Implementing the autonomous adaptive algorithm to manage ESP operation in harsh reservoir conditions

M Antonic¹, M Solesa¹, G Thonhauser², A B Zolotukhin³, M Aleksic⁴

¹Chair of Drilling and Completion Engineering, Department of Petroleum Engineering, Montanuniversität Leoben, Parkstraße 27, 8700 Leoben, Austria
²TDE Group Limited, Al Maqam Tower, ADGM Square, Abu Dhabi, UAE
³Gubkin Russian State University (NRU) of Oil and Gas, GSP-1, Leninsky Prospekt, 65, 119991, Moscow – Russia
⁴NIS Gazprom Neft, Narodnog Fronta 12, Novi Sad, Serbia

* Contact author: miroslav.antonic@unileoben.ac.at

Abstract. The well geometries with a shallow kick-off point in conjunction with surface infrastructure limitations have led to Electrical Submersible Pump (ESP) technologies’ application as one of the most suitable artificial lift methods for the harsh reservoir conditions. However, the harsh reservoir conditions in terms of the low reservoir pressure, high reservoir temperature, scaling problems in various forms, and high gas content at the pump intake have reduced the ESP system run life. Therefore, this research represents the Autonomous Adaptive Algorithm (A³) as a holistic approach to integrate analytical and machine learning models to assist production engineers in the early detection of operating problems. The A³ relies on different data sources and uses unique, well diagnostics logic to generate valuable features and prepare data for training. Finally, the paper evaluates different classifiers and explores the possibilities of application A³ as a flexible edge solution. The research benefits will be demonstrated for several problematic ESP wells.

1. Introduction

One of the core objectives of producing well optimization is minimizing operational costs while achieving the most favorable maintenance programs and designed production rates. Optimizing large-scale Artificial Lift (AL) operations and continuously evaluating their performance is a rather challenging task. The oil and gas industry has recognized the value of implementing digital technologies in AL optimization. However, the industry still faces many challenges to achieve the full value of digitalization [1].

The oilfield located in the South Pannonian Basin is characterized by low reservoir pressure, specific well geometry, deep interval depths, and unfavorable fluid properties. ESP (Electrical Submersible Pump) method was determined as the most suitable for application in these conditions. There are currently more than 42% of all producing wells in the field equipped with ESP, contributing around 73% to the total production. Due to the high share in the total production and high costs associated with installing the ESP, any system failure will result in significant differed production losses and capital expenditures (CAPEX) investment. These two aspects are major pain areas for operators. [2]
The opportunity to guide the decision-makers towards actions that would optimize the ESP performance is hidden in the real-time data from the downhole and surface sensors. Traditionally, sensors measurements are fed into the Supervisory Control and Data Acquisition System (SCADA) or any other distributed Control System, which is responsible for monitoring and controlling the equipment. These systems allow engineers to set alarms and predefined rules to keep the production parameters within designed envelopes [3]. However, the described approach represents a reactive maintenance program where an alarm is triggered when a change has already happened.

Recently, many authors worked on the implementation of predictive maintenance programs that involve the application of digital technologies. Abdelaziz at el. [4] described application of Principal Component Analysis as a tool to detect developing ESP failures and predict the remaining run life. Alamu et al. [3] used real-time data in combination with deep autoencoders to detect anomalous behavior in ESP operation. Other authors worked intensively on developing Deep Learning algorithms to analyze sensor behavior and detect trends [5]. Some other authors focused their research on analyzing the slope trends of different data channels [6] and applied fuzzy expert systems [7] to identify the root cause of the failures. More sophisticated approaches go beyond data and expert-driven modeling. New trends in ESP predictive maintenance include implementing digital twins and predictive models based on the integration of well models and real-time data. Several references show the implementation of previously mentioned models, where the ESP digital twins were used to mimic the physical behavior of the system and used to identify well operation problems. [8] Real-time data were used to automatically feed the validated well models in order to identify the production optimization opportunities. [9]

The ESP predictive models have already found their place in the companies' digital ecosystems. However, the scalability and applicability of these models remain questionable. Therefore, this study presents a comprehensive approach that integrates physical, expert, and data-driven modeling and investigates how such an integrated model can enrich the accuracy of ESP predictive models. Well-known empirical correlations describing a system's physical behavior are used to generate additional valuable feeds and get more insights into the system's health condition. After generating the new feeds, real-time data are correlated with episodic data. A unique data set was prepared for expert evaluation and historical data processing. The available data were labeled so that stable well operation conditions could be easily differentiated from other unplanned events. Eventually, the machine learning algorithms were trained to predict the well events.

2. Methodology
The predicted outcome of this research work is that a promising ESP diagnostic approach can be made if there is integration between physical, expert, and data-driven modeling. We assumed this approach is achievable if data are passed through three phases of generating additional insights into the ESP system health conditions.

2.1. Data preparation
The first step in a continuous improvement process is a performance measurement. Metrics established to measure the ESP performance are generally classified based on the sensor location (surface or downhole), the nature of the measured variables (electrical, hydraulic, or thermal), or the data acquisition approach. The application of advanced well monitoring and predictive maintenance procedures depends highly on data acquisition, aggregation, structuring, and storing technologies. Considering the data acquisition approach, all fundamental ESP performance data are separated into three major categories, as shown in Figure 1. Data acquired with high frequencies and constant sampling rates are treated as real-time data, while data documented in irregular intervals (few times per day/week/month) are treated as episodic data. For example, the downhole sensors represent real-time data, and fluid rates or water cut measurements belong to episodic data. The third category, named miscellaneous data, comprises the attributes that represent the ESP system or well itself. For instance, the miscellaneous data are the well completion details or ESP pump and motor performance curves. A more detailed overview of data
organization and structuring in the form of a comprehensive Petroleum Data Management System was given in related work [10].

![Figure 1. General workflow of the data preparation procedure](image)

In modern times, data are a principal asset that enables digital transformation. However, it is not easy to capitalize the value only by looking into raw data, and the data analysis starts with data preprocessing algorithms. Figure 1 depicts the general steps performed to prepare the raw data for further analysis. As shown in Figure 1, two different algorithms were deployed to preprocess data. The real-time data usually contained missing values, and the specific feature of Variable Speed Driver (VSD) responsible for pump control caused that events are recorded along with numerical values making discontinuities in the data. Therefore, the function was created to handle missing data and remove textual data. In contrast, the periodical datasets had discontinued recordings that generate problems in applying machine learning procedures. The second set of preprocessing functions included the tools for generating missing values so that the physical meaning of recording is kept (e.g., the production rates are filled with the nearest previous sample). Functions and code explained in the paper are built in the MATLAB programming language. Automating the preprocessing algorithms allowed data ingestion on a larger scale and their easy integration with the overall logic of the A³.

### 2.2. Physical models

In the first phase, the physical or well-known industry analytical/empirical models were used to generate missing feeds or create indicators closely related to determining the root cause of the operation problem. For illustration, the downhole sensor block in the available wells did not measure the pump discharge pressure. The pump discharge pressure is a valuable diagnostic tool of ESP downhole operating conditions and an essential parameter needed for system/nodal analysis. [11] Therefore, the pump discharge pressure (PDP) was calculated by relying on the validity of the periodical dataset, the extensive set of pressure-volume-temperature (PVT) correlations, and models to estimate the pressure drop in multiphase flow conditions. Beggs and Brill's model [12] was selected to describe multiphase flow due to its applicability in various well geometries (different inclinations and dogleg severities). The well-known routines for solving ordinary differential equations facilitated the implementation process. Additionally, PDP was calculated using the Hazen-Williams model [13] to estimate the frictional pressure drop in the tubing for quality checking purposes. By using the pump performance curve, and the total number of stages, the final check, and comparison of the pump head are made. **Figure 2** represents the sequences in which physical parameters are calculated and incorporated into the original dataset.
The given wells at the oil field of the South Pannonian basin had failures mostly linked to excessive gas production and scale problems caused by the low bottomhole pressure and high fluid temperature. Once the gas enters the pump, it restricts its volumetric efficiency, degrade the pump performance, causing a decline in the pump head and fluid flow rate. Therefore, workflow (Figure 2) shows that indicators exploring the influence of the gas were calculated. At this research stage, the black oil calculations were performed to calculate the amount of gas at the local conditions (at the pump intake). The utilized models showed very good matching with fluid properties available in the company's reports. The presence of the free gas at the pump intake is in accordance with low pump intake pressure (PIP), but also it depends on the well completion details and gas handling devices. Since the considered wells were completed without a packer and with a gas separator, the Alhanti mechanistic model [14] was applied to estimate the natural separation efficiency. Consequently, the actual gas amount ingested by the pump was calculated by using the total gas-phase volumetric rate and total efficiency of the gas separation. Turpin coefficient [13], as an essential key performance indicator, is used to differentiate between stable and unstable pump operation. Turpin coefficient was calculated on a daily level and is used in the expert-driven as well as in data-driven modeling.

2.3. Expert-driven modeling
The second phase of the predictive maintenance model is expert-driven modeling. In this phase, the expert knowledge is translated into labels or classifications of the ESP working conditions. Expert modeling means that well operation symptoms (e.g., pump intake pressure, downhole temperature, current, etc.) are observed in representative periods. During the behavior observation, the expert is able to recognize and evaluate patterns. Pattern evaluation usually means that the trends in the data are quantitatively or/and qualitatively described. Subsequently, the expert rules built on the practical experience relate the behavior of individual trends to identify a developing problem.

In the research, historical data were ingested, and the long history of the ESP operation has been analyzed on a day-by-day basis. Each noted event, change in the shape of parameters, or any detected anomalies are classified during the expert-driven observation and labeling procedure. In many cases, the labeling process had to be done for relatively short periods because changes in operating parameters happened suddenly with having unregular patterns and trends. All these changes were captured during the expert-driven analysis based on the formalized rules. The observed labels are listed in Table 1.
Table 1. List of labels used in the expert-driven modeling

| Label                   | Meaning                                                                 |
|-------------------------|-------------------------------------------------------------------------|
| Stable operation        | Minor or no changes in operating parameters                            |
| Unstable operation      | Small fluctuations in the majority of the key symptoms                 |
| Erratic operation       | More pronounced fluctuations in the majority of the key symptoms       |
| Pump trip               | Pump idle time                                                          |
| Power outage            | Sudden loss of all sensor information due to instabilities in power supply |
| Undercurrent condition  | Current decreasing due to operating in the underload condition, mainly caused by the high gas presence |
| Insufficient cooling    | The sudden increase in the downhole temperature                         |
| Restricted inflow       | Identified problem on the sandface because of the near-wellbore zone damage or problem at the pump intake |
| Scale Build-up          | Increased resistance to flow reflected in the change of current, downhole temperature, and PIP |
| Gassy condition         | Fluctuations in current followed by increased free gas at the pump intake and within the pump, while Turpin coef. <1 |
| Near gas lock           | More severe influence of gas, Turpin coef. >1                           |
| Failure                 | End of the operation                                                    |

To increase the level of automation in the data labeling process, data were visualized in a few different contexts. Figure 3 demonstrates the graphical representation of episodic and real-time data for the selected case. The same figure shows that some data have been duplicated, while some are not shown at all. The PIP is available as a real-time recording, but also it is aggregated on a daily level to allow the necessary implementation of physical models. The interpretation of the real-time data can be a cumbersome procedure, thereby to aid the expert data labeling, the first derivative of the time series was calculated and used to help in the qualitative description of changes in symptom trends and patterns.

The applied A3 methodology automatically generates pump performance curves and amper charts, giving ESP operating symptoms in more common forms for further diagnostic evaluation (Figure 4). Figure 4 shows the modeled pump head data along with the pump performance curve calculated using the pump operating frequencies and recorded flow rates. On the right side of the figure, the fluctuations of the current caused by the severe impact of free gas are correlated with head changes.
To summarize the expert-driven approach, an algorithm is presented in Figure 5. Since the work has been done with a large amount of different data types, data had to be managed appropriately. The File Ensamble Datastore as an object type was used to store time series (real-time data) and numeric (episodic) data along with a label assigned during the expert analysis.
2.4. Data-driven modeling

Machine learning models have been tested to make a predictive model that will automatically recognize the operation states, and for example, detect the unstable operation that usually comes before the pump trip. Since the data ensemble contains time-series and numeric data, it was necessary to generate features representing summarized information of original data. There were around one hundred generated features in the time and frequency domain. In the time domain, signal-based statistical metrics have been calculated, such as: mean, standard deviation, root mean square (RMS), shape factor, kurtosis, skewness, and peak values. In the frequency domain, features like peak amplitude, peak frequency, and band power were calculated. The features were determined for original time-series data (e.g., current) and for some derived feeds extracted for expert-driven models (e.g., the first derivative of current). The Analysis of Variance (ANOVA) statistical method was used to rank features (Figure 6) to avoid overfitting and keep only indistinctive features that will help estimate the ESP operation accurately. After the feature ranking, the most indistinctive features were those related to statistical metrics (mean and standard deviation) of load, power, and current data channels. The final number of features was reduced to 51.

2.4.1. Machine learning. The motivation for applying machine learning in predictive maintenance comes from the complexity to explicitly program the expert rules that will be enough general to detect the ESP operation problem/status. Another advantage of using machine learning algorithms in predictive maintenance is that improvements in hardware technologies follow advancements in artificial intelligence. The innovation in Industrial Internet of Things (IIoT) technologies has increased the connectivity possibilities, and practically, machine algorithms could be deployed on the field, perform data interpretation, and transmit only critical key performance indicators.
2.4.2. Machine learning. The motivation for applying machine learning in predictive maintenance comes from the complexity to explicitly program the expert rules that will be enough general to detect the ESP operation problem/status. Another advantage of using machine learning algorithms in predictive maintenance is that improvements in hardware technologies follow advancements in artificial intelligence. The innovation in Industrial Internet of Things (IIoT) technologies has increased the connectivity possibilities, and practically, machine algorithms could be deployed on the field, perform data interpretation, and transmit only critical key performance indicators.

In this research, the implementation of machine learning models was the final stage. In order to verify the value of integration of different models, multiple machine learning models have been trained. In the first attempt, the models were trained only with features extracted from real-time data. The maximal model accuracy of different classification algorithms (Neural Networks, Support Vector Machines, etc.) was around 70%. In the second attempt, the features and episodic data (including calculated parameters, e.g., pump head) were passed to the same architecture of classifiers, and the maximum model accuracy of 78% was achieved by one of the ensemble learning algorithms (AdaBoost). A similar performance was reached with decision tree algorithms (76.6%). The obtained accuracy depends on the amount of available data, and future research work will be focused on increasing the number of wells/cases to include additional labels and assure enough data for the model training. It is expected that this additional work will increase the model validation and classifier accuracy level so that timely decisions can be taken on time.

Figure 7 shows the confusion matrix for the multiclass classification algorithm with the highest accuracy. Additionally, tuning of hyperparameters could be done to achieve better accuracy and get more valid statistical error estimation. However, that was not the objective of the research at this stage. The figure shows that stable operation is predicted with the highest accuracy. However, the unbalanced dataset resulted in relatively low accuracy for other events, especially for the detection of unstable operations. This leads to the conclusion that involving additional physical models such as heat transfer modeling is essential and could help in a better and more precise evaluation of scale problems.
3. Conclusions
The presented approach is a part of the research work on A³ that will be used for the IIoT deployed as an edge solution. The model looks at the real-time raw data from a different perspective by extracting health indicators, predicting the events, and suggesting actions. This paper explored how successfully physical, expert, and data-driven models can be integrated to enable efficient predictive maintenance of ESP systems. The conclusions drawn from this research are following.

- Empirical correlations and physical models that have been developing over the years in the oil and gas industry are valuable input for data-driven predictive maintenance algorithms, and an increase in accuracy is observed.
- The integration of different models and datasets has been successfully applied, and predicting ESP operating conditions was done with an accuracy of 78%.
- The proposed model generalization can be improved by including the new training examples and involving new parameters.

Further research developments will be focused on generating additional episodic indicators (specific power consumption, scaling index, shaft stress, and load indicator, etc.), which will further improve the accuracy of machine learning models.

![Figure 7. Confusion matrix for AdaBoost machine learning algorithm.](image-url)
Abbreviations
ESP Electrical Submersible Pump
VSD Variable Speed Drives
A³ Automatic Adaptive Algorithm
AL Artificial Lift
SCADA Supervisory Control and Data Acquisition System
QC Quality Control
PVT Pressure Volume Temperature
PIP Pump Intake Pressure
PDP Pump Discharge Pressure
IIoT Industrial Internet of Things

References
[1] Carvajal G, Maucec M and Cullick S 2017 Intelligent digital oil and gas fields: Concepts, collaboration, and right-time decisions (Amsterdam: Gulf Professional Publishing) p 13.
[2] Andrade MA, Busaïdy S, Murad M, Balushi I, et al 2019 ESP Well and Component Failure Prediction in Advance using Engineered Analytics - A Breakthrough in Minimizing Unscheduled Subsurface Deferments Proc. of Abu Dhabi Int Pet Exhibition Conf (Abu Dhabi, UAE, 11-14 November) (Society of Petroleum Engineers)
[3] Alamu O A, Pandya D A, Warner O and Debacker I 2020 ESP Data Analytics: Use of Deep Autoencoders for Intelligent Surveillance of Electric Submersible Pumps Proc. of Offshore Technol Conf (Houston, Texas, USA, 4-7 May) (Offshore Technology Conference)
[4] Abdelaziz M, Lastra R and Xiao J J 2017 ESP Data Analytics: Predicting Failures for Improved Production Performance Proc. of Abu Dhabi Int Pet Exhibition Conf (Abu Dhabi, UAE, 13-16 November) (Society of Petroleum Engineers)
[5] Cimmino M, Calabrese M, Kapetis D, Lygren S, Vanzan D, Piantanida M, Carminati S, Esposito G, Raimondi L and Amendola A 2020 Predictive Modelling and Monitoring of Industrial Equipment Through Advanced Deep Learning Algorithms Proc. of Int Pet Technol Conf (Dhahran, Saudi Arabia, 13 – 15 January) (International Petroleum Technology Conference)
[6] Bermudez F, Carvajal G A, Moricca G, Dhar J, Md Adam F, Al-Jasmi A, Goel H K and Nasr H 2014 Fuzzy Logic Application to Monitor and Predict Unexpected Behavior in Electric Submersible Pumps (Part of the KwIDF Project) Proc. of SPE Intell Energy Conf Exhibition (Utrecht, The Netherlands, 1-3 April) (Society of Petroleum Engineers)
[7] Grassian D, Bahatem M, Scott T and Olsen D 2017 Application of a Fuzzy Expert System to Analyze and Anticipate ESP Failure Modes Proc. of Abu Dhabi Int Pet Exhibition & Conf (Abu Dhabi, UAE, 13-16 November) (Society of Petroleum Engineers)
[8] Lastra R 2019 Electrical Submersible Pump Digital Twin, the Missing Link for Successful Condition Monitoring and Failure Prediction Proc. of Abu Dhabi Int Pet Exhibition & Conf (Abu Dhabi, UAE, 11-14 November) (Society of Petroleum Engineers)
[9] Kulkarni R, Al-Bimani A, Marin A A, Awaïd A, Balushi I, Al-Busaïdy S, Al-Murad M, Kumar N, Gala R and Marin E 2018 Towards a Fully Automated Well Management System for ESP Surveillance and Optimization in PDO Proc. of SPE Middle East Artificial Lift Conf and Exhibition (Manama, Bahrain, 28-29 November) (Society of Petroleum Engineers)
[10] Antonic M, Solesa M, Zolotukhin A B, Rakic D and Aleksic M 2019 Integration of expert and data-driven workflows to manage reservoir and well life cycle in Arctic conditions using innovative SICLO methodology, IOP Conf. Ser.: Mater. Sci. Eng. 700, 012056
[11] Bruijnien P M 2016 Nodal Analysis by Use of ESP Intake and Discharge Pressure Gauges, SPE Prod & Oper 31, 76–84
[12] Beggs D H and Brill J P 1973 A study of two-phase flow in inclined pipes, J Pet Technol 25, 607–17
[13] Takacs G 2018 Electrical submersible pumps manual: Design, operations, and maintenance (Cambridge MA: Gulf Professional Publishing) p 175.
[14] Alhani FJS 1993 Bottomhole Gas Separation Efficiency in Electrical Submersible Pump Installations PhD dissertation University of Tulsa