermal-dominated responses from genuine recharge behavior, improving consistency across cycles and wells. The framework also handled noisy datasets effectively: in cases such as the production casing of Well A5 and the intermediate casing of Well A8, local R^2 values dropped for some cycles due to irregular pressure fluctuations, yet the predicted SCP types remained physically plausible and aligned with the overall pressure evolution, with only isolated misclassifications (e.g., a single no pressure vs high-rate recharge case in the A8 surface casing and a borderline Constant vs Recharge cycle in the A12 intermediate casing).

From an operational perspective, the quantified results demonstrate how the PIML outputs can be translated into actionable decisions. Persistent recharge or high-rate recharge classifications in intermediate or production casings point to likely barrier degradation (e.g., cement sheath defects or casing leaks) and can trigger targeted diagnostics such as cement bond logging, annulus venting strategy review, or wellhead seal inspection. Conversely, repeated thermal pressure or trapped pressure patterns with high R^2 fits and stable end-pressures may justify continued surveillance rather than immediate intervention. Because the framework processes thousands of cycles ($N = 2{,}326$ in this study) using a consistent set of physics-guided rules, it offers reproducible and scalable screening across large well inventories, in contrast to traditional SCP diagnosis workflows that depend on engineer-by-engineer judgment applied to a small subset of wells.

Table 2
Summary of PIML classification performance and operational interpretation across wells and annuli.

Well	Annulus type	Accuracy, %	Avg R^2	Most common SCP types	Operational notes
A5	Surface casing	100	>0.99	No pressure, high-rate recharge	High transmissivity leak path; spikes easily detected; minimal misclassification.
A5	Intermediate casing	100	0.93	Recharge pressure, trapped pressure	Gradual recharge trends; strong agreement with observed patterns.
A5	Production casing	>90	0.88*	High-rate recharge, recharge pressure	Intermittent spikes; noisy data caused lower R^2 for certain cycles.
A8	Surface casing	~95	0.80	No pressure, high-rate recharge	One misclassification due to borderline slope/pressure match thresholds.
A8	Intermediate casing	>90	0.93	Recharge pressure, trapped, thermal	Highly variable patterns; strong adaptability to changing conditions.
A8	Production casing	>95	0.975	High-rate recharge, recharge pressure	Stable classification; high accuracy even in multi-phase trends.
A12	Surface casing	100	0.95	Recharge, high-rate recharge, thermal	Correct classification despite sharp pressure transitions.
A12	Intermediate casing	>95	>0.90	Recharge, thermal, constant pressure	One borderline constant/recharge cycle; possible threshold refinement needed.
A12	Production casing	100	>0.90	Recharge, high-rate recharge	Distinct diagonal spikes classified correctly; robust under rapid changes.

Despite these promising results, several limitations of the current framework should be acknowledged. First, the training and validation datasets are limited to 26 wells from a single region, with SCP behavior dominated by Gulf of Mexico–style completion practices and operating envelopes. The decision thresholds for slope and pressure matching were tuned to this dataset and may require recalibration before application to wells with markedly different venting procedures, sensor resolutions, or thermal regimes (e.g., HPHT or geothermal wells). Second, the present implementation relies solely on annular pressure time series; temperature, acoustic, and cement-evaluation data are not yet integrated, which restricts the ability to localize the leak path or distinguish between multiple concurrent failure mechanisms. Third, the classifier is based on a single decision tree for transparency and ease of deployment; while this choice enhances interpretability, it may under-utilize potentially richer patterns that could be captured by ensemble methods or more advanced PIML architectures. Finally, borderline cases—such as cycles that sit between constant and low-level recharge behavior—highlight the sensitivity of the method to tolerance settings and illustrate that some level of expert review remains necessary for high-consequence decisions.

These limitations suggest several directions for future work. Extending the dataset to include additional basins, completion designs, and operating conditions will allow more rigorous statistical validation and stress-testing of the model. Incorporating multi-sensor data streams and ensemble-based PIML models could further improve diagnostic resolution and robustness, while adaptive or data-driven threshold optimization would help reduce ambiguity in borderline cycles. Coupling the diagnostic classes with quantitative risk ranking and economic impact models would complete the workflow, linking SCP classification directly to intervention prioritization and lifecycle well integrity management.

8. Conclusions

This study developed a PIML framework for automated diagnosis of SCP from continuous annular pressure time-series data. The workflow combines cycle segmentation with physics-guided feature extraction (slope matching and pressure matching) and a decision-tree classifier constrained by fluid compressibility, thermal expansion, and leak-path mechanics to generate physically plausible and interpretable SCP classifications. Across 26 wells, with detailed demonstrations for wells A5, A8, and A12, the framework reliably distinguished six SCP behavior types (no pressure, thermal pressure, trapped pressure, recharge pressure, constant pressure, and high-rate recharge) and achieved classification accuracy exceeding 90% with average curve-fit performance (R^2) above 0.90. These results are operationally important because they convert SCP evaluation from a manual and often subjective interpretation process into a reproducible diagnostic output that can be integrated into integrity surveillance to prioritize wells exhibiting recharge-driven behavior indicative of barrier degradation. The approach is expected to generalize to other wells because it relies on physically meaningful signatures (pressure rebuild rate, residual pressure after bleed-off, and thermal-driven trends) rather than purely statistical pattern matching; in practice, transfer to new assets requires only limited calibration of slope/pressure tolerances to reflect local venting procedures, sensor noise levels, and operating envelopes, while preserving the same physics-guided feature set. The method is therefore suitable for deployment within SCADA or well integrity management systems to support scalable, near-real-time SCP classification and integrity triage across large well populations. Limitations primarily relate to data quality and borderline cycles where constant and recharge behavior are subtle; future work will expand validation across broader operating conditions and incorporate adaptive threshold optimization and risk-ranking linkages that translate diagnostic class into recommended intervention priority.

CRediT authorship contribution statement

Ahmed Alsubaih: Conceptualization. **Evan Wetmore:** Methodology, Formal analysis. **Watheq J. Al-Mudhafar:** Writing – review & editing, Visualization, Validation. **Kamy Sepehrnoori:** Conceptualization. **Alberto L. Manriquez:** Investigation. **Mojdeh Delshad:** Data curation.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethical approval

Not applicable.

Funding

This research received no external funding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abbas, M.A., Al-Mudhafar, W.J., Wood, D.A., 2023. Improving permeability prediction in carbonate reservoirs through gradient boosting hyperparameter tuning. Earth Sci. Inf. 16, 3417–3432. https://doi.org/10.1007/s12145-023-01099-0.
Ahmed, S., Salehi, S., 2021. Failure mechanisms of the wellbore mechanical barrier systems: implications for well integrity. J. Energy Resour. Technol. Trans. ASME 143 (7). https://doi.org/10.1115/1.4050694/1105947.
Al-Mudhafar, W.J., Wood, D.A., 2022. Tree-based ensemble algorithms for lithofacies classification and permeability prediction in heterogeneous carbonate reservoirs. In: Offshore Technology Conference, OTC pp. D031S041R003. https://doi.org/10.4043/31780-MS.
Alam, M., Minor, J., 2024. Solving sustained casing pressure problems at the source. In: SPE symposium on well integrity management 2024. https://doi.org/10.2118/220571-MS.
Alawad, M.N.J., Mohammad, H.A., 2016. Engineering management and inspection schedule of petroleum well integrity. J. Emerg. Trends Eng. Appl. Sci. 7 (3), 109–117.
Almulhim, B.K., Ikhsanov, N.A., 2023. Identification of multiple annuli leakages using acoustic sensor. In: ADIPEC, Richardson: Society of Petroleum Engineers pp. SPE 216939–MS. https://doi.org/10.2118/216939-MS.
API, 2013. API RP 96. https://standards.globalspec.com/std/1586765/api-rp-96API.
Alsubaih, A.A.S., Sepehrnoori, K., Delshad, M., Alsaedi, A., 2025a. A comprehensive review of well integrity challenges and digital twin applications across conventional, unconventional, and storage wells. Energies 18 (17), 4757. https://doi.org/10.3390/en18174757.
Alsubaih, A., Sepehrnoori, K., Mojdeh, D., Lopez, M.A., 2025b. A comprehensive review of CO2 subsurface storage: Integrity, safety, and economic viability. Energy Geoscience 6 (3), 100441. https://doi.org/10.1016/J.ENGEOS.2025.100441.
Ashena, R., 2025. Barriers. In: The practical handbook of well control, pp. 20–31. https://doi.org/10.1201/9781003473770-2.
Chakravorty, A., 2021. Use of machine learning algorithms to predict well barrier elements and envelopes in drilling operations. https://uis.brage.unit.no/uis-xmlui/handle/11250/2829441.
Choueri, N., Jaculli, M.A., Leite, A.G.A.S., Mendes, J.R.P., 2024. Use of graphs to assess well safety in drilling projects and during operations by identification of available barrier elements and consolidation of barrier envelopes. Petrol. Res. 9 (3), 418–431. https://doi.org/10.1016/j.ptlrs.2024.02.002.
Corneliussen, K., Sørli, F., Haga, H.B., Tenold, E., Menezes, C., Grimbert, B., Owren, K., 2007. Well integrity management system (WIMS)—A systematic

way of describing the actual and historic integrity status of operational Wells. Proc. SPE Annu. Tech. Conf. Exhib. 4, 2576–2583. https://doi.org/10.2118/110347-MS.

Davies, R.J., Almond, S., Ward, R.S., Jackson, R.B., Adams, C., Worrall, F., Herringshaw, L.G., Gluyas, J.G., Whitehead, M.A., 2014. Oil and gas wells and their integrity: implications for shale and unconventional resource exploitation. Mar. Petrol. Geol. 56, 239–254. https://doi.org/10.1016/j.marpetgeo.2014.03.001.

Dou, N., Hu, Z., Wang, Z., He, H., Yang, S., 2023. Risk assessment of annular pressure caused by tubing leakage in offshore gas Wells with high CO_2. Geofluids 2023, 9072217. https://doi.org/10.1155/2023/9072217.

Fanailoo, P., Buchmiller, D., Ouyang, S., Allen, E., 2017. Risk based approach to well plugging & abandoning - reducing costs while verifying risk. Offshore Technol. Conf. 4, 2728–2744. https://doi.org/10.4043/27921-MS.

IADC, 2020. Well control barrier philosophy.

Ibukun, M., Elyan, E., Amish, M., Njuguna, J., Oluyemi, G.F., 2024. A review of well life cycle integrity challenges in the oil and gas industry and its implications for sustained casing pressure (SCP). Energies 17 (22), 5562.. https://doi.org/10.3390/en17225562.

International Organization for Standardization, 2017. ISO 16530-1:2017 Petroleum and natural gas industries — Well integrity — Part 1: Life cycle governance. https://www.iso.org/standard/63192.html.

Kalwar, G.M., Al-Sagr, A., Rowaie, M.A., Malyazin, A., 2024. Closing the loop with the well integrity: Deploying high-definition acoustic and temperature logging for detecting complex multi annuli communication and integration of geochemical analysis. In: SPE Europe Energy Conference and Exhibition, pp. SPE 220115–MS. https://doi.org/10.2118/220115-MS.

Kazemi, M., Wojtanowicz, A.K., 2022. Improved method for testing integrity loss of Wells with sustained casing pressure. Energies 15 (10), 3632. https://doi.org/10.3390/en15103632.

Khalifeh, M., Saasen, A., 2020. General principles of well barriers. In: Well Barrier Element Design, pp. 11–69. https://doi.org/10.1007/978-3-030-39970-2_2.

Khassaf, A.K., Al-Hameed, Z.M., Al-Mohammedawi, N.R., Al-Mudhafar, W.J., Wood, D.A., Abbas, M.A., Ameur-Zaimeche, O., Alsubaih, A.A., 2025. Physics-informed machine learning for enhanced permeability prediction in heterogeneous carbonate reservoirs. In: Offshore Technology Conference. Houston, TX, USA. https://doi.org/10.4043/35892-MS.

Kinik, K., 2012. Risk of well integrity failure due sustained casing pressure. LSU Master's Theses. https://doi.org/10.31390/gradschool_theses.3456.

Kiran, R., Teodoriu, C., Dadmohammadi, Y., Nygaard, R., Wood, D., Mokhtari, M., Salehi, S., 2017. Identification and evaluation of well integrity and causes of failure of well integrity barriers: A review. J. Nat. Gas Sci. Eng. 45, 511–526. https://doi.org/10.1016/j.jngse.2017.05.009.

NORSOK, 2021. NORSOK D-010. https://online.standard.no/en/norsok-d-010-202.

Rusch, D.W., Sabins, F., Aslakson, J., 2004. Microannulus leaks repaired with pressure-activated sealant. In: SPE Eastern Regional Meeting, pp. 83–90. https://doi.org/10.2118/91399-MS.

Tayab, M.R., Abu Aishah, A.N.M., Valappil, S., Draz, A.M.A., 2016. Assessment and strengthening of well control barriers to prevent well blowout. In: Abu Dhabi International Petroleum Exhibition and Conference 2016. https://doi.org/10.2118/182899-MS.

Torbergsen, H.E.B., 2012. An introduction to well integrity.

Volkov, M., Yurchenko, A., Ellis, R., 2019. The loss of primary barrier: An integrated approach of downhole leak location and quantification. Case study. In: SPE Nigeria Annual International Conference and Exhibition, pp. SPE 198838–MS. https://doi.org/10.2118/198838-MS.

Xu, R., 2002. Analysis of diagnostic testing of sustained casing pressure in wells. LSU Dr. Diss. 1192.LSU Dr. Diss. https://doi.org/10.31390/gradschool_dissertations.1192.

Yakoot, M.S., Elgibaly, A.A., Ragab, A.M.S., Mahmoud, O., 2021. Well integrity management in mature fields: A state-of-the-art review on the system structure and maturity. J. Pet. Explor. Prod. Technol. 11 (4), 1833–1853. https://doi.org/10.1007/s13202-021-01154-w.

Zhu, H., Lin, Y., Zeng, D., Zhang, D., Chen, H., Wang, W., 2012a. Mechanism and prediction analysis of sustained casing pressure in "A" annulus of CO_2 injection well. J. Petrol. Sci. Eng. 92–93, 1–10. https://doi.org/10.1016/j.petrol.2012.06.013.

Zhu, H., Lin, Y., Zeng, D., Zhang, D., Wang, F., 2012b. Calculation analysis of sustained casing pressure in gas wells. Pet. Sci. 9 (1), 66–74. https://doi.org/10.1007/s12182-012-0184-y.