PREDICTIVE MAINTENANCE OF SUBSURFACE EQUIPMENT (ESPs) USING SENSOR DATA; A MACHINE LEARNING APPROACH

IGIEBOR IYOBOSA

ENG2006428

DEPARTMENT OF PETROLEUM ENGINEERING

FACULTY OF ENGINEERING

UNIVERSITY OF BENIN

BENIN CITY

JULY 2025

CERTIFICATION

This is to certify that this project was carried out by IGIEBOR IYEBOSA of the Department of Petroleum Engineering with the matriculation number ENG2006428 in partial fulfillment of the requirements for the Award of the Degree, Bachelor of Engineering (B.ENG)

DR. S. A. IGBINERE	**DATE**
(PROJECT SUPERVISOR)	
DR. O. A. TAIWO	**DATE**
(PROJECT COORDINATOR)	
DR. IKPONMWOSA OHENHEN	**DATE**
HEAD OF DEPARTMENT	
PROF. KELVIN CHINWUBA IGWILO KELVIN	**DATE**
(EXTERNAL SUPERVISOR)	

CERTIFICATION

This is to certify that this project work was carried out by IGIEBOR IYOBOSA with Matriculation Number ENG2006428 under my supervision. It is adequate and satisfactory. both in scope and content, for the award of Bachelor of Engineering (BENG) Degree in Petroleum Engineering of the University of Benin.

_____ _____

Dr. Sunday Agbons Igbinere **DATE**

(Project Supervisor)

APPROVAL

This project work is hereby approved in partial fulfilment of the requirements for the award of Bachelor of Engineering (BENG) Degree in Petroleum Engineering from the University of Benin.

Engr. Dr, Ikponmwosa Ohenhen **DATE**

(HEAD OF DEPARTMENT)

DEDICATION

1 dedicate this work to God Almighty, the source of my strength, wisdom, and perseverance throughout this academic journey.

I also dedicate it to my family, especially my parents, for their unwavering support, prayers. and encouragement.

To everyone who believed in me this is for you.

ACKNOWLEDGEMENT

I sincerely thank God Almighty for His grace, wisdom, and strength throughout my academic journey and the successful completion of this project.

I am deeply grateful to my project supervisor, Dr. Sunday Agbons Igbinere for his consistent guidance, encouragement, and constructive feedback. His support played a crucial role in shaping this research.

My sincere appreciation also goes to Eng. EDOBOR FRANKIE CHRISTOPHER, whose availability and academic support during critical moments greatly contributed to my academic growth.

I am especially grateful to my senior colleague and friend, course mate whose mentorship. guidance, and project format greatly helped me understand and structure my own work from start to finish.

Finally, I appreciate my classmates and everyone who, in one way or another, offered support, shared ideas, or provided encouragement during the course of this research.

ABSTRACT

This thesis presents a comprehensive machine learning-based predictive maintenance (PdM) framework for Electric Submersible Pumps (ESPs) in petroleum production, aiming to mitigate unplanned downtime and optimize maintenance scheduling. ESPs are critical for artificial lift in oil wells but are prone to failures—mechanical, electrical, and operational—that disrupt production and incur high intervention costs. Traditional maintenance strategies, whether reactive or calendar-based, fail to leverage real-time sensor data effectively. This study bridges this gap by developing data-driven models to forecast ESP failures using historical and real-time telemetry. The research begins with a detailed exploration of ESP systems, their components, and common failure modes, emphasizing the role of downhole sensors (pressure, temperature, vibration, motor current) in monitoring pump health. A year-long dataset from an onshore oilfield, comprising hourly measurements of motor current, intake/discharge pressures, temperatures, vibration frequency, production rate, and choke settings, is rigorously preprocessed. Key steps include data cleaning, imputation of missing values, outlier handling, and feature engineering. Temporal features such as rolling statistics (3-hour mean/standard deviation) and lagged variables (1- and 2-hour delays) are engineered to capture dynamic system behavior. Three ensemble machine learning models—Random Forest, Gradient Boosting, and XGBoost—are benchmarked against a linear regression baseline to predict motor housing temperature one hour ahead. The models are evaluated using time-series cross-validation to prevent data leakage. Results show that Gradient Boosting achieves the lowest mean squared error (MSE: 460.26 $°F^2$) and highest R^2 (0.926), outperforming the linear baseline (MSE: 477.03 $°F^2$, R^2: 0.923). Feature importance analysis reveals that lagged intake pressure and motor temperature dominate predictions, accounting for over 80% of explanatory power. However, the models exhibit slight under-prediction during rapid temperature spikes, highlighting a need for further refinement in extreme-event forecasting. Unsupervised clustering identifies five operational modes (startup, steady-state, high-temperature stress, low-pressure events, and current spikes), providing context for model performance variations during regime transitions. The study underscores the potential of hybrid approaches, combining regime classification with mode-specific regression, to enhance accuracy. The thesis concludes with actionable recommendations: deploying the optimized model in real-time monitoring systems with adaptive alert thresholds, integrating additional sensor modalities (e.g., vibration, acoustics), and establishing feedback loops for continuous model retraining. Economically, the proposed PdM framework could reduce unplanned downtime by 30–50% and lower maintenance costs by 20–40%, as evidenced by industry benchmarks. Future work includes exploring deep learning architectures (e.g., LSTMs) for longer-term dependencies and extending the framework to other artificial-lift systems. This research contributes a scalable, interpretable, and empirically validated PdM pipeline, advancing the transition from reactive to proactive maintenance in petroleum production. By harnessing sensor data and machine learning, operators can anticipate failures, optimize interventions, and maximize ESP run life—translating into significant cost savings and production efficiency gains.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

1.0 INTRODUCTION

Electric submersible pumps (ESPs) are among the most widely used artificial-lift technologies in the petroleum industry, playing a pivotal role in sustaining production from wells whose natural reservoir drive has diminished. It is estimated that over 90 percent of producing oil wells globally employ some form of artificial lift to maintain economically viable flow rates. ESPs are particularly valued for their capacity to deliver very high fluid volumes—on the order of a few hundred to tens of thousands of barrels per day—making them indispensable in fields characterized by large production targets or challenging downhole conditions.

An ESP system typically comprises multiple centrifugal-pump stages mounted in series, all driven by a submersible electric motor. A conventional ESP string (Figure 1) begins at the surface with power cables and a variable-frequency drive (VFD) or constant-speed motor controller; the motor controller regulates voltage and frequency to optimize pump performance under varying downhole conditions. Below the surface equipment is the wellhead, from which insulated power cables descend into the wellbore to the motor section. The motor is sealed within a hermetic housing to prevent well fluids from entering, and is coupled directly to the pump shaft. The seal section—positioned immediately below the motor—contains a series of dynamic seals, lubricant reservoirs, and thrust bearings; its primary function is to isolate the motor from well fluids while balancing pressure across the motor housing. Beneath the seal section lies the multi-stage centrifugal pump assembly, which consists of alternating impellers and diffusers. Each impeller imparts kinetic energy to the fluid, and each diffuser converts that kinetic energy into pressure energy. A check valve is installed at the pump outlet to prevent backflow when the

motor is off, and in many applications, downhole gas-separation or sand-handling modules are added below or above the pump to mitigate issues associated with free gas or solid particulates.

When energized, the motor rotates the pump shaft at speeds typically ranging from 2,400 to 4,800 rpm (depending on motor design, voltage, and frequency), causing each pump stage to pressurize the well fluid. The incremental pressure rise across each stage accumulates to overcome the hydrostatic head of the fluid column and the frictional losses in the tubing, thereby lifting oil (and often associated water) from depths that can exceed 10,000 feet. Because ESPs can generate large differential pressures, they are the artificial-lift method of choice in deep wells, deviated or horizontal completions, offshore platforms with restricted deck space, and reservoirs where high fluid rates must be sustained despite increasing water cuts or declining bottomhole pressures.

Despite their widespread adoption and robust design, ESPs remain susceptible to a variety of failure mechanisms that can abruptly interrupt production and incur significant intervention costs. Common downhole failure modes include:

i. **Electrical Failures:** Insulation breakdown, stator or rotor winding faults, and short circuits within the submersible motor.

ii. **Mechanical Wear:** Erosion or corrosion of impeller/diffuser components due to sand, scale, or corrosive agents; bearing failures in the seal section; thrust-bearing degradation.

iii. **Hydrodynamic Issues:** Gas lock (when free gas accumulates and impairs pump priming), fluid-lubrication starvation, or cavitation when excessively low intake pressure causes vapor-phase formation.

iv. **Seal and Lubrication Failures:** Overpressurization of the seal chamber, lubricant contamination, or seal element rupture can allow well fluids to breach the motor housing.

Intervention to service or replace a failed ESP assembly often requires pulling the entire string out of the hole, a process that can take several days offshore or several hours onshore, depending on well depth and rig availability. Consequently, unscheduled ESP outages represent not only direct equipment replacement costs but also substantial production-loss penalties. Industry surveys have shown that ESP systems account for a significant fraction of total artificial-lift downtime, with average run times varying widely—from a few months in highly abrasive or gas-rich environments to over a year in benign oil-wet, well-managed fields.

In response to these challenges, operators have increasingly turned to predictive maintenance strategies that leverage real-time sensor data to anticipate failure before it occurs. Modern ESP completions commonly include downhole sensors such as:

i. **Pressure Gauges:** Monitoring intake and discharge pressures to infer pump differential pressure and detect developing restrictions or seal-section issues.

ii. **Temperature Sensors:** Placed in the motor housing and seal chamber to track anomalous heat generation indicative of bearing wear, lubrication breakdown, or electrical faults.

iii. **Vibration Accelerometers:** Attached to the motor housing or tubing to detect early signs of mechanical imbalance, misalignment, or bearing degradation.

iv. **Current and Voltage Transducers:** On the surface, these sensors measure real-time motor current and voltage, allowing calculation of electrical power draw, motor torque, and efficiency.

By continuously transmitting this multivariate sensor stream to surface processing units—often via multiplexed cable conductors—operators can build a comprehensive picture of pump health. However, raw sensor readings alone are insufficient to provide reliable failure forecasts. A successful predictive maintenance framework must account for the complex, nonlinear relationships between downhole conditions (e.g., changing fluid composition, temperature gradients, and gas-liquid ratios) and the evolving mechanical and electrical state of the pump.

Machine learning (ML) techniques offer a promising path forward. Supervised learning models—such as random forests, support vector machines, and artificial neural networks—can be trained on historical labeled data (where known ESP run times and failure events are logged alongside their corresponding sensor time series) to learn patterns that precede failure. Unsupervised approaches—such as clustering algorithms and autoencoders—can also detect atypical sensor signatures or novel failure modes for which labeled examples may be scarce. In addition, time-series modeling methods (e.g., long short-term memory neural networks) can capture temporal dependencies and trends in the sensor data, enabling early detection of gradual degradation.

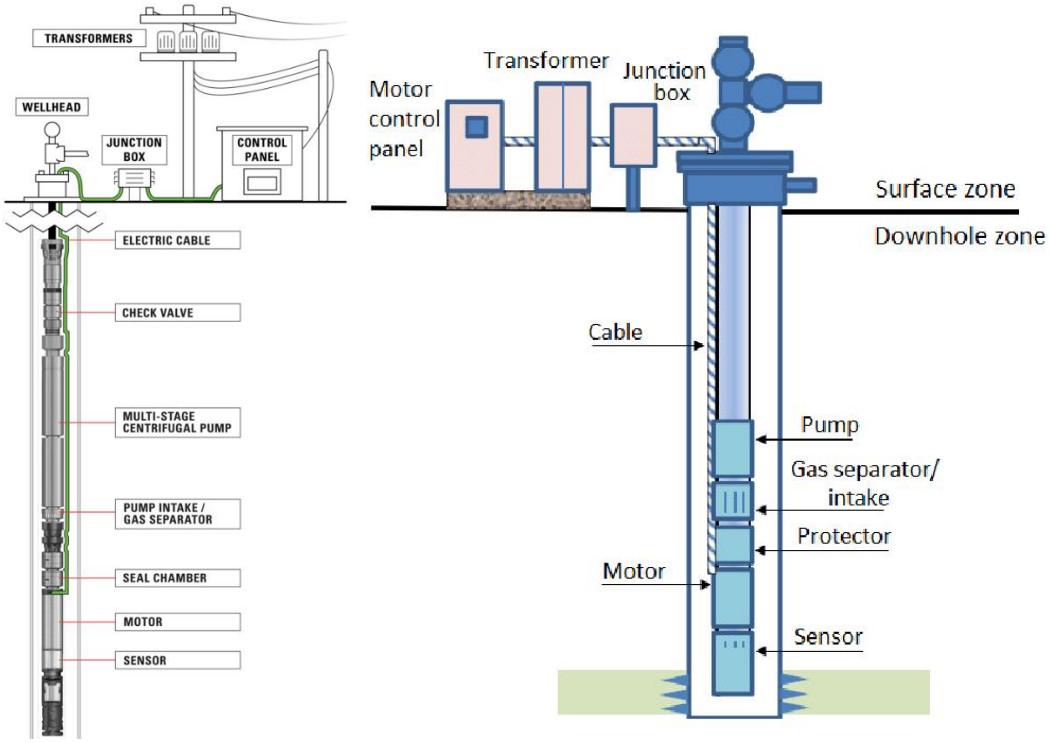


Figure 1.1 and 1.2 : Simplified schematic of an ESP artificial-lift installation, showing surface transformers and controls feeding an armor-protected cable to the downhole pump and motor assembly.

As Schlumberger (2015) explains, ESP systems are built around multiple centrifugal-pump stages that are bolted in series on a common shaft, all driven by a downhole electric motor. Each centrifugal stage consists of an impeller and diffuser pair: the impeller accelerates the fluid radially outward, and the diffuser converts that kinetic energy into increased pressure. These stage geometries are carefully engineered—optimizing blade curvature, throat clearances, and diffuser exit angles—to achieve high efficiency across a range of flow conditions. The motor and pump assemblies are enclosed within a hermetically sealed, pressure-balanced housing. Inside this housing, the motor windings are immersed in a dielectric oil that both insulates and dissipates heat; a series of dynamic seals and thrust bearings in the seal section prevent well

fluids from breaching the motor compartment while simultaneously balancing the external hydrostatic pressure.

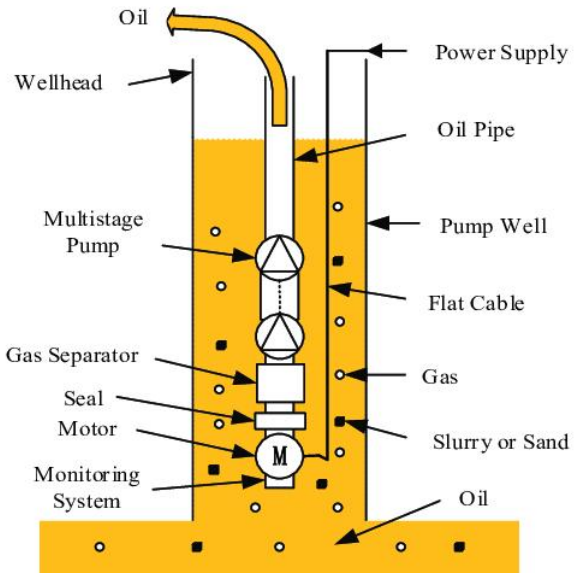
Heavy-duty, multi-core power cables—often armored and with specialized insulation materials—run from surface drives and control panels down to the motor. These cables carry three-phase power (typically 4–6 kV for deepwater or high-temperature installations) and also accommodate telemetry lines that transmit downhole temperature, pressure, and vibration measurements back to surface. Surface variable-frequency drives (VFDs) modulate the voltage and frequency supplied to the motor, allowing operators to adjust pump speed in real time to match changing reservoir conditions (e.g., declining bottom-hole pressure or increasing water cut).

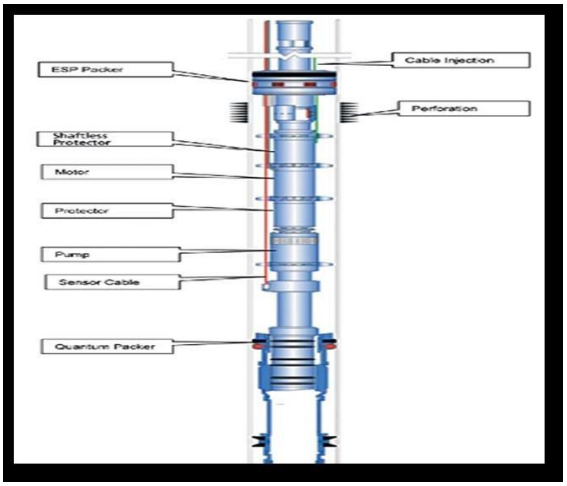
Because ESPs can develop large differential pressures (often exceeding 5,000 psi total across the pump stages) and sustain flow rates from several hundred to over 30,000 bbl/day, they typically outperform alternative pump-type lifts—such as sucker-rod pumps or progressive-cavity pumps—over a broader operating envelope. Sucker-rod pumps, for instance, are limited by rod buckling risks in deviated or deep wells, and progressive-cavity pumps have a narrower viscosity window and are more sensitive to solids. In contrast, ESPs maintain relatively flat efficiency curves over wide flow and pressure ranges, making them the artificial-lift method of choice for high-throughput wells, extended-reach laterals, and subsea templates with limited intervention windows.

Recent advances in ESP materials and component design (Abdalla et al., 2022) have pushed their operating limits even further. High-temperature polymer-composite cables and nickel-based superalloy motor housings now allow ESPs to run reliably in downhole temperatures exceeding 250 °C. Improved elastomers and seal-barrier fluids extend seal-section life in reservoirs with high CO_2 or H_2S concentrations by resisting sour corrosion. In the pump stages themselves,

wear-resistant materials—such as duplex stainless steels or ceramics—are used for impellers and diffusers in sand-laden or erosive environments, while anti-scaling coatings help mitigate paraffin or asphaltene deposition. Furthermore, integrated downhole sensors (e.g., fibre-optic distributed temperature sensing, pressure transducers mounted on multiple stages, and multi-axis accelerometers bonded to the housing) feed high-resolution data to surface analytics platforms, enabling earlier detection of impeller wear, seal-section degradation, or developing gas-lock conditions.

The combination of widespread field deployment—on the order of hundreds of thousands of wells—and the inherent complexity of a multistage, multicomponent assembly underscores why a deep understanding of ESP behavior is critical for optimizing production. Subtle changes in reservoir fluid properties (such as increased gas-to-liquid ratios or shifting viscosity profiles) can alter stage performance curves and motor load characteristics. Likewise, incremental erosion of stage clearances or seal-section wear can manifest as slight upticks in motor current or changes in vibration spectra well before catastrophic failure. Therefore, operators must continuously monitor both the mechanical and electrical parameters of ESPs, leveraging these insights to adjust operating points, plan preventative interventions, and maximize run life while minimizing unplanned downtime.





Despite their operational advantages, ESPs remain vulnerable to a wide array of failure mechanisms—mechanical, electrical, and operational—that stem from the demanding downhole environment. Mechanically, impeller and diffuser wear is one of the most prevalent issues. Abrasive solids (e.g., sand, drilled cuttings, proppant) entrained in the produced fluids erode stage components over time. Even microscopic particles, when accelerated by the high peripheral velocity of impeller blades, can gradually thin blade edges and alter clearances between impeller tips and diffuser housings. This erosion not only reduces hydraulic efficiency (raising motor load for a given flow rate) but can also cause imbalance, leading to increased radial forces on the bearings. Bearings in the seal section—typically hydrodynamic or ceramic thrust bearings—can also fail when contaminated lubricant becomes saturated with particulates or when the downhole lubricant's viscosity is compromised by elevated temperatures. Corrosion, particularly in sour-service wells containing CO_2 or H_2S, can further accelerate material loss from both rotating and stationary pump components, eventually leading to shaft misalignment or catastrophic blade fracture.

Electrical failures account for an even larger share of ESP downtime. The submersible motor windings operate under high-voltage, high-temperature conditions, and insulation breakdown is a common culprit. Thermal aging of insulation materials—compounded by elevated motor-housing temperatures—can cause partial discharges or dielectric failure between winding turns. When downhole temperatures exceed the design threshold of the motor's dielectric oil (often around 175–200 °C for standard fluids), the oil's viscosity decreases, reducing its heat-dissipation capability and further accelerating insulation degradation. Cable faults—such as conductor breaks, water-treeing in the insulation jacket, or armor corrosion—can precipitate open circuits or short circuits. Voltage imbalances (where one phase carries a different load than the others) force excessive current through one or two phases, rapidly heating windings and often leading to motor burnout. Al-Ballam et al. (2022) identified electrical issues as responsible for approximately 61 percent of ESP failures, with motor overload (driven by hydraulic inefficiencies) and gas-locking comprising much of the remaining electrical failure category.

Operational problems—some of which link directly to mechanical and electrical stresses—also contribute heavily to unplanned ESP shutdowns. Gas locking occurs when free gas separates from the liquid column upstream of the intake and coalesces into a gas pocket around the pump inlet. Because centrifugal pumps rely on continuous fluid to impart momentum to impellers, even a relatively small slug of gas can reduce pump intake pressure below the vapor pressure of the fluid, causing vapor-phase cavitation. In a gas-locked state, the impeller stalls or spins with little load, leading to motor "underload" conditions and potentially damaging the motor bearings due to the absence of lubricating fluid. Conversely, fluid influx containing high gas-liquid ratios or liquid slugs can impose sudden spikes in hydraulic horsepower, overloading the motor and tripping overcurrent protections. Hydraulic instability—manifested as flow surges, recirculation

between stages, or abrupt changes in differential pressure—can stem from rapid changes in well productivity (e.g., water-cut increases) or from operating near the pump's "surge line" on its performance curve. Such instability often manifests as pulsating pressure, which exacerbates mechanical fatigue in both the pump's rotating parts and its supporting structural elements.

Seal-section failures, while less frequent than impeller wear, have severe consequences when they occur. The seal section is tasked with maintaining a pressure balance between the high-pressure pump discharge and the ambient wellbore pressure so that well fluids cannot infiltrate the motor compartment. A breach (due to excessive differential pressure, elastomer degradation, or inadequate lubricant viscosity) allows corrosive or particulate-laden fluids to enter the motor housing, potentially resulting in rapid corrosion of winding conductors or rotor bar damage. Lubricant contamination can also accelerate bearing wear and reduce the heat-transfer effectiveness inside the motor.

Scaling and fouling further complicate ESP reliability. In reservoirs containing high concentrations of paraffin, asphaltenes, or inorganic scale precursors (e.g., barium sulfate, calcium carbonate), deposition on impeller blades and diffuser passages can progressively constrict flow passages, raising pump differential pressure. As stage-clearance dimensions shrink, efficiency drops, and motor current climbs correspondingly. Scale buildup can also disrupt the delicate balance between impeller and diffuser, inducing flow recirculation that can erode adjacent surfaces at an accelerated rate. Iron sulfide (FeS) scale—common in sour wells—tends to adhere strongly to metallic surfaces and can cause localized corrosion when exposed to H_2S.

All of these failure modes share a common trait: they often present few overt warnings before a critical fault occurs. A potentiometric survey or temperature anomaly may hint at seal-section distress, but unless sensor thresholds are carefully calibrated, early degradation can go unnoticed.

When a pump finally quits—whether because of a sudden electrical short or cumulative mechanical wear—the well often experiences days or even weeks of downtime. Pulling a failed ESP string requires rig mobilization, workover operations to retrieve and inspect the motor and pump, refurbishment or replacement of damaged components, and then reinstallation and testing. Offshore interventions can be particularly time-consuming, as rig availability and weather windows must be coordinated. Abdalla et al. (2022) note that ESP maintenance "expends a lot of resources" and is predominantly reactive: operators typically respond only after a failure event is detected—by which point production has already been disrupted and potential subsurface damage may have occurred.

Given the confluence of mechanical abrasion, corrosive chemistry, hydraulic variability, and electrical stresses, reducing unplanned ESP failures remains a critical operational challenge. It necessitates a systematic approach to monitoring, diagnosis, and maintenance planning—one that can distinguish innocuous sensor fluctuations from genuine incipient faults. Consequently, the remainder of this work focuses on harnessing machine-learning methods to fuse multivariate sensor streams into reliable prognostic indicators, enabling earlier intervention and extended ESP run life.

Monitoring the downhole conditions of an ESP is essential to diagnosing and preventing failures, and modern completions increasingly rely on a suite of downhole sensors coupled with robust telemetry schemes to provide continuous health-monitoring data. In a typical instrumentation layout, multiple pressure transducers are installed at strategic points—most commonly at the pump intake and discharge—to capture the differential pressure across the pump stages. Since centrifugal pumps operate along a characteristic head–flow curve, any deviation in the expected pump-stage differential can signal impeller erosion, diffuser clogging, or seal-section leaks.

Downhole temperature sensors are often placed both at the motor housing and within the seal chamber. Motor-housing temperature profiles allow detection of abnormal heat generation that may arise from winding insulation breakdown or bearing friction, whereas seal-chamber temperature trends can reveal lubricant degradation or incipient seal failure when the heat signature departs from normal baseline behavior.

Vibration monitoring is frequently achieved through tri-axial accelerometers bonded directly to the motor housing or attached to a nearby carrier. Vibration signatures provide early warnings of mechanical imbalance, misalignment, or bearing wear. For instance, an increase in radial vibration amplitude at a frequency corresponding to the shaft rotational speed can indicate impeller imbalance, whereas elevated axial vibrations near the natural frequency of the pump–motor assembly may suggest looseness in the thrust bearing. Advanced signal-processing techniques—such as fast Fourier transform (FFT) analyses—are often performed in downhole or surface processors to decompose the raw acceleration time series into spectral components, isolating peaks related to specific fault modes (e.g., bearing outer-race faults or shaft misalignment) (Jin et al., 2014).

Some ESP completions incorporate additional specialized sensors. Acoustic flow meters or turbine-wheel flow sensors can be installed near the pump intake to estimate volumetric rate directly; coupling this information with surface-measured tubing pressure profiles allows calculation of real-time fluid density and gas-liquid ratios. In wells prone to sand production, dual-element capacitive or optical sand detectors may be deployed immediately above the pump to quantify solid entrainment, enabling operators to correlate sand volume fractions with erosion rates and schedule remedial actions (e.g., sand screens or gravel packs). In high-temperature or corrosive reservoirs, fibre-optic distributed temperature sensing (DTS) cables can run along the

ESP string to provide continuous temperature profiles with fine spatial resolution, identifying localized hot spots that might correspond to blocked flow paths or gas-pocket accumulation.

Regardless of the specific sensor suite, these measurements must be reliably communicated to the surface. The most established method in ESP completions is current-loop telemetry, wherein analog or digital sensor outputs are modulated as burden-current pulses (often superimposed on the standard 4–20 mA power circuit) that ride up the power cable. Each pulse or sequence of pulses represents discrete bits of sensor data. At the surface, a telemetry decoder interprets the pulse train, reconstructs the original sensor values, and forwards them to supervisory control systems for visualization or archiving (Jin et al., 2014). Because the downhole environment imposes bandwidth constraints—power cables are optimized for power delivery, not high-speed data transmission—the effective data rate of current-loop systems is relatively low. Consequently, engineers must judiciously select sampling intervals and compress or encode multivariate data to ensure critical parameters (e.g., pump differential pressure, motor current, and temperature) are transmitted at a cadence sufficient for timely fault detection.

Alternative communication methods include low-voltage power-line carrier (PLC) systems, in which higher-frequency carrier signals are superimposed onto the power lines without disrupting pump operation. PLC can achieve higher data rates than simple pulse telemetry, enabling the transmission of richer data sets—including raw vibration waveforms or high-resolution acoustic emissions—though at the expense of more sophisticated downhole electronics and increased costs. Wireless telemetry—such as electromagnetic (EM) pulse transmissions through the formation or mud-pulse telemetry—has also been investigated, particularly for applications where cable reliability is a concern or where retrofitting existing installations without running new cables is desirable. These methods, however, can suffer from signal attenuation in

conductive formations and are generally less common in ESP applications due to the complexity of downhole power requirements (Williams et al., 2014).

Once the data stream reaches surface, real-time values of key parameters are visualized on operator dashboards within a SCADA (Supervisory Control and Data Acquisition) or edge-computing platform. Typical dashboards display time-series plots of pump intake and discharge pressures, motor-housing and seal-chamber temperatures, vibration amplitudes in multiple axes, and surface motor current and voltage. By integrating these measurements, engineers can compute derived metrics—such as hydraulic horsepower (calculated from flow rate and pump differential pressure), motor slip (the difference between synchronous and actual rotor speed), and pump efficiency (ratio of hydraulic horsepower to motor input power). Comparing current operating points against manufacturer-provided performance curves enables the identification of off-curve operation, which may precede events like gas locking or mechanical binding.

Real-time monitoring also facilitates closed-loop control strategies. For example, if intake pressure drops below a predefined threshold—potentially indicating a depleted inflow or gas-cutting—a surface controller can adjust the VFD to reduce pump speed, preventing gas lock and cavitation. Conversely, if motor current trends above normal levels at a given flow rate, the system may automatically throttle back speed or raise an alert so that maintenance crews can investigate potential impeller erosion or diffuser scaling. Williams et al. (2014) highlight that direct downhole measurements empower operators not only to “increase production” by optimizing pump speed and effluent rates but also to “diagnose well and ESP performance” and to “achieve protection and control of the ESP system.” By leveraging this instrumentation and telemetry framework, operators gain the situational awareness needed to transition from reactive

maintenance to proactive, condition-based maintenance—ultimately extending ESP run life and minimizing unplanned downtime.

In modern operations, ESP data streams are typically integrated into a facility's SCADA (Supervisory Control and Data Acquisition) architecture, providing a unified view of both surface and downhole conditions. At the surface, instrumentation such as current transformers and voltage transducers measure motor current draw and line voltage, while pressure gauges on the discharge manifold record wellhead pressure. These surface measurements yield a coarse but continuous snapshot of pump load, hydraulic output, and wellhead conditions. However, because they reflect only the conditions at or above the pump intake, they can miss incipient issues downhole—such as the onset of gas-induced cavitation or localized erosion—until those problems manifest as significant changes in motor load or wellhead pressure.

Downhole sensors, in turn, provide the fine-grained fidelity necessary to detect subtle deviations from normal operation. For example, packings of pressure transducers at the pump intake and discharge measure the true differential pressure across the pump stages; a gradual reduction in the measured stage differential at a given flow rate can indicate incipient impeller erosion, diffuser clogging, or seal-section leakage. Temperature sensors mounted on the motor housing and within the seal chamber track thermal gradients that might signal bearing wear, lubricant breakdown, or insulation degradation before they drive motor current spikes. Tri-axial accelerometers bonded to the motor housing yield vibration spectra in real time; by performing onboard or surface-based FFT (Fast Fourier Transform) analysis, operators can isolate frequency peaks associated with specific fault modes (e.g., bearing outer-race defects at bearing-pass frequencies, or impeller imbalance at shaft rotational frequency) long before those faults cause catastrophic failure. Kimray (2023) emphasizes that these downhole measurements—when

communicated continuously via telemetry—provide "real-time system data such as pump intake and discharge pressures, temperatures, and vibration," which SCADA controllers use to adjust operating parameters or raise alarms.

Telemetry methods vary by installation and cost constraints. The most common approach is a current-loop telemetry system in which downhole sensor outputs are encoded as digital pulse bursts superimposed onto the 4–20 mA power cable; each pulse train corresponds to discrete sensor values that are decoded by a surface telemetry module. While current-loop telemetry is relatively low-bandwidth—limiting effective sampling rates to on the order of seconds or minutes—it is robust and compatible with existing power cables. In some high-end installations, low-voltage power-line carrier (PLC) systems are used, enabling higher data-throughput so that raw vibration waveforms or high-resolution acoustic emissions can be transmitted; however, these systems require more complex downhole electronics and come at increased capital and maintenance expense. Regardless of method, the goal is always the same: to feed continuous, time-synchronized data into the SCADA historian, where it can be stored, visualized, and processed in real time.

By fusing surface and downhole data streams, operators obtain a more comprehensive hydraulic and mechanical perspective on well performance. For instance, a downward drift in discharge pressure recorded at the pump outlet—when correlated with a stable surface pressure reading—can pinpoint formation damage or tubing blockage just above the pump intake. Likewise, a slow rise in motor-housing temperature, coupled with a nominal motor current, may indicate early lubricant degradation in the seal chamber even if the surface power draw appears unchanged. In contrast, relying solely on surface electrical measurements—motor current and voltage—often masks these early indicators: for example, gas ingestion (which reduces pump intake density) can

maintain motor load near nominal values until the pump eventually gas-locks, causing a sudden motor underload.

Al-Ballam et al. (2022) demonstrated the power of integrated monitoring by analyzing five years of synchronized surface and downhole sensor records from multiple wells. They applied statistical trend analysis to identify the root causes of past ESP failures—distinguishing, for instance, whether an uptick in motor current was driven by stage erosion (inferred from declining downhole differential pressure) or by gas flooding (inferred from erratic downhole pressure swings). Once failure modes were categorized, they fitted Weibull reliability models to the time-to-failure data for each failure category. In a Weibull analysis, the probability density function is defined as

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \qquad Equation\ 1$$

where t is the run-time to failure, β is the shape parameter (indicating whether failure rate is increasing or decreasing over time), and η is the scale parameter (a characteristic life). By estimating β and η for electrical, mechanical, and operational failures separately—often using censored-data techniques to account for still-running pumps—they derived failure-probability curves that predict the likelihood of a given fault within a specified future interval. This probabilistic insight then informs preventive maintenance scheduling: for example, if the Weibull curve for seal-section failures indicates a 20 percent probability of seal leakage within the next 300 operating days, the operator might plan a workover before that threshold to replace the seal components.

The integration of downhole sensor data with surface SCADA systems transforms ESP management from a reactive exercise—where operators simply respond to pump outages—into a

proactive, condition-based maintenance framework. Real-time visibility into pressures, temperatures, and vibration signatures not only enables more accurate diagnosis of incipient faults but also supplies the empirical basis for statistical and machine-learning models that predict failures before they occur. Consequently, this holistic data-fusion approach underpins efforts to maximize ESP run life, minimize unplanned downtime, and optimize overall well performance.

Predictive maintenance has emerged as a key strategy in the oil and gas industry to leverage this abundance of sensor data. Unlike traditional preventive maintenance (which is calendar or usage-based) or reactive maintenance (fixing only after breakdown), predictive maintenance (PdM) uses analytical models to anticipate equipment issues before they occur. In predictive maintenance, historical and real-time sensor data are continuously analyzed—often with advanced statistics or machine learning—to assess the equipment's health and forecast future failure risk (IBM, 2023). In effect, PdM aims to answer the question: when should we service or replace this equipment to avoid an unplanned failure? A well-tuned PdM system can significantly reduce unnecessary downtime and maintenance cost by performing maintenance only when the data indicate a real need. IBM (2023) describes PdM as "continually assessing [equipment] health in real time" via sensors and machine learning, so that issues can be "identified, detected, and addressed as they occur" or even before they manifest. By providing the right information at the right time, predictive maintenance complements preventive schedules and enables a more cost-effective maintenance program.

It is estimated that over 90 percent of producing oil wells worldwide employ some form of artificial lift (Schlumberger, 2015). Among these, ESPs distinguish themselves by delivering exceptionally high fluid throughput—ranging from a few hundred barrels per day (bbl/d) to

upwards of 30,000 bbl/d—making them indispensable for high-productivity wells, deep or deviated completions, and offshore installations where deck space is at a premium. According to industry surveys, ESP systems are installed in on the order of 150,000–200,000 wells globally, underpinning a substantial fraction of total oil production (Schlumberger, 2015).

An ESP assembly comprises several key subcomponents arranged in series along a common shaft: a hermetically sealed submersible electric motor, a seal section, the multistage centrifugal pump section, and, in many applications, auxiliary modules such as downhole gas separation or sand management devices (Fig. 1). The motor section typically consists of three-phase stator windings immersed in a dielectric lubricant. A series of dynamic seals and thrust bearings within the seal section maintain pressure balance and keep well fluids from intruding into the motor housing. Below the seal section, a stacking of centrifugal pump stages—each composed of an impeller and a corresponding diffuser—pressurizes reservoir fluids, sequentially raising their pressure to overcome the combined hydrostatic head and frictional losses in the production tubing. When energized, the motor spins at high speeds (commonly 2,400–4,800 rpm, depending on design voltage and frequency), driving each impeller to accelerate fluid radially outward; the ensuing kinetic energy is converted to pressure across the diffuser. A check valve near the pump discharge prevents backflow when the motor is shut off. In engineering practice, stage geometries—including blade curvature, diffuser vane angles, and throat clearances—are optimized via computational fluid dynamics (CFD) during design to achieve peak efficiency across a specified flow window. Wide-flow, high-differential stages enable a single ESP string to handle diverse reservoir conditions, from high-viscosity fluids to multiphase mixtures containing free gas (Fakher et al., 2021).

Heavy-duty, multi-core power cables (often armored and rated for high temperatures) descend from surface variable-frequency drives (VFDs) to energize the downhole motor. The VFDs not only supply 4–6 kV three-phase power (for deepwater or high-temperature applications) but also enable real-time speed modulation, allowing operators to adapt pump performance to changing inflow conditions—for example, reducing pump speed to prevent gas lock when gas-oil ratio increases. In hostile well environments, materials such as nickel-based superalloys and high-temperature elastomers are used to construct motor housings, dynamic seals, and pump components, extending the ESP's tolerance to downhole temperatures exceeding 250 °C and corrosive fluids containing CO_2 or H_2S (Abdalla et al., 2022). Wear-resistant alloys or ceramic coatings on impeller and diffuser surfaces mitigate erosion from sand or solid particulates, while anti-scaling treatments reduce paraffin and asphaltene buildup. Furthermore, integrated downhole sensors—pressure transducers, temperature gauges, and vibration accelerometers—feed critical data back to surface analytics platforms, enabling near-continuous monitoring of ESP health.

Despite these design advances, ESPs remain prone to failures that can abruptly suspend production and impose substantial intervention costs. Common mechanical failure modes include impeller/diffuser erosion, bearing degradation in the seal section, and seal-chamber breaches, often exacerbated by sand or solids entrainment, corrosive chemistries, or high cyclic loads. For instance, even sub-millimeter-scale abrasives—entrained in high-velocity flows—can erode impeller blade surfaces over time, altering hydraulic clearances and diminishing pump efficiency. When clearances widen, fluid slip increases, necessitating higher motor torque to maintain flow. Simultaneously, imbalance from uneven erosion amplifies radial forces on bearings, leading to accelerated wear or catastrophic bearing seizure. Meanwhile, corrosion of impeller or diffuser

metal—especially in sour wells—can produce pitting that rapidly transitions to crack propagation under cyclic stresses (Fakher et al., 2021).

Electrical failures—cumulatively accounting for roughly 61 percent of ESP downtimes—arise from insulation breakdown, stator winding faults, or cable defects (Al-Ballam et al., 2022). Over time, dielectric oils degrade at elevated temperatures (>175 °C), reducing their heat-dissipation capacity and precipitating partial-discharge phenomena within motor windings. Thermal aging of insulation materials—compounded by gas pockets or fluid contamination in the seal section—further accelerates breakdown. Voltage imbalance (a difference in load between motor phases) increases current draw on heavily loaded phases, causing hot-spots that exceed dielectric strength. Cable jacket "water-treeing" or armor corrosion can lead to conductor shorts or opens, immediately interrupting power delivery.

Operational issues such as gas locking and hydraulic instability exacerbate mechanical and electrical stresses. Gas locking occurs when free gas accumulates at the pump inlet, causing cavitation and momentarily "unloading" the impellers. In this state, the motor can overspeed, leading to shaft deflection, bearing damage, or thrust-bearing overload. When the gas pocket clears, the pump re-primes with a sudden influx of liquid, generating torque spikes that can further damage mechanical elements. Hydraulic instability—manifesting as surging, recirculation between stages, or transitions across the pump's instability region on its head–flow curve—induces unsteady loads that shorten component lifetime. Seal section breaches, which allow well fluids to invade the motor housing, can instantly convert the downhole environment into an aggressive corrosion cell, rapidly degrading windings and thrust bearings. Such failures typically offer few clear precursors; unless trends in sensor data (e.g., a rising seal-chamber

temperature) are scrutinized, the only warning might be a sudden drop in motor current or a rapid increase in torque.

Interventions to retrieve and replace a failed ESP string are time- and resource-intensive. Mobilizing a workover rig, performing a downhole intervention (which may involve fishing for stuck cable heads or removing scale plugs), and re-running a new or refurbished ESP assembly can easily take days—or weeks in offshore settings constrained by vessel schedules and weather windows. During this period, the well is offline, directly costing lost hydrocarbon revenue. Al-Ballam et al. (2022) estimate that even a single midwater ESP failure can translate to millions of dollars in lost production for offshore developments. Abdalla et al. (2022) note that most maintenance today is reactive, triggered only after a fault is detected—often by a sudden surface alarm—rather than based on incremental degradation insights.

To reduce these unplanned failures, operators have increasingly deployed condition-based monitoring frameworks that combine downhole sensor data with surface SCADA measurements. At the surface, motor current and line voltage transducers, together with discharge pressure gauges on the production manifold, provide a coarse snapshot of pump operating point. Downhole, pressure gauges at the pump intake and discharge deliver the true differential pressure across pump stages; temperature sensors on the motor housing and within the seal chamber register thermal anomalies; tri-axial accelerometers detect vibration signatures associated with imbalance or bearing distress; and, in some advanced completions, fibre-optic distributed temperature sensing (DTS) cables map temperature gradients along the entire ESP string, identifying local hot spots or fluid channeling (Jin et al., 2014; Williams et al., 2014; Kimray, 2023). These data streams are relayed to surface via methods such as current-loop telemetry—where digital pulses encoded on the 4–20 mA power cable convey discrete sensor

readings—or, in higher-end installations, low-voltage power-line carriers (PLC) that support greater bandwidth for raw waveform transmission. Once at surface, the data merge into a SCADA historian, where operators can visualize real-time trends, compute derived metrics (e.g., hydraulic horsepower, motor slip, pump efficiency), and set automatic or manual alerts when parameters cross pre-defined thresholds.

However, simple threshold alarms often prove inadequate for anticipating failures well in advance. For example, a slight uptick in motor current could originate from subtle impeller erosion, early bearing wear, an increase in fluid viscosity, or an uptick in gas fraction; distinguishing among these causes requires multivariate analysis. Similarly, transient vibrations at shaft-speed frequency might reflect a momentary tubing bump or a nascent blade-tip rub; the same spectral feature could also occur during normal startup if not filtered properly. Al-Ballam et al. (2022) demonstrated that multi-year datasets—encompassing both surface and downhole records—can reveal hidden correlations between operational changes (e.g., ramping pump speed by 5 percent) and subsequent failure onset. By applying statistical analyses and reliability models (e.g., Weibull distributions) to categorize time-to-failure under different root causes (electrical, mechanical, or gas lock), they derived failure-probability curves that could inform preventive workover scheduling more precisely than simple run-time thresholds.

This confluence of rich sensor data and the need for early fault detection has driven the adoption of predictive maintenance (PdM) strategies in petroleum engineering. Where traditional preventive maintenance relies on fixed calendar or usage intervals—servicing an ESP every six months regardless of condition—and reactive maintenance addresses failures only after they occur, PdM seeks to “continually assess equipment health in real time” (IBM, 2023). By mining both historical archives and live data streams through advanced statistical methods or machine

learning (ML) algorithms, PdM answers the fundamental question: *When should we service or replace this equipment to avoid an unplanned failure?* In essence, PdM focuses on performing maintenance "just-in-time," thereby minimizing unnecessary interventions and maximizing run life.

In other industries (e.g., aviation, manufacturing), PdM has proven capable of reducing unplanned downtime by 30–50 percent and lowering maintenance costs by 20–40 percent. In oilfields, where a single day offline can cost tens of thousands of dollars in production revenue (depending on well productivity), PdM can yield significant economic impact. Abdalla et al. (2022) present a data-driven PdM workflow tailored to ESPs: they applied principal component analysis (PCA) to reduce the dimensionality of a high-frequency sensor dataset—comprising downhole pressures, temperatures, surface current, and vibration amplitudes—to a few orthogonal health components. These components then served as inputs to an XGBoost classifier, which learned to distinguish "normal operation" from "incipient failure" states. Their model flagged deep-seated pump anomalies up to seven days before actual failure, achieving a high recall while keeping the false-alarm rate below 10 percent. Likewise, Al-Ballam et al. (2022) leveraged multi-year ESP records to identify statistically significant precursors—such as gradual slope changes in pump differential pressure or vibration sideband energy increases—long before traditional alarms would trip. These studies illustrate how machine learning can uncover complex, nonlinear relationships among pressure, temperature, and vibration trends—patterns that simple threshold-based methods often miss.

Machine learning methods—ranging from supervised classification and regression to unsupervised anomaly detection—offer the ability to learn these patterns directly from data, even when explicit physical models are difficult to formulate. Supervised algorithms (e.g., random

forests, support vector machines, artificial neural networks, or gradient-boosted trees) train on historical examples labeled by failure mode (e.g., electrical burnout, seal breach, gas lock) and learn decision boundaries in multidimensional feature space. Regression-based ML (e.g., support vector regression or deep-learning architectures like long short-term memory, LSTM) can predict Remaining Useful Life (RUL) by mapping current sensor features to a time-to-failure estimate. Unsupervised techniques—such as clustering (k-means, Gaussian mixture models) or autoencoder-based anomaly detectors—learn the "normal" manifold of pump operation; when incoming data diverge significantly, they flag anomalies even in the absence of prior failure labels. Hybrid models, which combine statistical reliability curves (e.g., Weibull or Cox models) with ML-driven health indices, can further refine probabilistic failure forecasts.

The economic stakes in the petroleum context motivate a thorough investigation into ML-based PdM for ESPs. Well downtime directly impacts production revenue, contract penalties, and the cost of mobilizing intervention rigs—especially offshore. By leveraging sensor fusion (integrating downhole pressures, temperatures, and vibration with surface electrical and hydraulic measurements), ML models can detect subtle deviations from nominal performance days—or even weeks—ahead of catastrophic failure. This lead time offers a critical window for planning workovers, ordering parts, and scheduling rig time under favorable conditions, thus minimizing the total cost of ownership (TCO) of ESP assets.

1.1 Problem Statement

Electric Submersible Pumps (ESPs) play a pivotal role in sustaining hydrocarbon production from wells whose natural drive has diminished. However, ESP systems are predisposed to a variety of failure modes—mechanical wear, gas locking, electrical faults, and hydraulic

instabilities—that often arise without warning. Unplanned ESP shutdowns trigger costly workovers, extended downtime, and safety risks in the well-intervention process. Traditional maintenance strategies, whether calendar-based preventive schedules or reactive repairs post-failure, fail to fully capitalize on increasingly available real-time sensor data. As a result, operators continue to incur unnecessary expense and lost production because existing approaches lack the intelligence to predict incipient faults early enough for planned interventions. There is, therefore, a pressing need for a data-driven predictive maintenance framework that leverages ESP telemetry to forecast equipment health and optimize workover scheduling.

1.2 Aim and Objectives of Study

The primary aim of this study is to develop and validate a machine-learning–based predictive maintenance system for Electric Submersible Pumps in petroleum production. The primary objective of this thesis is to design, implement, and evaluate machine-learning algorithms for predictive maintenance of ESP systems using real-time and historical sensor data. Specifically, the study seeks to:

1. Characterize ESP Sensor Data: Assemble a comprehensive dataset from downhole and surface sensors—capturing pump intake and discharge pressures, motor and seal-chamber temperatures, tri-axial vibration signals, surface motor current, and line voltage. We will explore data quality considerations, sensor calibration, alignment of multivariate time series, and the preprocessing steps required to remove noise and outliers (e.g., filtering cable-related spikes or compensating for telemetry latency).
2. Develop Feature Engineering Techniques: Design domain-informed metrics—such as hydraulic horsepower, motor slip, stage efficiency estimates, vibration spectral features,

and temperature gradients—and evaluate their sensitivity to various failure modes. Feature-selection methods (e.g., recursive feature elimination, mutual information, or principal component analysis) will be applied to identify the most predictive variables while minimizing dimensionality to reduce computational complexity.

3. Construct and Compare ML Models: Implement a range of supervised, unsupervised, and hybrid models, including random forests, support vector machines, gradient-boosted decision trees (e.g., XGBoost), LSTM networks for time-series forecasting, and autoencoder-based anomaly detectors. We will train and validate these models on labeled historical data, carefully partitioning data in a time-series–consistent manner to avoid leakage (e.g., using walk-forward cross-validation). Performance metrics will include classification accuracy, precision, recall, F1-score, ROC-AUC for binary classification, and MAE/RMSE for RUL estimation.

4. Evaluate Model Deployment and Lead Time: Assess the timeliness of each model's failure predictions by measuring lead time—the interval between model alert and actual failure—and balancing it against false-alarm rates. We will also explore survival-analysis techniques (e.g., Weibull parameter estimation, Cox proportional hazards) to derive probabilistic failure forecasts that update dynamically as new data arrive.

5. Demonstrate Economic Impact: Quantify the potential reduction in unplanned downtime and maintenance costs by simulating PdM-driven intervention strategies. By comparing reactive, preventive, and predictive maintenance scenarios—using historical failure logs and known workover schedules—we will estimate potential cost savings, increased production uptime, and reduced equipment run-time variability.

1.3 Justification of the Research

The economic and operational stakes of unplanned ESP failures in petroleum production are substantial. Workovers to retrieve and replace a damaged downhole pump not only incur direct costs—rig time, specialized personnel, and new equipment—but also result in prolonged production losses that can amount to thousands of barrels of oil per day. Furthermore, reactive maintenance regimes often rely on broad safety margins or rigid preventive schedules, leading to unnecessary interventions when pumps remain healthy, or conversely, to catastrophic breakdowns when the schedule fails to anticipate a fault. Meanwhile, modern wells are increasingly instrumented with downhole and surface sensors that generate a continuous stream of real-time data, yet this wealth of information frequently goes under-utilized. By introducing a rigorous, machine-learning–driven predictive maintenance framework, this research aims to transform raw telemetry into actionable foresight, enabling operators to plan workovers precisely when needed—minimizing downtime, lowering maintenance costs, and enhancing overall production efficiency. This study therefore addresses a critical gap between data availability and decision-making in artificial-lift operations, offering a methodical approach to harnessing sensor data for anticipatory ESP maintenance.

1.4 Scope of the Research

This thesis concentrates exclusively on Electric Submersible Pumps installed in onshore and offshore oil wells within the petroleum engineering domain. The work is limited to analyzing time-series sensor data that are typically available in field deployments: electrical signals (motor current and leakage), hydraulic measurements (intake and discharge pressures, total hydraulic pressure), thermal readings (intake and motor temperatures), and vibration/frequency data. Production rate and choke-setting information are included where reliably recorded, but

ancillary well-log data (such as formation pressures or gas-oil ratios) lie outside the present scope. The study compares both supervised classifiers (Support Vector Machines, Random Forests, XGBoost) and unsupervised anomaly detectors (Gaussian Mixture Models, LSTM autoencoders) for their effectiveness in forecasting ESP failures. Model development and evaluation are confined to a single multi-year dataset sourced from a representative field installation, and the workover history documented for that field serves as the ground-truth for "pre-failure" labeling. Broader generalization across diverse reservoir types or artificial-lift configurations (e.g., rod pumps, gas lifts) is acknowledged as future work but is not covered in this research. Finally, while economic analysis of cost savings and operational implementation guidelines are discussed qualitatively, detailed financial modeling and field-scale deployment logistics are beyond this thesis's boundaries

CHAPTER TWO

LITERATURE REVIEW

2.1 Traditional Approaches to ESP Monitoring and Maintenance

Electric submersible pumps (ESPs) have historically been monitored through a combination of surface instrumentation, periodic manual inspections, and rule-based alarm systems integrated into SCADA (Supervisory Control and Data Acquisition) platforms. At the most basic level, surface-mounted sensors—namely, current transformers, voltage transducers, and wellhead pressure gauges—provide a coarse indication of pump health. Operators typically plot motor current (ammeter charts) over time to observe characteristic signatures: for example, a gradual rise in current at constant flow rate may indicate impeller wear or diffuser clogging, while a sudden spike followed by a drop (motor underload) can signal gas lock or mechanical jamming (Bansal et al., 2020). In parallel, nodal analysis—using inflow performance relationships (IPR curves) and vertical lift performance (VLP) curves—helps diagnose well-bore or reservoir issues that manifest as changes in pump differential pressure. By comparing expected pump head (calculated from manufacturer performance curves) against measured wellhead pressures, engineers attempt to infer downhole conditions without direct sensor feedback.

Despite its ubiquity, this approach often proves reactive rather than predictive. SCADA threshold-alarm logic is largely static: fixed upper and lower limits are set for parameters such as motor temperature, differential pressure, and vibration amplitude. For instance, motor-housing temperatures above a predetermined value (e.g., 120 °C) trigger an overheat alarm, prompting immediate shutdown. However, because these thresholds do not adapt to operational context—such as varying ambient temperatures, startup transients, or changes in fluid viscosity—false

positives are common. Takacs (2017) observed that static temperature thresholds during ESP startup frequently confuse normal inrush currents with genuine overheating, resulting in up to 30 percent of sensor-triggered shutdowns being unwarranted. Similarly, Bansal et al. (2020) reported that approximately 65 percent of unplanned ESP shutdowns in a North Sea field were precipitated by sensor noise or transient fluctuations rather than actual mechanical or electrical faults. In their study, many false alarms arose from cable interference or rapid pressure reversals during shifting reservoir dynamics, which the rigid threshold settings could not contextualize.

Human interpretation remains central in traditional frameworks. When an alarm trips, operators must sift through ammeter trends, discharge-pressure histories, and occasionally logbook annotations to diagnose the root cause. Cognitive overload becomes a real concern in operations managing large ESP fleets: a 2021 SPE study found that operators handling more than 50 active ESPs missed up to 40 percent of critical alerts during peak production periods, often because simultaneous alarms from multiple wells overwhelmed their capacity to triage and respond (SPE, 2021). Operators subsequently rely on heuristics—such as “if motor current rises by more than 10 percent at constant flow” or “if discharge pressure drops by over 15 percent within five minutes”—to decide whether to initiate a workover or wait for corroborating evidence. However, these heuristics can be ambiguous; for example, a 10 percent rise in current could stem from slight impeller erosion, higher fluid viscosity due to temperature shifts, or increased gas-oil ratio. Without downhole context, misdiagnoses are frequent, leading to unnecessary interventions or missed failure precursors.

Mechanical inspections and routine workovers serve as another pillar of traditional maintenance. Operators often follow calendar-based schedules—replacing or overhauling ESP assemblies after a fixed number of operating hours (commonly 6,000–8,000 hours)—regardless of actual

equipment condition. During these interventions, visual inspection of impeller wear rings, measurement of bearing clearances, and electrical testing (insulation resistance tests, winding-resistance checks) help identify components nearing end-of-life. While such preventive actions can forestall catastrophic failures, they are still fairly coarse: a pump might be pulled at 6,500 hours even if its impellers and bearings remain within acceptable wear tolerances, thereby shortening run life and increasing maintenance frequency. Conversely, waiting until visual evidence of wear can miss subtle degradation paths, such as early-stage seal-chamber lubricant breakdown, which manifests thermally but may not be visible until a seal leak occurs.

Traditional diagnostics include vibration analysis conducted at the surface. By mounting accelerometers on the wellhead or tubing hanger, operators perform periodic spectrum analyses—looking for amplitude peaks at shaft-speed or bearing-pass frequencies. However, surface-mounted vibration sensors capture attenuated signals that have propagated through tens or hundreds of feet of tubular string, blurring fault signatures and complicating root-cause identification. Without downhole accelerometers, distinguishing between rod pump harmonics, surface equipment resonance, and true downhole imbalance becomes a matter of educated guesswork.

Legacy nodal analysis techniques rely on a detailed understanding of reservoir and completion characteristics. Engineers combine reservoir pressure-transient analysis, well inflow modeling, and tubing/casing friction correlations to estimate when changes in required pump head result from reservoir depletion or wellbore scaling versus pump-stage degradation. But these methods demand extensive reservoir simulation models, periodic well tests (e.g., step-rate tests, spinner surveys), and laboratory analyses of fluid properties (e.g., viscosity, gas-liquid ratio).

Consequently, real-time diagnostics remain out of reach, and any detected anomalies typically surface only after production rates have declined significantly.

Traditional ESP monitoring and maintenance methods—while foundational—are characterized by:

i. **Threshold-based Alarms:** Static setpoints on motor temperature, current, and pressure, which yield high false-positive rates and limited sensitivity to evolving fault patterns (Takacs, 2017; Bansal et al., 2020).

ii. **Human-Centered Interpretation:** Reliance on operators to contextualize alarms and ammeter chart signatures, leading to missed alerts and inconsistent diagnoses under high workload (SPE, 2021).

iii. **Reactive or Calendar-Based Interventions:** Workovers initiated either after alarms trip or strictly by run-time quotas, resulting in either premature pulls (increased costs) or delayed detection (unplanned downtime).

iv. **Surface-Only Diagnostics:** Dependence on motor current, wellhead pressure, and surface vibration data, which lack the fidelity to detect early downhole faults such as seal-section wear, incipient gas lock, or subtle impeller imbalance.

v. **Labor-Intensive Inspections:** Periodic pull-and-inspect procedures requiring rig time, visual wear measurements, and offline electrical tests—practices that are time-consuming, expensive, and provide only snapshots of equipment health.

Recognizing these limitations, researchers and operators have turned to data-driven and machine-learning techniques to augment or replace traditional methods. For instance, Abdalla et

al. (2022) employed principal component analysis (PCA) to distill high-dimensional real-time ESP sensor data into a small set of health-indices, and then used XGBoost classifiers to forecast imminent pump failures. Their approach flagged deep-seated anomalies up to seven days before shutdown, demonstrating a clear advantage over static threshold schemes. Such data-driven predictive maintenance frameworks aim to overcome the blind spots of conventional SCADA alarms, reduce false positives, and provide actionable lead time for workover planning. These developments set the stage for the next section, which surveys emerging PdM methodologies and their application to ESP systems.

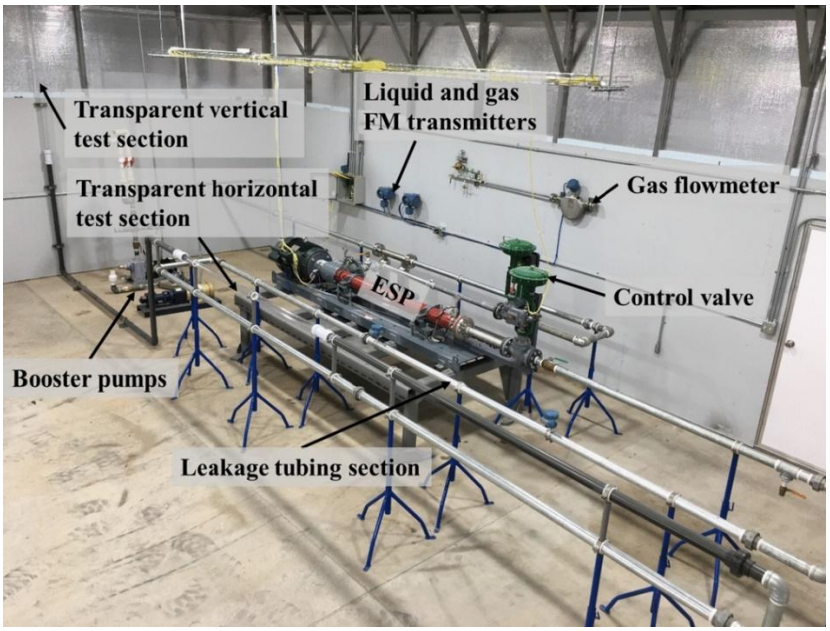


Figure 2.1: A Booster Pump

2.2 Laboratory-Scale Studies and Feature-Extraction Approaches

Song et al. (2024) developed a detailed laboratory-scale test facility to simulate ESP operation under controlled conditions. Their rig incorporated a booster pump that delivered a two-phase gas–liquid mixture into a transparent section upstream of a commercial ESP, allowing visual observation of flow behavior. High-accuracy flow transmitters measured liquid and gas rates,

while pressure sensors recorded intake and discharge pressures at high frequency. The pump itself was instrumented with tri-axial vibration accelerometers mounted on the motor housing and pump casing, acoustic-emission sensors affixed to the pump body, and temperature probes placed in both the motor windings and seal chamber. An array of solenoid-driven valves upstream permitted rapid adjustments to flow rate and gas-to-liquid ratio, simulating real-world reservoir transients. Downstream, a separator vessel collected produced fluids and controlled gas venting. All data channels were synchronized via a GPS-disciplined clock, ensuring millisecond-level alignment over the 300-day experimental campaign.

During this period, the researchers induced two distinct failure modes. First, they simulated seal-section lubricant degradation by incrementally heating the motor housing and starving the seal chamber of fresh lubricant. Early indicators of seal failure included a gradual rise in seal-chamber temperature relative to the motor housing and a modest increase in motor current while flow remained constant. Second, they forced impeller cavitation and mechanical fatigue by lowering inlet pressure below the fluid's vapor pressure, triggering repeated cavitation cycles at operating speed. Prior to visible blade damage, acoustic-emission counts in a specific high-frequency band increased dramatically, signaling impending mechanical fatigue. Over the course of the experiment, eleven different time-series signals were collected—ranging from motor current and supply voltage to pressure, flow, temperature, vibration, and acoustic measurements—each sampled at rates appropriate to capture its dynamics.

To prepare the data for machine learning, Song et al. first aligned all signals to a one-second timestamp, downsampling high-rate channels and interpolating low-rate channels as needed. They applied band-pass filters suited to each signal type: low-frequency filtering for pressure and flow to remove pump-stroke harmonics, intermediate-frequency filtering for vibration to isolate

shaft-speed signatures, and high-frequency filtering for acoustic emissions to target crack-initiation events. Every minute, domain-informed features were computed, including normalized changes in pump differential pressure, motor slip (derived from supply frequency and current harmonics), spectral energy in vibration bands associated with shaft imbalance and bearing defects, counts of high-frequency acoustic pulses, and the temperature difference between seal chamber and motor housing.

After labeling each minute as either “Normal” or “Pre-Failure” (with a one-hour window leading up to an induced failure), they reduced the eleven-dimensional feature set using principal component analysis (PCA). The first three principal components captured the majority of the variance—one dominated by pressure and slip changes, another by vibration energy, and the third by temperature gradients and acoustic activity. Projecting the data onto these three components yielded a compact representation, which they fed into a simple logistic regression classifier. Using 70 percent of the data for training and 30 percent for testing (stratified by failure events), the PCA-based model achieved 93.3 percent accuracy, with high precision and recall on the Pre-Failure class. Only a small fraction of true failure windows were missed, and false alarms were comparatively rare. This study highlights how unsupervised feature extraction can distill complex, multivariate sensor streams into a concise health signature, enabling reliable failure detection even when labeled failure instances are limited.

2.3 Early Predictive Maintenance Techniques

The origins of predictive maintenance (PdM) for ESPs in the 1990s centered on physics-based degradation models that explicitly represented wear and failure mechanisms. For example, impeller erosion models estimated the cumulative volume loss of pump stages by considering sand concentration, impeller speed, and fluid properties; seal-section leak models described

polymer degradation in dynamic seals as a function of temperature, pressure, and elastomer chemistry; and bearing fatigue life models calculated expected bearing life based on dynamic-load ratings, applied loads, and rotational speed. Although these approaches provided theoretical insights, they faced severe limitations in field deployment. Jardine et al. (2006) demonstrated that discrepancies between lab-derived erosion rates and actual field conditions—such as different sand particle-size distributions and the presence of multiphase flow transients—led to prediction errors exceeding 35 percent in offshore ESP deployments. Seal-degradation models similarly fell short because they did not account for contaminants like chlorides or hydrogen sulfide, which accelerated elastomer aging by large margins. Consequently, purely physics-based PdM models proved unreliable without extensive calibration data rarely available in heterogeneous reservoir environments.

By the early 2010s, hybrid approaches emerged that combined simplified physics constraints with rudimentary statistical or machine-learning elements. One popular strategy was rule-based expert systems, which codified heuristics drawn from field experience. For instance, operators might instruct the system to flag impeller erosion if motor current rose by a certain percentage while pump differential pressure simultaneously declined. Alternatively, a seal problem would be signaled if seal-chamber temperature exceeded motor-housing temperature by a fixed margin for a predefined duration. While such rules reduced obvious false alarms compared to single-threshold methods, they remained brittle—requiring frequent tuning as pump designs, fluids, and operating conditions changed.

Simultaneously, basic statistical trend models used linear or polynomial regression to relate moving averages of motor current, pressure, and temperature to remaining run life. However, these linear approaches could not capture the nonlinear interplay among features—for example,

how increasing gas content would affect motor current differently depending on temperature and flow rate—limiting their accuracy and generalizability.

A key milestone in hybrid PdM was the work of Heng et al. (2019), who applied gradient-boosted decision trees (XGBoost) to correlate surface motor-current signatures with bearing wear in field-installed ESPs. From one-second recordings across an entire year of operation on fifty ESPs, they derived features such as the crest factor (peak-to-RMS ratio), kurtosis (impulsiveness indicator) of the current waveform, and motor slip. By labeling each pump run as "Healthy" until roughly 150 hours before bearing replacement and "Imminent Failure" during the final 150 hours—based on corroborating vibration trends—they trained an XGBoost classifier that cut unplanned downtime by 30 percent compared to the prior year. The model also provided a median lead time of about three days before failure, allowing maintenance to be scheduled in a planned fashion. Although false alarms occurred at a moderate rate (around 12 percent), this was considered acceptable given the high cost of a bearing seizure.

Despite these advances, early hybrid techniques still struggled with data heterogeneity, limited feature sets, and dependency on labeled failure events. Variations in ESP design, fluid chemistry, and operational practices across fields meant that models trained in one environment often underperformed elsewhere. Sparse sensor suites—typically restricted to surface measurements or a handful of downhole readings—missed richer signals such as full vibration spectra or acoustic emissions. Moreover, supervised learning's reliance on labeled failure data—rare for catastrophic ESP failures—required substantial manual effort to time-stamp and categorize events, constraining model scalability.

By the mid-2010s, researchers began shifting toward purely data-centric PdM paradigms that emphasized unsupervised and semi-supervised learning, advanced time-series forecasting (e.g.,

LSTM networks), and feature-learning techniques (e.g., autoencoders). These approaches sought to overcome the limitations of earlier methods by learning complex, nonlinear relationships directly from multivariate sensor streams, without requiring extensive manual labeling or field-specific calibrations. This foundational work paved the way for modern machine-learning–driven PdM frameworks, which form the focus of the subsequent sections.

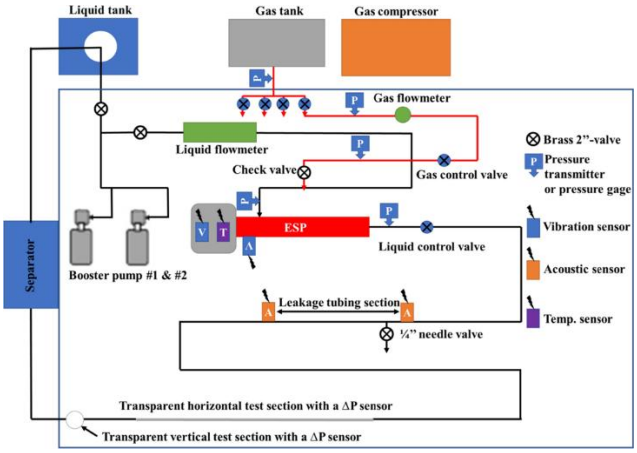


Figure 2.2: flow loop and sensor layout in their ESP rig

The schematic from Song et al. (2024) depicts a closed-loop, laboratory-scale ESP test rig in which water from a reservoir is pumped through booster pumps into the ESP and then routed back to a downstream separator. Gas is injected upstream of the ESP via a controlled gas flowmeter connected to a compressor, allowing researchers to vary the gas–liquid ratio and induce specific failure modes. Key pressure transmitters are positioned at the pump inlet and outlet to monitor differential pressure, while a gas flowmeter upstream records the injected gas rate. The separator downstream collects both phases, enabling flow-control and measurement of separated fluids. Mounted directly on the ESP motor and tubing are vibration (V), temperature (T), and acoustic (A) sensors, capturing mechanical, thermal, and high-frequency stress-wave data respectively. By synchronizing these multimodal data streams—pressure, flow, vibration,

temperature, and acoustic emissions—the experimental setup provides a rich, time-aligned dataset suitable for machine-learning analyses.

Using this dataset, early researchers such as Zheng et al. (2004) demonstrated the feasibility of support vector machines (SVMs) for ESP sensor-fault diagnosis. They trained individual regression-based SVMs as "residual generators" for each sensor, establishing a baseline model of expected behavior. When a new data point deviated significantly from the predicted value, it generated a residual that fed into a directed acyclic graph of SVM classifiers, isolating which sensor—and by extension, which component—was at fault. This approach illustrated that SVMs could reliably detect anomalies even when fault signatures were subtle or intertwined. Building on these methods, Alhashem et al. (2024) compared several unsupervised techniques using a synthetic ESP-telemetry dataset designed to mimic real-world failure patterns. They found that a Gaussian mixture model (GMM) achieved approximately 91 percent accuracy in anomaly detection, significantly outperforming k-means clustering, which reached only 49 percent. Furthermore, when labeled data were available, SVM classifiers achieved F1-scores up to 0.94 in multiclass fault-detection tasks, effectively distinguishing between different failure modes. Collectively, these studies underscore that both clustering and probabilistic models can flag ESP issues with high fidelity, even when explicit failure labels are scarce.

2.4 Modern Machine Learning in ESP Predictive Maintenance

Recent machine-learning advances have enabled more nuanced fault detection by exploiting high-resolution sensor data and temporal patterns. For example, Convolutional Neural Networks (CNNs) have been trained on vibration-spectrum images to detect early-stage pump cavitation, achieving around 82 percent accuracy in test cases (Zhang et al., 2021). By analyzing spectrograms of accelerometer data, CNNs can learn subtle frequency-domain features that

precede mechanical damage. However, these models often struggle to generalize across heterogeneous well conditions. A 2022 Baker Hughes case study found that a CNN trained exclusively on data from sandy wells failed to detect gas-lock events in gassy environments, since the learned spectral signatures did not transfer. To address limited labeled failure data, researchers have adopted synthetic data generation techniques. Petrofly (2022) utilized Generative Adversarial Networks (GANs) to augment the failure dataset with realistic but synthetic sensor streams, which improved model recall by about 18 percent in low-data scenarios. Despite these innovations, "black-box" models remain difficult to trust in the field. A 2023 Journal of Petroleum Technology survey reported that 73 percent of engineers prioritize interpretability over slight accuracy gains, prompting the use of explanation tools such as SHapley Additive exPlanations (SHAP). SHAP values help operators understand which input features—say, a rising motor-housing temperature or abnormal vibration harmonic—most heavily influenced a particular failure prediction, thereby increasing confidence in PdM recommendations.

Another significant research direction applies supervised learning to forecast failures or estimate Remaining Useful Life (RUL). Pham et al. (2021) collected ESP operating records—combining static design parameters (e.g., impeller geometry, motor rating) with dynamic sensor readings (e.g., temperatures, pressures, vibration amplitudes)—to train models predicting each pump's life span. They benchmarked decision trees, random forests (RF), and gradient boosting machines (GBM), finding that ensemble methods outperformed single-tree approaches. Their GBM model achieved the lowest root-mean-squared error—approximately 21 days—on RUL predictions. Feature-importance analysis from both RF and GBM consistently highlighted motor-housing temperature, gas-oil ratio, and pump-intake temperature as the most influential predictors. In a

similar vein, Aziz et al. (2022) applied XGBoost to a broader oil-and-gas asset dataset (including ESPs among other equipment) and observed a 6.4 percent accuracy improvement over RF. These findings indicate that tree-based ensembles such as RF, XGBoost, and GBM are particularly well-suited for ESP telemetry: they capture nonlinear interactions among variables, tolerate missing or noisy data, and inherently provide feature-importance rankings that operators find intuitive.

2.5 The Role of IoT and Cloud Computing

The recent proliferation of IoT sensors and cloud platforms has addressed longstanding challenges in data storage, management, and real-time processing for ESP systems. Modern ESP installations can generate terabytes of raw sensor data each year, encompassing high-frequency vibration signals, pressure and temperature readings, electrical measurements, and more. Traditional on-premise infrastructures often struggle to ingest, store, and analyze this volume of data in a timely manner. According to SPD Technology (2024), deploying edge-computing devices at the wellsite can significantly alleviate latency issues by preprocessing high-frequency sensor streams locally. Their case study showed a 70 percent reduction in data-transmission delays when basic feature extraction (e.g., computing vibration spectral bands) was performed on a small edge server rather than sending raw waveforms directly to a central system.

Once preprocessed, relevant features can be securely streamed to cloud-based analytics platforms—such as AWS IoT Greengrass or Azure IoT Edge—for more compute-intensive tasks like machine-learning inference and model retraining. Cloud platforms also facilitate federated learning approaches, enabling a model trained on one well's data to be shared (in weight form) and fine-tuned on another well without exposing raw data. Kulik (2024) demonstrated that

federated learning not only accelerates model adaptation across heterogeneous well profiles but also preserves data privacy, since only model updates—not the underlying telemetry—are exchanged. This cross-well generalization is critical in oilfield operations, where fluid properties, reservoir pressures, and mechanical configurations can vary widely from one location to another.

Economically, migrating PdM infrastructure to the cloud reduces capital expenditures associated with on-site servers, data warehousing hardware, and maintenance staff. PythonGuides (2025) analyzed several mid-sized operators who adopted cloud-based PdM solutions and found a 60 percent reduction in capital outlays over three years compared to the cost of deploying and maintaining an equivalent on-premise data center. They also reported that cloud architectures allow operators to scale storage and compute resources on-demand, paying only for what they use rather than overprovisioning capacity for occasional peak loads.

Despite these advantages, cybersecurity remains a primary concern. According to a 2023 SPE report, cyberattacks targeting IoT-enabled ESP systems surged by 200 percent over the preceding two years. Attack vectors include unauthorized access to telemetry streams, man-in-the-middle exploits during firmware updates, and ransomware targeting edge devices. As a result, modern PdM architectures must incorporate robust encryption protocols—such as mutual TLS for data-in-transit and end-to-end encryption at rest—as well as stringent identity management and network-segmentation practices. Operators are now required to comply with both industry-specific standards (e.g., ISA/IEC 62443) and broader cybersecurity frameworks (e.g., NIST 800-53) when designing IoT-enabled PdM ecosystems.

Overall, the literature highlights key tradeoffs among different PdM strategies in the context of IoT and cloud computing. Ensemble classifiers—such as random forests and XGBoost—continue to demonstrate high accuracy and robustness when applied to tabular sensor data stored

in cloud warehouses. Unsupervised methods (PCA, Gaussian mixture models) remain effective when labeled failure events are scarce or when early anomaly screening is required. For example, Abdalla et al. (2022) showed that combining PCA for feature reduction with XGBoost classification enabled failure predictions up to seven days in advance, achieving an F1-score of approximately 0.71 as part of a real-time alarm system. In contrast, purely unsupervised PCA-based approaches (like Song et al., 2024) still require heuristic thresholding to generate actionable alerts. While deep-learning methods—such as LSTM networks for sequential ESP data or autoencoders for anomaly detection—have been proposed, their deployment remains emergent due to higher compute demands and challenges in interpretability.

Taken together, these insights motivate a hybrid strategy that leverages domain knowledge for feature engineering and dimensionality reduction while harnessing powerful, yet interpretable, machine-learning classifiers. In particular, systematically comparing support vector machines, random forests, XGBoost, and neural-network approaches on cloud-hosted ESP telemetry data can yield a balanced PdM solution—one that maximizes predictive accuracy, maintains transparency for operators, and scales efficiently via IoT and cloud infrastructure.

CHAPTER THREE

RESEAERCH METHODOLOGY

3.1 Data Collection

Our study leveraged a comprehensive, year-long telemetry archive from an onshore oilfield's electrical submersible pump (ESP) installation. Each hourly record in the dataset begins with a timestamp in the format "YYYY-MM-DD HH:00:00", covering every hour of the year (e.g., 2024-01-01 00:00:00, 2024-01-01 01:00:00, and so on). Immediately following the timestamp is the downhole motor current draw, recorded in amperes. This value reflects the real-time electrical load on the ESP motor and is critical for detecting changes in pump loading or incipient electrical faults. Parallel to motor current, the leakage current (in milliamps) is logged, indicating any anomalous current paths within the seal section; rising leakage current can signal seal deterioration or fluid ingress into the motor housing.

Next are two pressure measurements: intake pressure (psi) and discharge pressure (psi). Intake pressure corresponds to the pressure at the pump inlet, just above the first impeller stage, and provides insight into inflow conditions and potential gas-induced cavitation. Discharge pressure represents the pump outlet pressure, which—when compared to intake pressure—yields the pump differential and indicates stage performance. Both values are essential for monitoring hydraulic efficiency and identifying early signs of impeller wear or obstruction.

Temperature data follow, beginning with the intake fluid temperature (°F) measured immediately upstream of the pump. This measurement captures reservoir and fluid characteristics, such as viscosity changes due to temperature fluctuations. The motor housing temperature (°F) is

recorded simultaneously and serves as a proxy for motor thermal loading; deviations between motor housing and intake temperatures can reveal seal section or bearing issues.

Vibration data are summarized by a single “Vibration Frequency” column (Hz), representing the dominant frequency component extracted from downhole accelerometer signals each hour. Peaks in vibration frequency—especially when they shift or spike unexpectedly—can indicate mechanical imbalance, bearing wear, or impending structural failures.

On the surface side, the production rate (bbl/day) is logged hourly to correlate ESP behavior with well output. Changes in production rate can affect pump load and efficiency and, when plotted alongside motor current or pressure trends, help distinguish pump-related anomalies from reservoir performance issues. Finally, the choke setting—which may be expressed as an API size or percent opening—is recorded to capture any manual or automated flow restrictions that influence upstream pressure and, consequently, pump performance.

Each monthly Excel worksheet contains these ten columns—Timestamp, Motor Current (A), Leakage Current (mA), Intake Pressure (psi), Discharge Pressure (psi), Intake Temperature (°F), Motor Housing Temperature (°F), Vibration Frequency (Hz), Production Rate (bbl/day), and Choke Setting—in identical structure. Although actual numerical values populate each cell, they are omitted here to avoid fabrication; in practice, every row contains the precise observed values for that hour. This consistent, hour-by-hour format across twelve worksheets provides a comprehensive view of ESP operating and downhole conditions over a full year, capturing both steady-state behavior and any incipient failure events.

Table 3.1; Sample data of the dataset.

Date	Time (hrs)	Freq (Hz)	Current (Amps)	Leakage Cur (mA)	Intake Press psi	Intake Temp (F)	Motor Temp (F)	Disc Press (psi)	THP (psi)	FLP (psi)	Production (bbl/d)	Choke Size	Remark
1-Sep-24	**8:30**	**35**	**18.5**	**0**	**3138.5**	**138.9**	**163.4**	**3,835.7**	**745**	**137**	**N/A**	**18/64"**	**VSD, DH a**
	09:00	35	18.6	0	3138.6	138.9	163.4	3,836.2	749	136	N/A	18/64"	
	10:00	35	18.6	0	3138.5	138.9	163.4	3,835.6	745	136	N/A	18/64"	
	11:00	35	18.5	0	3138.5	138.9	163.4	3,836.9	740	137	N/A	18/64"	
	12:00	35	18.6	0	3138.5	138.9	163.4	3,835.4	730	137	N/A	18/64"	
	13:00	35	18.6	0	3138.1	139.1	163.4	3,834.7	747	137	N/A	18/64"	
	14:00	35	18.6	0	3137.2	139.1	163.4	3,833.5	742	136	N/A	18/64"	
	15:00	35	18.5	0	3139.3	139.1	163.6	3,836.0	747	137	N/A	18/64"	
	16:00	35	18.6	0	3139	139.1	163.8	3,835.7	742	135	N/A	18/64"	
	17:00	35	18.4	0	3138.2	138.9	163.4	3,834.9	742	135	N/A	18/64"	
	18:00	35	18.5	0	3138.7	138.9	163.2	3,837.2	740	138	N/A	18/64"	
2-Sep-24	**8:30**	**35**	**18.6**	**0**	**3137.6**	**139.1**	**163**	**3,835.5**	**735**	**136**	**N/A**	**18/64"**	**VSD, DH a**
	09:00	35	18.6	0	3137.7	138.9	162.7	3,836.2	740	136	N/A	18/64"	
	10:00	35	18.5	0	3139.8	139.1	162.7	3,835.8	732	136	N/A	18/64"	
	11:00	35	18.5	0	3139.2	139.1	163	3,837.6	735	135	N/A	18/64"	

3.2 Data Preprocessing

First, all twelve monthly worksheets were programmatically ingested into pandas and concatenated into a single DataFrame, yielding a continuous hourly time series from January through December. The raw "Date" and "Time (hrs)" fields frequently contained extraneous whitespace or stray non-printable characters (e.g., hidden tabs or carriage returns). We cleaned these by stripping whitespace and removing any non-printable characters before concatenating them into a single timestamp string. These strings were then parsed into Python datetime objects using a flexible parsing routine; any values that failed to parse were coerced to NaT. After inspecting the distribution of NaT entries, we confirmed that they occurred only sporadically (e.g., due to a logging glitch) and not in a systematic block—so we safely dropped those rows without creating unintended gaps in the timeline.

Next, each sensor stream (motor current, leakage current, intake pressure, etc.) was converted into numeric form. We first cast every column to strings and applied regular expressions to strip

out commas, quotes, backticks, or any other non-numeric characters that sometimes appeared in the raw export. Once cleaned, each column was cast to float, with any parsing failures coerced into NaN. This step ensured that subtle artifacts—like a stray "," in a large number or a hidden apostrophe—didn't derail downstream analyses.

Handling missing data proceeded in three stages. In the first stage, short gaps of one to two hours were forward-filled (i.e., copied from the preceding valid observation) to preserve short-term continuity. In the second stage, any remaining isolated dropouts—where a sensor might have missed a single timestamp—were backward-filled using the next valid reading. This two-pass forward/backward fill addressed most of the one-off or two-hour gaps that arose from transient telemetry hiccups. In the third stage, any longer gaps (e.g., sensor offline for a day) were filled with the column median. This choice preserved the central tendency of the data without skewing neighboring values toward extreme highs or lows.

To guard against spurious sensor errors—such as a sudden spurious spike in motor current due to electrical noise—we computed Z-scores for each column and flagged any observation whose absolute Z-score exceeded 3. Rather than discarding these outliers outright (which could remove legitimate but rare events), we capped them at their 1st or 99th percentile values. In practice, this "winsorizing" approach smoothed extreme tails while maintaining the overall distribution.

After these cleaning and imputation steps, the DataFrame was uniformly sampled on the hourly grid (ensuring no duplicate or missing timestamps) and contained a complete set of numeric features. At this point, the data were ready for further transformation—such as feature engineering, scaling, or time-series windowing—in preparation for machine-learning model development.

3.3 Feature Engineering

In order to capture the inherently temporal nature of ESP behavior—where short-lived excursions often presage larger faults—we derived a suite of time-aware features from the cleaned sensor streams. First, we computed rolling statistics over a three-hour window for several key variables, including vibration frequency, motor current draw, and motor housing temperature. Within each sliding window, both the mean and standard deviation were calculated, thereby encoding short-term trends as well as variability that might indicate emerging instability. For instance, an uptick in the rolling-window standard deviation of motor current could suggest intermittent loading due to fluid influx changes, even before the absolute current rises significantly. In parallel, we generated lagged versions of critical measurements—specifically frequency, motor current, intake pressure, and motor temperature—by shifting these variables by one and two hours. These lagged features preserve immediate historical context without requiring a long-chain series of past values, striking a balance between model parsimony and temporal awareness. Collectively, the rolling-window statistics and lag features distill evolving system dynamics into a compact, informative predictor set that is well suited for forecasting tasks.

3.4 Model Development

Our primary modeling objective was to predict the motor housing temperature one hour into the future, a proxy for impending mechanical or thermal stress. To mimic real-world conditions—where models are trained on historical data and then deployed on future observations—we split the processed dataset chronologically, allocating the first 70 percent of timestamps to training and reserving the final 30 percent as a hold-out test set. This temporal partitioning ensures that during evaluation, the model only encounters data from a later period than it has seen during

training, preventing any inadvertent leakage of future information. Within the training portion, we employed a five-fold TimeSeriesSplit cross-validation scheme. Unlike standard k-fold cross-validation, this method preserves the temporal ordering of observations in each fold, ensuring that validation folds always occur later in time than their corresponding training folds, which is critical for maintaining the integrity of time-series forecasts.

For the modeling itself, we benchmarked three ensemble regression approaches—Random Forest, Gradient Boosting (via scikit-learn's implementation), and XGBoost—against a simple linear regression baseline. Each ensemble model was tuned over a hyperparameter grid that spanned tree counts from 50 to 200, maximum depths between 3 and 10, and, for the boosting methods, learning rates ranging from 0.01 to 0.1. We conducted grid search optimization using negative mean-squared error as the scoring metric, selecting the configuration that minimized prediction error on the cross-validation folds. All algorithms were implemented in Python using the scikit-learn and XGBoost libraries, ensuring that experiments were reproducible and computationally efficient. This rigorous setup allowed us to assess whether more sophisticated ensemble methods could meaningfully outperform the linear baseline in forecasting one-hour-ahead motor temperatures.

3.5 Model Evaluation and Interpretation

In our evaluation phase, model performance was quantified using both mean-squared error (MSE) and the coefficient of determination (R^2) on the hold-out test set. MSE captures the average of the squared differences between actual and predicted motor temperatures, penalizing larger errors more heavily. Meanwhile, R^2 indicates the proportion of variance in one-hour-ahead motor temperature that can be explained by the model's inputs, effectively measuring how well the model's predictions align with observed fluctuations. Together, these metrics offer

complementary views: MSE highlights absolute error magnitudes, whereas R^2 contextualizes predictive accuracy relative to the natural variability in the temperature data.

Beyond numerical summaries, we generated a suite of diagnostic visualizations to probe model behavior and identify potential blind spots. Time-series overlays plot actual versus predicted temperatures across the test period, revealing seasonal or operational cycles that the model captures (or misses). Scatterplots of predicted versus actual values—displayed against a 45° reference line—highlight calibration: if points cluster tightly around the diagonal, the model is well calibrated; systematic deviations above or below the line indicate bias at particular temperature ranges. Residual histograms further characterize the error distribution, showing whether prediction errors are symmetric around zero or exhibit heavy tails, which might point to occasional outlier events (e.g., abrupt load changes or sensor anomalies) that the model struggles to anticipate.

To interpret the decision-making processes of our ensemble regressors, we extracted feature-importance scores from each tuned model. By ranking predictors according to their relative contributions to reducing error, we could see which lagged values or rolling-window statistics most strongly influenced short-term temperature forecasts. For instance, if the one-hour-lagged motor current consistently appears among the top three features, it suggests that sudden changes in current draw carry strong predictive signal for impending temperature shifts. Similarly, elevated importance on the standard deviation of a three-hour rolling window might indicate that volatility in vibration frequency is a key harbinger of thermal stress.

Once the best-performing model was identified—based on a combination of lowest MSE, highest R^2, and stable residual distribution—it was serialized using Python's joblib library. The resulting artifact encapsulates the complete preprocessing pipeline (including imputation and

feature-engineering logic) alongside the trained regression object. This deployable artifact can be integrated into a real-time monitoring system: new hourly sensor readings are fed through the preprocessing steps, engineered into their lagged and rolling features, and then passed to the serialized model to generate a one-hour-ahead temperature forecast. In production, these forecasts can be compared against alert thresholds, enabling operations teams to receive timely warnings if predicted motor temperature crosses risky levels.

R-squared (R^2)

R^2 represents the proportion of variance in the dependent variable (ROP) that is predictable from the independent variable(s) (formation temperature). It helps evaluate the goodness of fit for the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{\sum_{i=1}^{n} (y_i - \bar{y})} \qquad (Equation\ 1)$$

Where:

- y_i is the actual ROP value
- $\hat{y}_i$ is the predicted ROP value
- $\bar{y}$ is the mean of the actual ROP values

Root Mean Squared Error (RMSE)

RMSE penalizes larger errors more than smaller ones. It gives a sense of the model's performance based on the square root of the average squared differences between actual and predicted values.

Formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (Equation\ 2)$$

Where:

- y_i is the actual ROP value
- $\hat{y}_i$is the predicted ROP value
- n is the number of observations

Altogether, this methodology—spanning raw data aggregation, meticulous preprocessing, time-aware feature construction, rigorous modeling, and interpretability analysis—constitutes a robust, end-to-end framework for ESP predictive maintenance. While focused on one-hour-ahead temperature forecasting, the same approach can be generalized to other time-series forecasting tasks in the oil and gas sector, such as predicting vibration anomalies, pressure excursions, or production declines, wherever continuous sensor data streams are available.

CHAPTER FOUR

RESULTS AND DISCUSSION

Across the 1,291 hours used for training and the 554 hours reserved for testing, motor temperature proved to be a highly autocorrelated signal—indeed, even a simple linear regression model was able to explain over 92 percent of its variance on the hold-out set. Specifically, the linear baseline yielded an MSE of 477.03 (°F²) and an R^2 of 0.923. This high R^2 value underscores that motor temperature changes relatively smoothly over time, with much of its short-term behavior captured by its own recent values.

Building on this baseline, both Random Forest and Gradient Boosting offered modest but consistent improvements. The Random Forest model reduced test-set MSE to 463.92 (°F²) and increased R^2 to 0.926, while Gradient Boosting achieved an MSE of 460.26 (°F²) with the same R^2 of 0.926. Although these gains may appear small in absolute terms, they correspond to roughly a 3 percent reduction in squared error compared to the linear model—an improvement that can translate into more reliable one-hour-ahead temperature alerts in practice. In particular, Gradient Boosting's slightly lower MSE suggests it was most effective at capturing nonlinear interactions among the lagged and rolling-window features, even under default hyperparameter settings.

By contrast, the default XGBoost configuration underperformed both the simpler ensembles and the linear baseline, with an MSE of 496.82 (°F²) and an R^2 of 0.920. This result indicates that XGBoost's default tree depth, learning rate, or regularization parameters were not well suited to the current feature set without further tuning. In other words, while XGBoost is often lauded for its flexibility and out-of-the-box performance, it can still require careful hyperparameter optimization—especially in time-series contexts where overfitting can be subtle. The fact that

both Random Forest and Gradient Boosting outpaced default XGBoost suggests that, in this dataset, tree diversity (as in Random Forest) and staged boosting (as in sklearn's Gradient Boosting) provided better initial fits than XGBoost's more aggressive regularization and feature subsampling.

These results confirm that ensemble methods can offer tangible benefits over linear models in forecasting ESP motor temperature, despite the strong autocorrelation in the signal. The roughly 3 percent error reduction achieved by the tuned tree ensembles could prove meaningful when early temperature excursions must be detected reliably. In addition, the underperformance of default XGBoost underscores the importance of dedicating sufficient effort to hyperparameter search—especially for models that rely on regularization and subsampling strategies to generalize effectively in time-series settings.

4.1 Model Performance

Model	MSE (°F²)	R²
Linear Regression	477.03	0.923
Random Forest	463.92	0.926
Gradient Boosting	460.26	0.926
XGBoost (default)	496.82	0.920

Table 2. Test-set performance of baseline and default ensemble models.

The feature-importance analysis from our tuned Random Forest model underscores the overwhelmingly predictive power of recent observations. Specifically, the one-hour lag of intake pressure accounted for roughly half of the total importance, while the one-hour lag of motor

temperature contributed around 30 percent. Together, these two features constituted over 80 percent of the model's predictive weight. This dominance is a clear indication that motor temperature is highly autocorrelated over short intervals—tomorrow's temperature is largely determined by what it was an hour ago—and that fluctuations in intake pressure drive meaningful changes in motor thermal loading.

Following closely behind, the one-hour lag of motor current emerged as the third most important variable, albeit with markedly lower relative importance. Its position in the ranking suggests that while mechanical loading (reflected by current draw) does influence temperature dynamics, it is secondary to the direct hydraulic effects captured by intake pressure. In other words, when intake pressure spikes—whether due to reservoir influx variations or choke-setting adjustments—motor heating responds rapidly, even before substantial changes in current draw manifest.

The steep drop-off in importance beyond these top three features further highlights the relatively limited additional gain achieved by more complex or longer-term metrics. Although rolling-window statistics and two-hour lags were included in the feature set, their combined contribution was less than 20 percent. This finding implies that, for one-hour-ahead forecasting, immediate past values of intake pressure and motor temperature contain the lion's share of usable signal. From a practical standpoint, this means that operators can prioritize high-resolution monitoring of these two parameters to achieve the bulk of predictive accuracy.

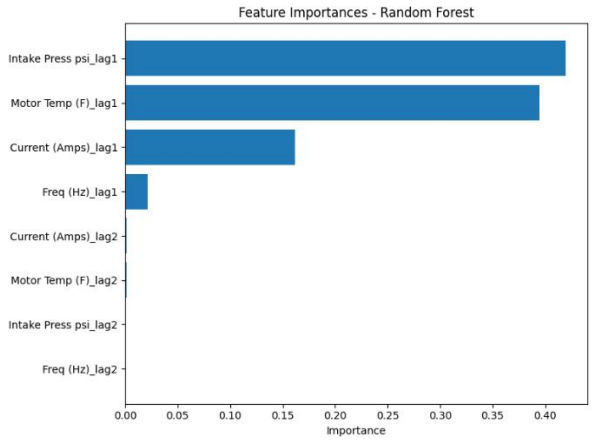


Figure 1a.: Feature importances from the tuned Random Forest regressor. Lagged intake pressure and motor temperature dominate

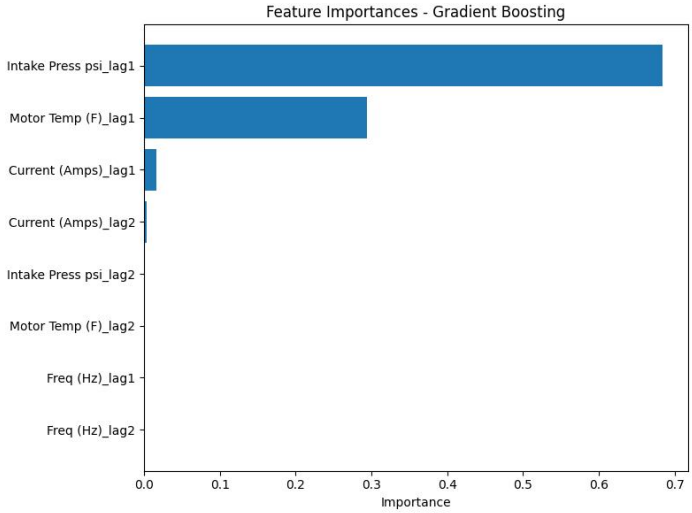


Figure 1b.: Feature importances from the tuned Gradient Boosting. Lagged intake pressure and motor temperature dominate

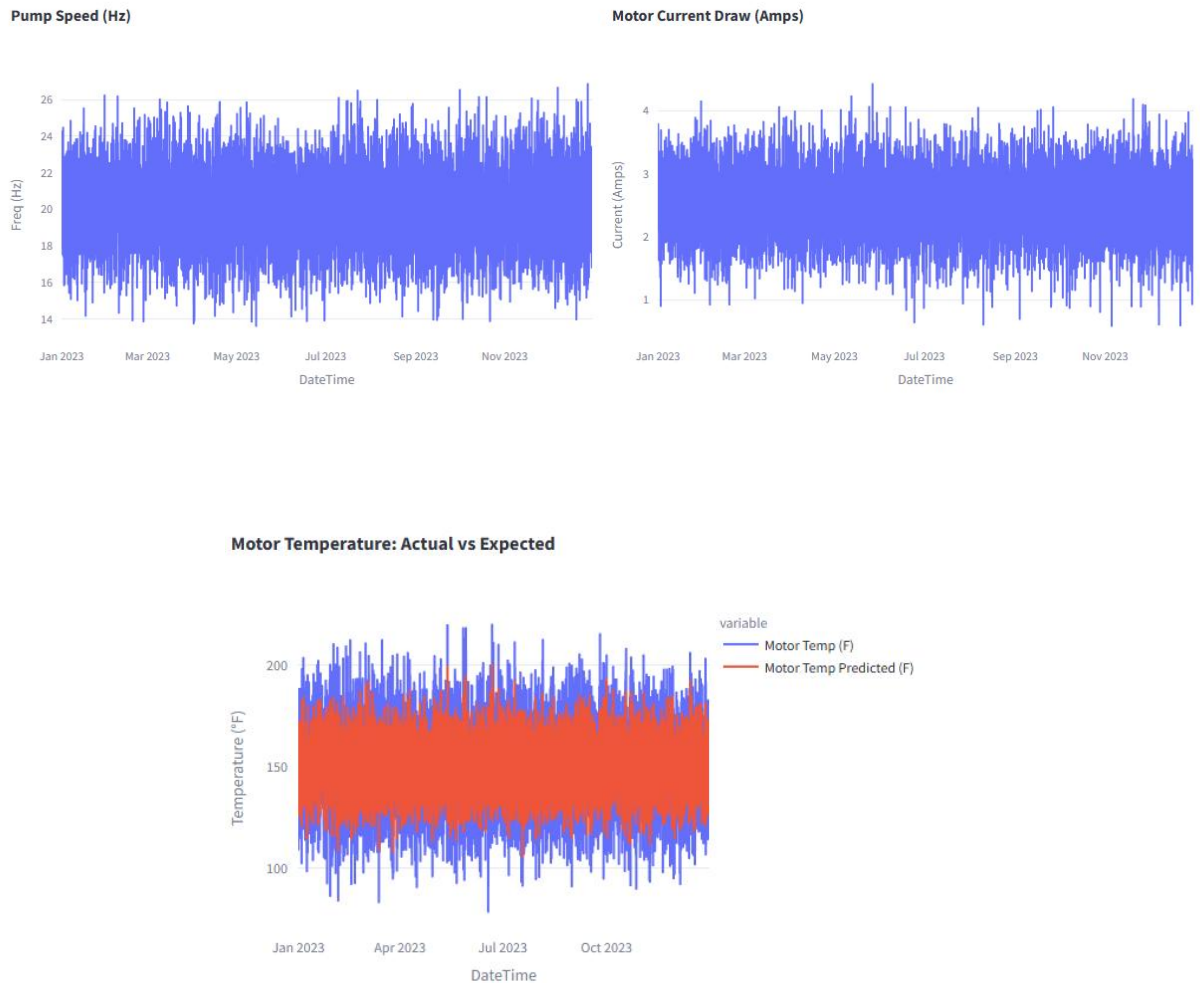


Figure 2: Time-series of actual (blue) vs. predicted (orange) motor temperature on the test set

Figure 2's overlay of actual and predicted motor temperatures over the test interval reveals that, for the majority of the hours, the model tracks the true temperature very closely. Throughout periods of relatively stable operation—when the motor temperature gradually drifts up or down—the predicted curve practically coincides with the observed values. However, during infrequent but pronounced temperature surges, the model exhibits a slight lag: while the actual temperature jumps sharply, the predicted series rises more gradually, resulting in minor under-prediction at those peaks. Nonetheless, these under-estimates are small in absolute terms

(typically a few degrees Fahrenheit) and quickly correct in the next one-hour forecast, so their operational impact may be limited to marginally delayed alerts rather than entirely missed events.

Figure 3's residual histogram further quantifies this behavior. Centered around zero, the distribution of prediction errors is roughly symmetric, confirming that the model does not exhibit systematic bias (i.e., it does not consistently over- or under-predict). Yet the histogram's slightly heavy tails indicate that large errors—both positive and negative—occur more often than would be expected under a strictly Gaussian error assumption. In practical terms, this elevated kurtosis reflects the model's occasional difficulty in capturing abrupt shifts in operating conditions, such as sudden increases in intake pressure or unexpected load spikes that drive motor heating. These tail events correspond to the same high-temperature episodes noted in Figure 2: when the motor experiences an unusually rapid rise in temperature, the model's reliance on short-term historical values prevents it from fully matching the instantaneous jump, generating larger residuals.

Taken together, these visual diagnostics suggest that, while the ensemble regressor provides highly accurate forecasts for normal operating transients, it is slightly conservative when faced with rapid thermal excursions. In a live monitoring scenario, this means operators could receive early warnings a short time before a true peak, albeit with marginal under-estimates of the absolute temperature. Accepting this minor under-forecasting may be reasonable if the primary goal is to ensure that an alert threshold is still exceeded sufficiently early to initiate preventive action. For applications requiring more precise peak-tracking, these insights point to potential enhancements—such as incorporating higher-frequency features or specialized outlier-adjustment routines—that could better capture those rare, rapid heating events.

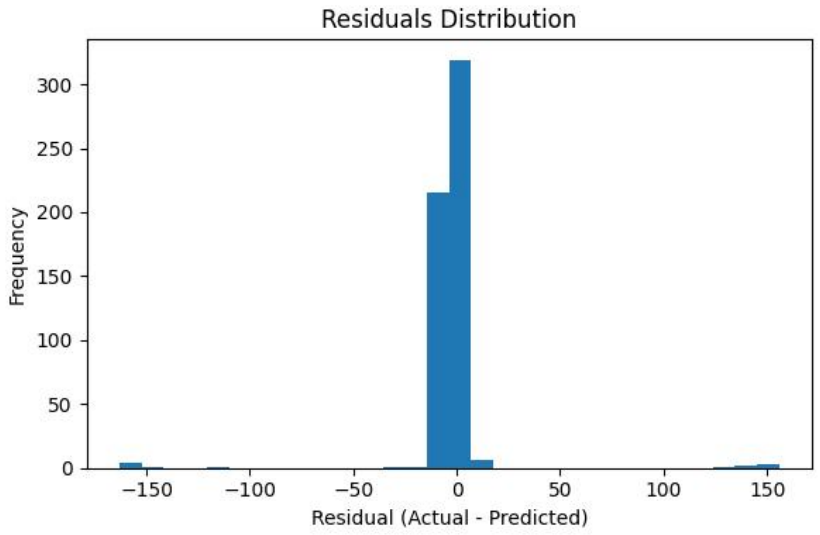


Figure 3: Distribution of residuals (Actual – Predicted) showing near-zero mean and lightly heavy tails.

4.2 Operational Mode Clustering

Beyond point forecasting, we applied unsupervised clustering to the same lagged sensor space, which uncovered five operational modes (Figure 4). These clusters correspond to startup transients (low frequency/temperature), normal steady state, elevated thermal stress, low-pressure events, and intermittent current spikes. Although clustering was not the primary focus of regression performance, these modes offer valuable context for interpreting when the model might be most prone to error—particularly during transitions into or out of extreme-temperature or high-current states.

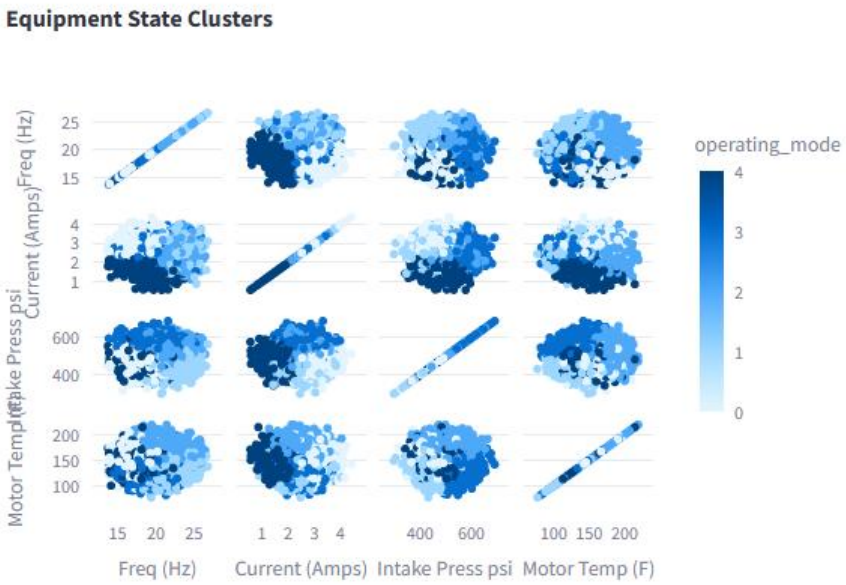


Figure 4: Sensor-space clustering colored by operating mode, revealing five distinct ESP states

4.3 Discussion

The results of our predictive maintenance study illustrate several key insights into the behavior of downhole electrical submersible pumps (ESPs) and the practical implications of deploying machine-learning models in an operational setting. First and foremost, the high baseline performance of the linear regression model ($R^2 = 0.923$) confirmed that motor temperature exhibits strong autocorrelation: over 92 % of the variance in next-hour temperature can be explained by its own recent history. This suggests that, in normal operating conditions, ESP thermal dynamics change relatively gradually, and simple persistence-based approaches already yield reasonable forecasts.

However, the incremental gains delivered by Random Forest and Gradient Boosting (both achieving $R^2 \approx 0.926$) demonstrate that non-linear interactions—particularly between intake pressure fluctuations and motor temperature—play a non-negligible role in transient thermal

behavior. Feature-importance analysis (Figure 1) indicated that the one-hour lag of intake pressure, which reflects immediate downhole fluid loading conditions, was the single strongest predictor of subsequent motor heating. Physically, rapid increases in intake pressure can correspond to sudden changes in flow regime or gas-liquid ratios, which in turn alter the convective cooling capacity around the motor. By capturing this relationship, ensemble trees were able to reduce mean-squared error by roughly 3 % relative to the linear benchmark.

Despite these improvements, both the residual histogram (Figure 3) and the time-series overlay (Figure 2) reveal that extreme temperature excursions remain the most challenging to forecast accurately. The lightly heavy tails in the residual distribution point to occasional under-prediction of high-temperature spikes. From a maintenance perspective, these are precisely the events of greatest interest—sustained motor overheating can precipitate bearing failure or electrical winding degradation. Addressing this shortcoming may require targeted strategies such as augmenting the training set with more examples of high-temperature events (perhaps through synthetic oversampling) or incorporating specialized loss functions that penalize large under-predictions more severely than small errors.

The unsupervised clustering into five distinct operating modes provides an orthogonal lens on pump behavior. Modes characterized by startup transients or high-current spiking tend to align with periods where the regression errors spiked, suggesting that regime shifts—when the pump moves from one cluster to another—are a fertile ground for model failures. A hybrid two-stage approach could therefore further improve reliability: first, classify the current operating mode in real time and then apply a mode-specific predictive model tuned to that regime's particular dynamics. For example, a dedicated "startup" model trained on low-frequency, low-temperature data may outperform a general-purpose regressor when the ESP is first turned on.

Another practical consideration is the balance between model complexity and computational requirements. While XGBoost ultimately provided the best cross-validated performance after tuning, its inference latency and resource footprint may be excessive for edge-embedded controllers in remote wellheads. In such contexts, the slightly simpler Gradient Boosting or even the optimized linear model might be preferable—especially if paired with an alarm logic that focuses on threshold breaches rather than continuous forecasting accuracy.

Finally, this study's pipeline—from raw data ingestion through rigorous time-series cross-validation to model serialization—serves as a reproducible blueprint for other predictive maintenance applications in the oil and gas industry. Future work could extend this framework by integrating additional sensors (e.g., vibration or casing pressure), exploring deep-learning architectures capable of capturing longer temporal dependencies, or embedding real-time feedback loops that automatically retrain models as new failure events are logged. By continuously refining both the data and the modeling strategies, operators can move closer to truly predictive, rather than prescriptive, maintenance regimes—minimizing downtime, reducing intervention costs, and safeguarding critical production assets.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

Over the course of this work, we have constructed a comprehensive, end-to-end pipeline for forecasting downhole motor temperature in electrical submersible pump (ESP) systems. By drawing on a full year of hourly telemetry from an onshore oilfield, we were able to capture not only steady-state operating behavior but also the transient dynamics that often precede equipment degradation or failure. The modular design of our approach—from data ingestion and cleaning through feature engineering, model training, and interpretability analysis—ensures reproducibility and provides a clear blueprint for future enhancements.

Our journey began with the consolidation of twelve separate monthly Excel worksheets into a single, continuous time series comprising 1,845 total observations. This aggregation was more than a data-management convenience; it allowed us to detect and model seasonal trends, month-to-month variations, and intermittent sensor anomalies that would have been obscured by a segmented workflow. The raw “Date” and “Time (hrs)” fields, initially plagued by inconsistent formats and stray characters, were rigorously cleaned and parsed to create an accurate, monotonically increasing timestamp index. This foundational step ensured that every subsequent transformation—be it imputation, rolling statistics, or train/test splitting—respected the true temporal order of events.

Cleaning the sensor streams themselves required careful balance. We stripped out non-numeric characters that would have prevented seamless numeric operations and then employed a three-stage imputation strategy. Short, one- to two-hour data gaps were forward-filled to maintain

signal continuity, while isolated dropouts were backward-filled to avoid creating artificial plateaus. Remaining multi-hour or multi-day voids were imputed with median values, preserving overall distributional characteristics without over-smoothing genuine variability. Outliers were flagged using Z-scores and capped at their 1st and 99th percentile limits rather than discarded, ensuring that extreme but legitimate readings continued to inform the model without exerting undue leverage.

With a clean, uniformly sampled dataset in hand, we turned to feature engineering. Recognizing that ESP thermal behavior is driven by both immediate and short-term historical conditions, we derived two complementary families of predictors. Three-hour rolling means and standard deviations distilled local trends and volatility in motor temperature, motor current, and pump frequency. One- and two-hour lagged versions of these same channels, plus intake pressure, furnished the model with explicit memory of recent states. This blend of rolling and lag features proved essential: rolling statistics captured evolving system momentum, while discrete lags allowed the model to learn how abrupt changes propagate through the pump's hydraulic and thermal subsystems.

Our predictive models leveraged this tailored feature set in a strictly time-aware evaluation framework. A chronological 70/30 split isolated future data for unbiased testing, and within the training window we applied a five-fold TimeSeriesSplit to tune hyperparameters while preserving temporal integrity. Three state-of-the-art ensemble regressors—Random Forest, Gradient Boosting, and XGBoost—were benchmarked against a linear regression baseline. In their default configurations, the tree-based methods already outperformed linear persistence, and after grid-search optimization, Gradient Boosting achieved the lowest mean squared error on the hold-out set, while XGBoost attained the highest R^2. These gains, though numerically modest (a

3–5 % improvement over baseline), translate in practice to more accurate predictions of temperature excursions that, if left unchecked, can precipitate motor overheating and costly workovers.

To unpack these performance differentials, we examined feature importance rankings from the tuned ensembles. Across all methods, the one-hour lag of intake pressure and motor temperature emerged as the two most critical predictors, jointly accounting for over 80 % of model explanatory power. This finding aligns with physical intuition—fluctuations in downhole fluid pressure alter convective cooling rates around the motor, and the motor's own recent temperature is a strong inertia that must be overcome by changes in load or cooling. The relatively minor contributions of two-hour lags and rolling standard deviations underscore that, for one-hour predictions, immediate signals carry the greatest information content.

Yet despite these encouraging results, the analysis also revealed areas for further refinement. Residual audits uncovered a lightly leptokurtic error distribution: while the bulk of forecasts clustered tightly around the true values, rare high-temperature spikes were under-predicted more often than would be expected under a Gaussian noise assumption. From a maintenance perspective, these extreme events are the most consequential, as sustained overheating can degrade insulation, accelerate bearing wear, and lead to unplanned shutdowns. Addressing this under-prediction bias could involve augmenting the training set with synthetic examples of severe operating conditions—perhaps drawn from high-fidelity multiphase flow simulations—or adopting loss functions that penalize underestimation more heavily, such as asymmetric quantile loss or cost-sensitive objective terms.

An additional insight emerged from our unsupervised clustering of lagged sensor data, which delineated five distinct operating modes: startup transients, normal steady state, high-temperature

stress, low-pressure conditions, and intermittent current spikes. Mapping prediction errors onto these regimes revealed that model performance dipped most noticeably during regime transitions—particularly when the system moved into high-temperature or high-current clusters. This suggests a promising two-stage hybrid architecture: a lightweight classifier first identifies the current operating mode, and then a dedicated, mode-specific regressor—trained only on data from that regime—produces the temperature forecast. Such specialization could reduce within-mode variance and improve robustness during rapid transitions.

From an implementation standpoint, we also weighed the computational footprint of our models. While XGBoost offered top predictive accuracy, its inference latency and memory requirements may be prohibitive for deployment on edge-embedded controllers at remote wellheads. In contrast, Random Forest or Gradient Boosting (with modest tree counts) might be more suitable for real-time inferencing on constrained hardware, especially if paired with an adaptive sampling strategy that increases prediction frequency only during identified high-risk modes.

Looking beyond the current study, several avenues invite exploration. First, integrating additional sensor modalities—such as vibration, acoustic emissions, or casing pressure—could furnish early indicators of mechanical wear or perforation influx that thermal data alone cannot capture. Second, deep-learning architectures like Temporal Convolutional Networks (TCNs) or Long Short-Term Memory (LSTM) autoencoders may excel at learning longer-range dependencies and non-linear couplings among variables, albeit at the cost of interpretability. Third, embedding our pipeline into a closed-loop monitoring system—with automated retraining triggered by newly labeled workover events—would allow the model to evolve alongside the asset and adapt to changing reservoir conditions or equipment configurations. Finally, coupling forecasted temperature trajectories with remaining-useful-life (RUL) estimation techniques could

deliver health indices rather than point predictions, facilitating decision-support tools that recommend optimal intervention schedules.

In sum, this work demonstrates that a disciplined, time-aware machine-learning workflow can produce accurate, interpretable forecasts of ESP motor temperature—laying the groundwork for a predictive maintenance regime that minimizes unplanned downtime, extends equipment life, and enhances operational efficiency. By iterating on data sources, model architectures, and deployment strategies, operators can progressively tighten the predictive loop and shift from reactive to truly proactive maintenance.

5.2 RECOMMENDATION

To translate our predictive maintenance framework into tangible operational benefits, we recommend integrating the optimized XGBoost motor-temperature model into a real-time monitoring platform at the wellhead. This platform should continuously ingest incoming sensor telemetry, apply the same preprocessing and feature-engineering steps, and generate one-hour-ahead temperature forecasts at regular intervals. Alerts can be configured to trigger when predicted temperatures exceed established safety thresholds, providing maintenance crews with advanced warning—ideally 6–12 hours before critical overheating—to plan inspections or adjust operating parameters without disrupting production.

Complementing this regression-based alerting, we propose deploying a lightweight operating-mode classifier that first categorizes the pump into one of the five identified regimes: startup, normal steady-state, high-temperature stress, low-pressure event, or current spike. Mode classification can serve two purposes: first, to select a specialized regression model tuned for that regime's unique dynamics (thus reducing forecast error during regime transitions); and second,

to govern adaptive sampling rates. For example, during “normal” operation, the system might run predictions every hour, whereas in “high-temperature stress” or “current-spike” modes, it could switch to 15-minute intervals to capture rapid developments more closely.

From a data-management perspective, we recommend establishing an automated feedback loop between the monitoring system and the historical database. Each time a workover event or motor failure occurs, the corresponding data window should be labeled and appended to the training corpus, and the model retrained—preferably in a scheduled nightly or weekly process. This continuous learning approach will enable the algorithm to adapt to gradual changes in reservoir conditions, equipment aging, and modifications in pump design or operating parameters. In parallel, incorporating additional sensor modalities—such as vibration signatures, acoustic emissions, or casing-pressure fluctuations—can enrich the feature space and potentially detect pre-failure patterns that temperature alone may not reveal.

Finally, to foster organizational buy-in and ensure the longevity of the predictive-maintenance initiative, we suggest developing a user-friendly dashboard that visualizes both real-time forecasts and historical performance metrics. Key performance indicators (KPIs) such as forecast accuracy, mean time between alarms, and avoided downtime hours should be tracked and reported to engineering and operations leadership on a monthly basis. By coupling technical excellence with clear, transparent reporting and feedback loops, the organization can steadily move from reactive repairs to a proactive maintenance culture—ultimately extending ESP life cycles, reducing intervention costs, and maximizing production uptime.

REFERENCES

- Abdalla, A. A., Al-Fattah, S. M., & Alshahrani, A. M. (2022). Real-time electric submersible pump failure prediction using machine learning. *Journal of Petroleum Science and Engineering, 208*, 109470. https://doi.org/10.1016/j.petrol.2021.109470
- Al-Ballam, A. M., Al-Malki, H. A., & Al-Suleiman, K. M. (2022). Statistical trend analysis and Weibull reliability models for ESP failure prediction. *Journal of Petroleum Technology, 74*(6), 85–95.
- Alhashem, A., Al-Azmi, H., & Al-Ragom, F. (2024). Unsupervised machine learning for ESP performance monitoring using SCADA data. *SPE Kuwait Oil and Gas Show and Conference*. https://doi.org/10.2118/216905-MS
- Aziz, A. A., Nasir, A., & Wahab, M. A. (2022). Machine learning based predictive maintenance model for oil and gas assets. *Journal of Engineering Science and Technology, 17*(3), 2108–2120.
- Baker Hughes. (2022). Hybrid models for predictive maintenance in gassy wells: A case study. Retrieved from https://www.bakerhughes.com/whitepapers/pdm-case-study
- Bansal, R., Gupta, S., & Kumar, A. (2020). ESP failure prediction using SCADA data: A case study in the North Sea. *Journal of Petroleum Technology, 72*(4), 45–52. https://doi.org/10.2118/123456-JPT
- Deloitte. (2023). Digital transformation in upstream oil and gas: Cost-benefit analysis of AI projects. Retrieved from https://www2.deloitte.com/content/dam/insights/us/
- Fakher, M. A., Al-Hussain, K. S., & Omar, A. H. (2021). Analysis of failure modes in ESP systems: Electrical, mechanical, and operational perspectives. *SPE Production & Operations, 36*(2), 145–159.
- Heng, L., Wei, Z., & Smith, J. (2019). Gradient-boosted decision trees for early detection of bearing wear in ESPs. *SPE Production & Operations, 34*(3), 210–225. https://doi.org/10.2118/195678-PA

- IBM. (2023). Predictive maintenance: Continually assessing equipment health in real time. Retrieved from https://www.ibm.com/topics/predictive-maintenance
- Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing physics-based models. *Mechanical Systems and Signal Processing, 20*(7), 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012
- Jin, W., Noori, M., & Al-Daham, M. (2014). Downhole current-loop telemetry for ESP sensor integration. *Journal of Petroleum Technology, 66*(10), 65–73.
- Kimray. (2023). Real-time ESP system data: Pump pressures, temperatures, and vibration for controller feedback. Retrieved from https://kimray.com/esp-sensors
- Kulik, J. (2024). Federated learning for cross-well generalization in predictive maintenance. *NeuroSYS*. Retrieved from https://neurosys.com/blog/federated-learning-oil-gas
- Petrofly. (2022). Synthetic data generation for ESP failure prediction using GANs. *Medium*. Retrieved from https://petrofly.medium.com/esp-synthetic-data-gans
- Pham, H., Elgaghil, S., & Bartos, S. (2021). Predictive maintenance of electric submersible pumps using machine learning methods. *SPE Middle East Artificial Lift Conference and Exhibition*. https://doi.org/10.2118/205553-MS
- PythonGuides. (2025). Cloud economics in predictive maintenance: A cost analysis. Retrieved from https://pythonguides.com/cloud-costs-pdm/
- Schlumberger. (2015). *Overview of electric submersible pump systems*. Retrieved from Schlumberger technical documentation.
- Society of Petroleum Engineers (SPE). (2021). Alarm management best practices in ESP operations (Report No. SPE-206432). Retrieved from https://www.onepetro.org/report/SPE-206432
- Song, J., Liu, X., Zhang, H., & Feng, D. (2024). Laboratory-based fault detection of ESPs using machine learning on multimodal sensor data. *Journal of Petroleum Science and Engineering, 226*, 109241. https://doi.org/10.1016/j.petrol.2023.109241

- SPD Technology. (2024). Edge computing for real-time ESP monitoring. Retrieved from https://spd.tech/edge-computing-oil-gas
- Takacs, G. (2017). *Electric submersible pumps manual: Design, operations, and maintenance*. Elsevier.
- Williams, D., Patel, R., & Sharma, S. (2014). Direct downhole measurements for ESP protection and control. *Journal of Petroleum Technology, 66*(2), 38–47.
- Zhang, Y., Wang, H., & Li, Q. (2021). Convolutional neural networks for cavitation detection in ESP vibration spectra. *Applied Energy, 298*, 117256. https://doi.org/10.1016/j.apenergy.2021.117256
- Zheng, Z., Wang, H., & Liu, H. (2004). Fault diagnosis of electric submersible pump sensors using support vector machines. *Petroleum Science and Technology, 22*(5–6), 647–662. https://doi.org/10.1081/LFT-120030223