Dynamic Linear Regression Models for Down Hole Safety Valve Remaining Useful Life Prediction

Harrouz Aymen[1], Toubakh Houari[1], and Kafi Mohamed Redouane[1], Moamar Sayed-Mouchaweh[2], Hajer Salem[3]

[1]*University of Kasdi Merbah Ouargla, Ouargla, 30000, Algeria.* [2]*Institute Mines-Telecom Lille Douai, Douai, France.*
[3]*Audensiel Technologies, Pôle R&D, Boulogne-Billancourt, France*

Harrouz.aymendjamel@univ-ouargla.dz
Toubakh.houari@univ-ouargla.dz
Kafi.redouane@univ-ouargla.dz
Moamar.sayed-mouchaweh@imt-lille-douai.fr
h.salem@audensiel.fr

ABSTRACT

Down hole safety valve plays and important role in the safety of the constructed wells of oil and gas; failures in this element are very problematic to both, the environment and the production losses; down hole safety valve failure prediction is critical to prevent such catastrophic events. This paper provides a comparative study and proposes a new feedback mechanism to enhance the performance of linear regression models in predicting the Remaining Useful Life of the Down Hole Safety Valve used in oil and natural gas well completion operations. The data used are part of the publicly available 3W database developed by Petrobras, the Brazilian oil holding. Three variations of linear regression models were investigated, and evaluated using prognosis performance metrics. Two experiments were conducted to evaluate the performance of the system and its reaction to new data, coefficient of determination reached 94% on average indicating a successful prediction with 95% confidence level reached once the feedback mechanism is applied.

1. INTRODUCTION

Natural gas is considered to be playing a crucial and important part in fulfilling energy demands, its reliability and cost-efficiency for power generation has been proven; and it is not stopping there, moreover it is expected to be the fastest growing major fuel source of the next two decades. Natural gas has many environmental benefits due to its low emissions, and reduced surface footprint. It is considered to be the cleanest burning fossil fuel, and less carbon intensive than other fossil fuels (XTO Energy, 2019).

Aymen Harrouz et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Oil and natural gas extraction goes through multiple steps and procedure; one of the final operations is "Well Completion", and it is the process of descending the production tubing into the cemented casing, and completing the well construction to make it ready for production. The production tubing contains numerous elements, some are necessary, others are complementary. One of the main elements of the completion tubing is the Down Hole Safety Valve (DHSV), this last have multiple names in the industry here are a few of them: Tubing Retrievable Safety Valve (TRSV), Surface Controlled Subsurface Safety Valve (SCSSV), Wireline Retrievable Safety Valve (WRSV), and many others, each one of them has some unique properties but they all share the same purpose. DHSVs are one of the critical element during the completion process of the well, as these are fail-safe devices, and act to prevent an uncontrolled release of oil and natural gas, that will cause a surge in the pressure inside the wellbore and create a blowout situation (Resato International, 2020).
Failure in the DHSV can be very detrimental, both to the environment as well as to the production losses and costs of repair. In general, the failure that occurs in the DHSV is the spurious closure, where there the DHSV closes and cuts the production; and in order to correct this failure a tool called 'Lock-Open Tool' is ran inside the production tubing to bypass the DHSV and resume the production.

To get a visual representation of a production well. Figure 1 presents a schematic of an offshore well. The oil and gas flow from a reservoir through production tubing and then through a production line to a platform. A subsea Christmas tree is a type of equipment installed on the seabed and is basically composed of valves and sensors operated remotely through an electro-hydraulic umbilical. A Permanent Downhole Gauge (PDG) and a Temperature and Pressure Transducer (TPT) are devices that contain pressure and temperature sensors, respectively. The PDG remains fixed in a certain position of the production tubing, and the TPT is part of the subsea Christmas tree (Vargas et al. 2019).

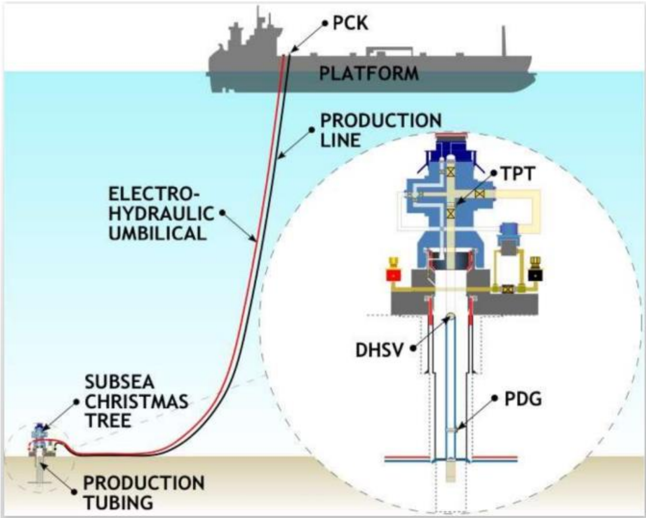

Figure 1: Simplified schematic of a typical offshore naturally (Vergas et al., 2019)

The time taken for the correction or the fix to take place is called Non-Productive Time (NPT). NPT is the time when any operation on the rig is interrupted for whatever reason, it is used as a measure of the effectiveness of the rig site operations. Usually it is represented as a percentage of idle time with respect to the total operation time (Krygier, Solarin and Ivanka, 2020), (Emhanna, 2018).

NPT and many other negatively influencing measures can be avoided or at least lowered if they have been correctly forecasted or foreseen. This is where fault prognosis techniques come in place. Fault prognosis is to use of current and past machine condition in order to predict how much time left for the failure to occur. The time left for the failure to occur is called Time to Failure (TTF), or the Remaining Useful Life (RUL) (Xiao-Sheng, Wenbin, Chang-Hua and Dong-Hua, 2011). Information about the fault propagation and failure mechanism is needed in order to predict the RUL. These mechanisms are found in sensor measurements and condition indicator that can be later used as features (inputs) for the prediction algorithm.

Prognosis methods fall into three main categories: Statistical/probability based techniques, Model-based techniques, and Data-driven techniques; each with its advantages and disadvantages (Salem & Sayed-Mouchaweh, 2020). The simplest form of prognosis techniques is the statistical approach; it is done by collecting statistical information from numerous component samples in order to indicate the time left for the failure to occur (Tran and Yang, 2009). Logistic regression has been used by Yan, Koc and Lee (2004), to calculate the probability function of the failure occurrence and used an autoregressive moving average to trend the condition variable for failure prediction. The advantage of statistical approaches that they do not require condition monitoring, population characteristics enable longer-range forecast, however, they only provide general estimates for the entire population of identical unites (Yan et al. 2004), (Phelps, Willett and Kirubarajan, 2001), (Banjevic and Jardine, 2005), (Vlok, Wnek and Zygmunt, 2004), (Chinnam and Baruah, 2003), (Kwan, Zhang, Xu and Haynes, 2003), (Lin and Makis, 2004), (Wang, Scarf and Smith, 2000), (Wang, 2002).

If a mathematical model is available, Model-based prognosis approaches can be constructed from physical systems. Features of these models are the residuals, which are the differences between the sensed measurements and the outputs of the mathematical model (Tran and Yang, 2009) (Derbal & Houari, 2018). Two defect propagation models were introduced by Li, Billington, Zhang, Kurfess, Danyluk and Liang (1999), a mechanistic modeling was used to estimate the RUL of bearings (Li, Kurfess and Liang, 2000).

Data-driven techniques or also known as machine learning techniques, use a large amount of historical failure data in order to build and train a prognostic model that learns the system behavior (Houari T., 2021). Due to the flexibility in generating a prognosis model, artificial intelligence techniques were used regularly. Artificial Neural Networks (ANN) were used in several approaches to model the system and estimate the RUL. Self-organizing neural networks were used by Zhang and Ganesan (1997) for multivariable trending of fault development in a bearing system in order to estimate its residual life.

Failure prediction has a very crucial impact on cost cutting and NPT reduction. We can imagine two scenarios, the first one is when using a failure prediction approach, the second one is when there is no use of a failure prediction approach; the first scenario will have less NPT than the second. This point of view can be applied to many engineering or manufacturing industries, however it is closely related to the oil and natural gas extraction field. The difference in production is very crucial to both, the client and the service company, where a decent amount of time is used in planning, preparing and transporting tools, equipment and trained personnel to the rig site to conduct the maintenance job. The more oil or natural gas produce the better, hence, the need for failure prediction is emphasized.

The major contributions of this paper can be stated as follows:

- Comparative study of three linear regression models to predict the RUL of the DHSV;
- A new proposed performance enhancement scheme using a feedback loop mechanism improving the RUL prediction.

The algorithm is evaluated using failure data of DHSV in order to predict its RUL. The results are evaluated using performance prognosis metrics mentioned in the article

(Saxena, Celaya, Balaban, Goebel, Saha, B., Saha, S. and Schwabacher, 2008).

This paper is organized as follows: Section 2 describes the function of the DHSV, failure description, and the used datasets; Section 3 introduces the proposed approaches and the workflow; experimentation methodology and the evaluation process; Experiments and results are discussed in section 4. Finally, Section 5 closes the paper emphasizing its main contributions.

2. CASE STUDY

2.1. Down Hole Safety Valve Working Principle and Failure

Down Hole Safety Valve is a Completion tools, considered as a safety device that allows the oil or natural gas to flow from the reservoir up to the surface (Brown, 1984), (Purser, 1997), however, in case of an emergency or physical disconnection of the hydraulic control line Figure 2, it closes automatically blocking the oil and natural gas to flow to the surface therefore avoid any surface spillage.

Figure 2 shows a vertical cut on a surface controlled subsurface safety valve, the working principle of a SCSSV is as follows: starting from the closed position, pressure flows from the surface though a hydraulic control line into the yellow chamber to press the control sleeve, the control sleeve presses down on a spring that has a Fail-safe mechanism in order for the flapper to open; this is the opening sequence of the SCSSV. Pressure inside the hydraulic control line can be controlled from the surface to close the flapper in case of a detected emergency; in addition, any damage to the hydraulic control line will result in a pressure drop and the flapper closes. Sometimes the closure function fails in a spurious manner without any indication on the surface, in order to correct this failure, an exercising tool is deployed into the production tubing to force the flapper back to its normal state, if this procedure fails, a 'Lock-Open Tool' is ran into the production tubing to lock the flapper in its open state.

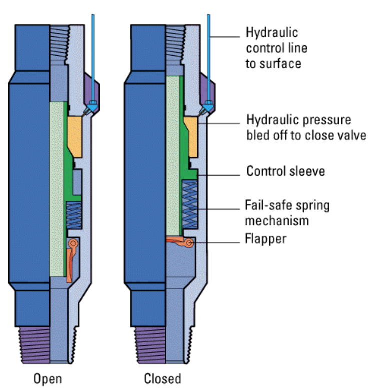

Figure 2 : Down Hole Safety Valve Vertical Cut and Components (Garcia, Jacinto, Lima, Pires and Droguett, 2006)

2.2. 3W Dataset description

The data used for this article is the 3W Dataset. It is structured as described in Table 1 and is available at (Vargas, Munaro, Ciarelli, Medeiros, Amaral, Barrionuevo, Dias de Araújo, Ribeiro, Magalhães, 2019). It is a collection of real sensor measurements taken from real operation in oil wells; the dataset also contains a collection of simulated data and sketched data. All instances were generated with observations obtained with a fixed sampling rate (1 Hz).

The dataset contains nine classes of events; the first class is the normal behavior; the eight other events are different faulty behaviors: 1) Abrupt increase of Basic Sediment and Water (BSW); 2) Spurious closure of DHSV; 3) Severe slugging; 4) Flow instability; 5) Rapid productivity loss; 6) Quick restriction in Production choke (PCK); 7) Scaling in PCK, and 8) Hydrate in production line.

Table 1: Quantitative Description of 3W database per event type.

Class	Description	Real	Simulated	Sketched	Total
0	Normal	597	0	0	597
1	Abrupt BSW Increase	5	114	10	129
2	Spurious DHSV Closure	22	16	0	36
3	Severe Slugging	32	74	0	106
4	Flow Instability	344	0	0	344
5	Rapid Productivity Loss	12	439	0	451
6	Quick PCK Restriction	6	215	0	221
7	Scaling in PCK	4	0	10	14
8	Hydrate in Prod. Line	3	81	0	84
	Total	1025	939	20	1984

The faulty behavior data contains full measurements starting from the healthy state, passing through the transient period, ending with the faulty steady state; Table 1 gives a quantitative description of 3W database per event type, where we count the number of instances belonging to each class. The number of instances is unbalanced due to the rarity of the events recorded; taking the "Scaling in PCK" event as an example, which is considered as a rare event in the industry (Vergas et al., 2019).

Each acquisition session contains eight process variables of different pressure and temperature values, in addition to the Gas life flow rate, the description of each measurement is presented in Table 2, with the corresponding units.

Table 2: List of tags in 3 W database, including tag names, descriptions, and measuring units.

Name	Description	Unit
P-PDG	Pressure at permanent downhole gauge (PDG)	Pa
P-TPT	Pressure at temperature/pressure transducer (TPT)	Pa
T-TPT	Temperature at temperature/pressure transducer (TPT)	℃
P-MON-CKP	Pressure upstream of production choke (CKP)	Pa
T-JUS-CKP	Temperature downstream of production choke (CKP)	℃

P-JUS-CKGL	Pressure downstream of gas lift choke (CKGL)	Pa
T-JUS-CKGL	Temperature downstream of gas lift choke (CKGL)	℃
QGL	Gas lift flow rate	m^3/s

3. PROPOSED APPROACH

In this section, the three proposed linear regression algorithms are described, in addition to the workflow of systems used in the comparative study, and most importantly the feedback loop mechanism used to improve or enhance the performance of the prediction; the validation process and method are also described alongside the prognosis performance metrics used to evaluate this work.

3.1. Linear Regression

Linear regression is used to study the linear relationship between a dependent variable (considered as the response to the regression algorithm), and one or more dependent variables (considered as features for the regression algorithm).

Linear regression can be univariable or multivariable, in many cases, the use of one dependent variable in inadequate to explain the dependent variable behavior; in such cases, the use of a multivariable linear regression is needed.

Multivariable linear regression model can be described in the following equation:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n \quad (1)$$

Where:

Y : is the dependent variable.

X_i : is the independent variable.

b_i : is the regression coefficient of X_i.

a : is a constant (Y intersect).

Where the dependent variable in this study is the Remaining Useful Life (RUL) of the DHSV, and the independent variables are the four tags (P-PDG, P-TPT, T-TPT, and T-JUS-CKP)

If we implement the previous variables in equation (1) we will get:

$$RUL = a + b_1PPDG + b_2PTPT + b_3TTPT + b_4\,JUSCKP \quad (2)$$

3.2. Stepwise Regression

Robust regression model needs to include independent variables that describe and explain a large part of the dependent variable. Variable selection methods vary from forward selection to backward selection and even stepwise selection (Fahrmeir, Kneib and Lang, 2009), (Schneider, Hommel and Blettner, 2010).

Forward variable selection, is a method that keeps adding these variables: P-PDG, P-TPT, T-TPT, and T-JUS-CKP; to the model as long as they contribute towards explaining the dependent variable which is the RUL in our case.

Backward variable elimination on the other hand, starts with a model that has all the independent variables (P-PDG, P-TPT, T-TPT, and T-JUS-CKP), the variables that worsen the prediction of the dependent variable (RUL) are then removed starting from the ones that has the most effect, to the least effect; this part is iterative until all the worsening variables are removed.

The stepwise variables selection combines aspects from forward selection and backward elimination. It starts with a null model, like the forward selection, and adds an independent variable that most describes the dependent variable and it keeps iterating. In addition, in every iteration a test is performed to see whether one of the added variables became invaluable, if this is the case, then the variable shall be removed. This selection is done to the four tags (P-PDG, P-TPT, T-TPT, and T-JUS-CKP) to test their performance and verify their value to predict the RUL.

3.3. Interactions Linear Regression

An interaction occurs not only when one independent variable has an effect on the dependent variable, but also its relation with other independent variables (Hong, Gui, Baran and Willis, 2010), (Ning and Hao, 2017).

If we consider a typical regression equation without interactions:

$$Y = a + b_1X_1 + b_2X_2 \quad (3)$$

However, with the inclusion of the interaction effect the previous equation will be as follows:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_1X_2 \quad (4)$$

Where:

X_1X_2 : is the interaction between X_1 and X_2.

This is called a two-way interaction, because there are two variables interacting with each other; higher order interactions are possible, however, the number of terms of the regression equation follows:

$$Number\ of\ terms = 2^{Interactions\ order} \quad (5)$$

If we pick two variables from our study the equation will become:

$$RUL = a + b_1.PTPT + b_2.TTPT + b_3.PTPT.TTPT \quad (6)$$

If we insert all four variables we would get 16 terms:

$$Number\ of\ terms = 2^4 = 16 \quad (7)$$

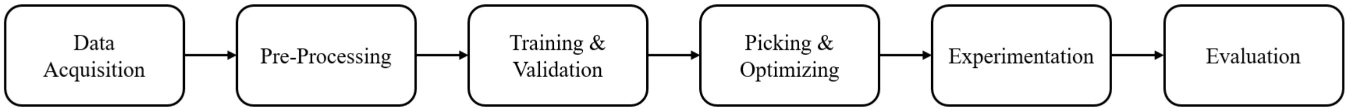

Figure 3: Block Diagram of the proposed comparative study system.

3.4. System Overview

3.4.1. Comparative Study System

The proposed system is represented in Figure 3.**Error! Reference source not found.** which shows the data flow starting from its raw form to be acquired to the system, passing through a pre-processing phase where modification is done to the data by keeping only the transient period where the deterioration started to reduce the procession power and concentrate only on the degradation phase; and adding the RUL variable from the last healthy behavior timestamp to the first faulty behavior timestamp to compensate for the absence of the RUL variable in the dataset. The pre-processed data are fed to the training model, the validation process using K-fold cross validation aims at validating the model for unseen instances in the training phase, the prognosis performance metrics were calculated afterwards; The best performing models will be picked to be optimized by changing their hyper parameters. Once the model is fully developed the testing phase takes place by feeding a new set of data to the model and check its performance with untrained data, and evaluate the results using prognosis performance metrics discussed in section 3.5.2.

3.4.2. Feedback Loop Mechanism

The best performing model from the comparative study is chosen for a furthermore performance improvement.

The performance improvement is conducted through using a new proposed scheme. The main idea is to influence the features used for the prediction; to do that a feature vector will be added to the feature space, therefore, instead of using "n" number of features, the model will be using "$n+1$" features. The implementation of this scheme will be discussed in this section.

The main questions are, where the added feature will be integrated from; how will this feature be added to the feature space; and can this feature be updated to provide a dynamic nature to the algorithm.

The added feature will be integrated from the previously predicted RUL, meaning that we will be using a Feedback Loop Mechanism in order to integrate the new feature in the feature space; therefore, at the start, the added feature vector will be empty due to the lack of predicted RUL data. After the first iteration the added feature will be constantly updated to provide a dynamic nature to the algorithm.

To further understand the proposed scheme Figure 4 shows a block diagram where the Feedback Loop Mechanism is implemented.

The "Training dataset" is fed into the trained model containing features (sensor measurements) and real RUL. The output of the "Model Training" block is a prediction function used to predict new data. The next step is to test the prediction function with a new "Testing dataset", the output of this phase is the predicted RUL. This last is used for computing the prognosis performance metrics, but most importantly it is used as a feedback feature to retrain the model and test it.

The update scheme is implemented following these main steps:

1. In the first iteration ($i = 0$), a new empty feature vector is created and used for training, however it won't have any effect due to its emptiness, therefore it will neither benefit or hurt the prediction.
2. The prediction function is created and will be used to test the model on new unseen data.
3. The model is tested and the predicted RUL is computed.
4. In the second iteration ($i = 1$), the empty feature vector is filled with the predicted RUL computed in the first iteration.
5. The new feature vector is used alongside the system features (sensor measurements) to retrain the model.
6. The model testing is done for the second time and the predicted RUL is computed.
7. The algorithm keeps iterating until satisfactory results are obtained, or a conversion happens.

In the fifth step, we kept the original features (sensor measurements) unmodified and unweighted to let the added feature act as a support and not as an original feature this will ensure that the prediction is not deviated from the original system.

This feedback mechanism can be described using the following set of equations for n iteration in the case of a regular Linear Regression:

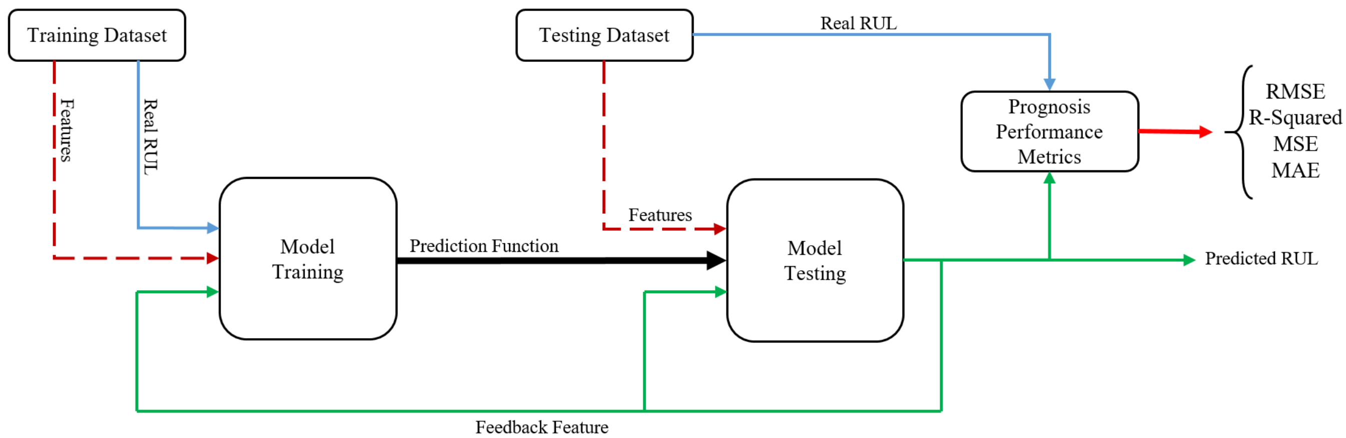

Figure 4: Block Diagram of the proposed Feedback Mechanism for performance enhancement.

$$
\begin{aligned}
Y_1 &= a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + 0 \\
Y_2 &= a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \hat{Y}_1 \\
Y_3 &= a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \hat{Y}_2 \\
&\vdots \\
Y_n &= a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \hat{Y}_{n-1}
\end{aligned} \tag{8}
$$

Or as

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} a & b_1 & b_2 & \cdots & b_n \\ a & b_1 & b_2 & \cdots & b_n \\ a & b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} 1 \\ X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} + \begin{bmatrix} 0 \\ \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_{n-1} \end{bmatrix} \tag{9}
$$

3.5. Evaluation Methodology

In order to evaluate the obtained results, prognosis performance metrics were computed using a K-fold cross validation technique. In this section, the validation and the performance metrics are explained.

3.5.1. K-Fold Cross Validation

During the model training, K-Fold cross validation technique is performed to evaluate and compare the performance of different machine learning models on the same dataset; by giving results of the training accuracy, and a broad picture of the testing accuracy. The number of folds is set to five due to the lack of differences in accuracy estimation with higher fold cross validations, compared to the processing time; 5-Fold cross validation divide the training set into five equal sections, the training is done using four sections and the fifth one is used for testing, then in the second iteration four different sections are selected including the section used in testing in the last iteration, and one is left for testing; the algorithm keeps iterating for five times and until all the sections have been used for training and testing the model; Figure shows the cross validation method.

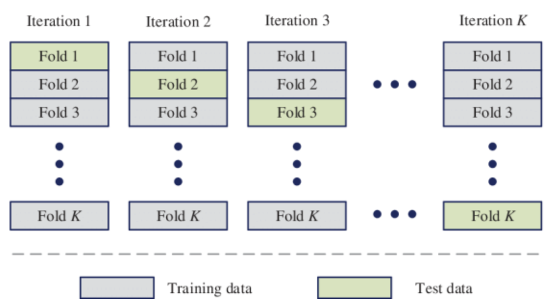

Figure 5: K-Fold cross validation method (Qiubing, Mingchao and Shuai, 2019).

3.5.2. Prognosis Performance Metrics

Evaluation of the model training and testing is done using the following prognosis performance metrics:

- **Root Mean Squared Error**

Root Mean Squared Error is used to evaluate the standard deviations of the residuals of the prediction results (the difference between the predicted and the observed data), equation 5 describes the $RMSE$:

$$RMSE = \sqrt{\overline{(\hat{y} - y)^2}} \tag{10}$$

Where:

$\hat{y}$: is the predicted response.

y : is the observed response.

- **Mean Squared Error**

Mean Squared error evaluates our model in terms of the variance of the residuals, using the following equation:

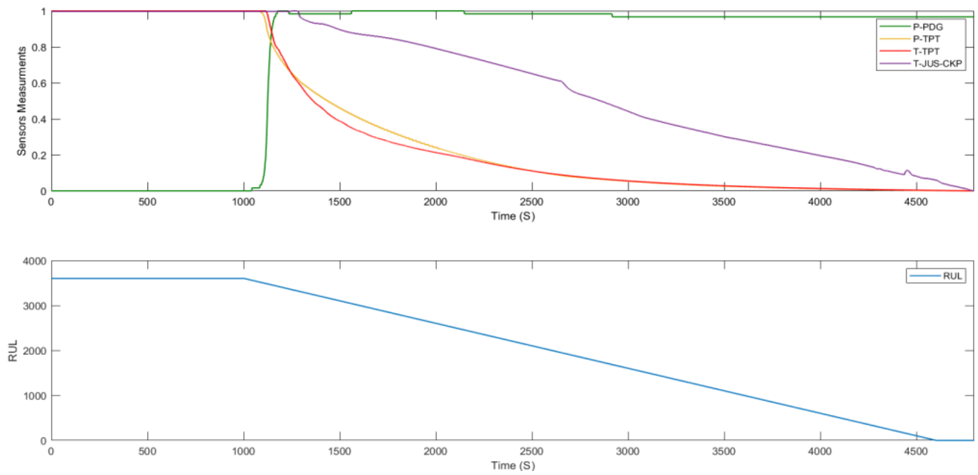

Figure 5: Sensor measurements used in model training (Plot 1), Remaining Useful Life (Plot 2).

$$MSE = \overline{(\hat{y} - y)^2} \quad (11)$$

- **R-squared (coefficient of determination)**

the proportion of the variance in the dependent variable that is explained by the linear regression model. It is calculated for the proposed models using the following equation:

$$R^2 = 1 - \frac{\sum(\hat{y} - y)^2}{\sum(y - \bar{y})^2} \quad (12)$$

Where:

$\bar{y}$: is the mean of the observed response.

- **Mean Absolute Error**

Mean Absolute Error measures the average of the residuals of the prediction results of the proposed models using the following equation:

$$MAE = |\overline{\hat{y} - y}| \quad (13)$$

After the computation of the prognosis performance metrics, different model hyper parameters were optimized through trial and error, until the best hyper parameters were obtained.

Once the optimization is finished and the model is fully developed; the models are then tested on two new sets that has not been trained on before, to check their performance with new data.

4. EXPERIMENTATION AND RESULTS

4.1. Experimentation

In order to test our proposed models, two kind of experiments were conducted, by testing the previously trained models on new unseen data. Both data sequences had an RUL of 3600 second (timestamp), however the sensor measurements were not the same. The first experiment recorded noisy measurements due the reaction between the closing mechanism of the DHSV and the high reservoir upward pressure. The second experiment was recorded from different well conditions, the sensor measurements showed cleaner behavior of the DHSV closing mechanism, due to the stabilized reservoir upward pressure.

The degradation process lasting 3600 second (1 hour) is quite fast, and the reaction time will be small in order to do the maintenance; however, it was the maximum length provided by the dataset creators (Vergas et al., 2019); these experiment will be used to prove that the proposed methods have a significance, future work will be evaluated with real data.

4.1.1. Model Training

The dataset used contains real and simulated data, for this comparative study only the simulated data was considered; transient period from healthy to faulty, only ten files have the full transient period, which are inadequate for a robust model training as pointed in section 2.2 to the unbalance of the recorded instances; therefore, data augmentation is needed, however, it will not be considered in this comparative study. Features discussed in section 4.1.2 were used as predictors and the RUL variable as the response.

4.1.2. Features Used

Four measurements are considered as features in training the regression models, P-PDG, P-TPT, T-TPT, and T-JUS-CKP, due to the physical characteristics of the system; other measurements are either been constant or have negligible values (Figure 5).

To explain the behavior of these measurements we can go back to Figure 2 and imagine that the DHSV is closing bit by bit until it is fully closed; knowing that the oil or gas is coming below from the reservoir, once it faces a closed DHSV, the pressure below the DHSV is increased which explains the increase in P-PDG. On the other side, the pressure and temperature above the closed DHSV is decreased, which explains the drop in P-TPT, T-TPT.

In order to further verify the usefulness of these measurements a correlation matrix was generated from the training data between these measurements (features) and the response. Figure 5 elucidate that no correlation above 0.73 exist between the features used.

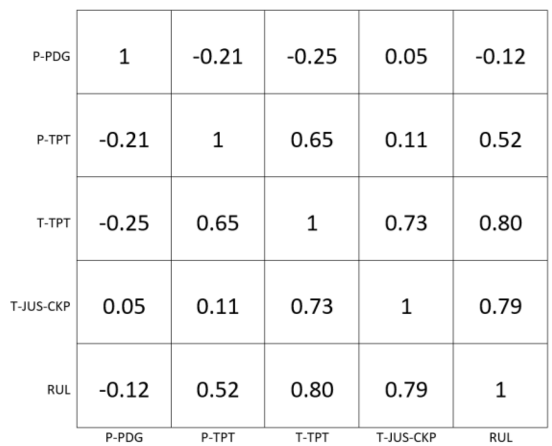

Figure 7: Correlation matrix for feature selection

4.2. Comparative Study Results

In this section results obtained during the validation process and during the testing phase are presented. Validation results were computed using a 5-Fold cross validation, prognosis performance metrics were calculated for each one of the regression models. *RMSE* for all the models indicated a promising training session, due to the fact that the total training sample size was around 50,000 timestamp. Interaction Linear Regression model performed better than the Linear Regression model by almost 10%, Stepwise Regression model on the other hand, showed the most promising results among the trained models, it performed better than the Linear Regression model by 33%, and by 26% better than the second model. Table 3 depicts the results obtained for all the trained models in details.

Table 3: Model training validation results for the three regression models

Performance Metrics	Linear Regression	Linear Interaction Regression	Stepwise Linear Regression
RMSE	425	384	**286**
$\boldsymbol{R^2}$	0.83	0.86	**0.92**
MSE	$1.8 * 10^5$	$1.5 * 10^5$	**81550**
MAE	326	283	**212**

In order to test our proposed models on new unseen data, two kind of experiments were conducted (Section 4.1).

The obtained results show that the stepwise regression model performs better than the linear regression model and the interaction linear regression model in terms of *RMSE* and R-squared values reaching up to 92% accuracy and 0.91 coefficient of determination. which is due to selection mechanism of the model. The first experiment results are shown in Table 4.

Table 4: First experiment results of the three linear regression models

Performance Metrics	Linear Regression	Linear Interaction Regression	Stepwise Linear Regression
RMSE	398	380	**304**
$\boldsymbol{R^2}$	0.85	0.87	**0.91**
MSE	$1.6 * 10^5$	$1.4 * 10^5$	**92695**
MAE	304	305	**230**

Performance pattern was the same for the second experiment, however there was no huge difference numbers wise, Linear Regression model, Interactions Linear Regression, and Stepwise Regression model, got a coefficient of determination of 0.93, 0.96, and 0.97 respectively. Very low *RMSE* values were recorded reaching 172 for the Stepwise Regression model. The obtained results from the second experiment were expected to be better than the first one, due to the clean sensor measurement. Table 5 depicts the second experiment results in detail.

Table 5: Second experiment results of the three linear regression models

Performance Metrics	Linear Regression	Linear Interaction Regression	Stepwise Linear Regression
RMSE	278	215	**172**
$\boldsymbol{R^2}$	0.93	0.96	**0.97**
MSE	77800	46368	**29802**
MAE	221	178	**131**

Figure 6 and Figure 7 give a visual interpretation of the results obtained in the first experiment and second experiment respectively. One my notice that the best performing model is the Stepwise Linear Regression, followed by the Linear Interaction Regression model, then the traditional Linear model. The difference between the models is not huge, however it may affect the results if the RUL is longer, or the sensor measurements are noisier. One may also notice that the results in the second experiment were better than the first one, due to the clean sensor measurement

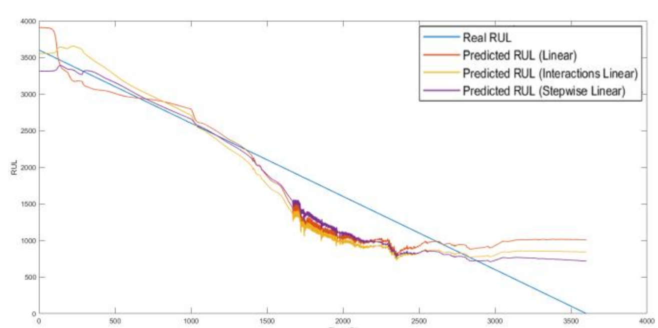

Figure 6: First test results plot showing the predicted RUL using different regression models and the real RUL

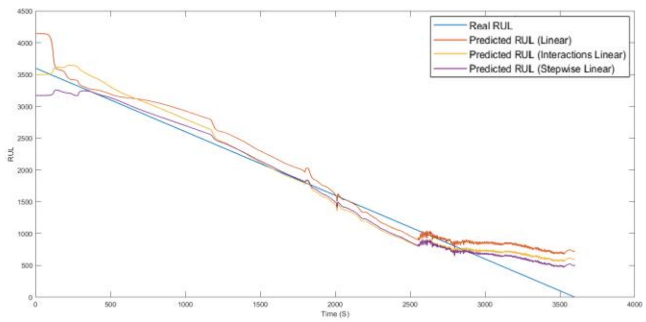

Figure 7: Second test results plot showing the predicted RUL using different regression models and the real RUL

Figure 8 shows the predicted RUL using the Stepwise Linear Regression and the upper and lower bounds of a 85% confidence interval.

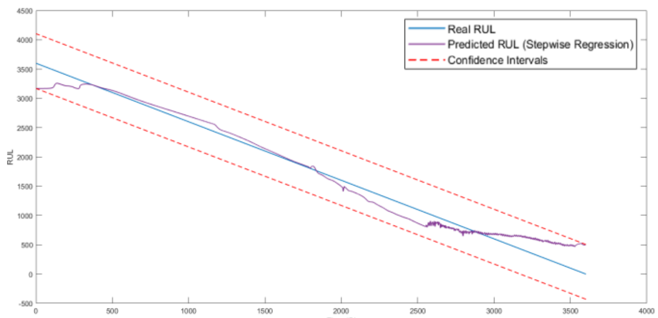

Figure 8: Predicted RUL with Stepwise regression within 85% confidence interval.

4.3. Feedback Loop Mechanism Results

In this section, the results of the proposed feedback loop mechanism are presented. After the selection of the best model from the comparative study; that model is iterated through the feedback loop mechanism. Figure 9 shows the results of the original prediction, as well as the predicted RUL of the first and third iteration.

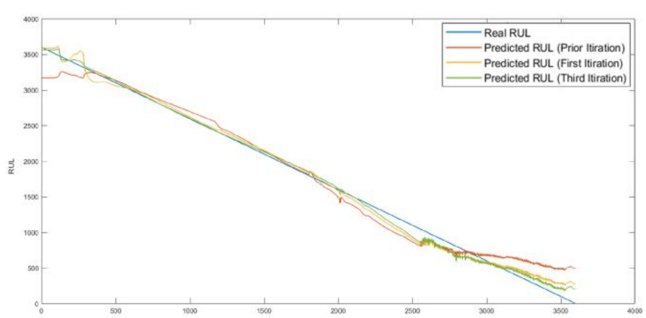

Figure 9: Feedback Mechanism results of the first three iterations

One can clearly notice the positive effect of the proposed mechanism, where the green plot of the predicted RUL (third iteration) is more accurate than the plot prior iterations. The improvements happened at the beginning of the prediction and across the full timeframe.

From Table 6 one can notice that the training validation *RMSE* was decreasing after each iteration, on the other hand the confidence level kept increasing reaching 95% in the third iteration, which proves the effectiveness of the proposed dynamical approach in terms of accuracy. (Figure 10).

Table 6: Validation *RMSE* and confidence interval for each iteration

Iterations	*RMSE* (validation)	Confidence interval
0	268	85%
1	113	92%
2	80	93%
3	**71**	**95%**

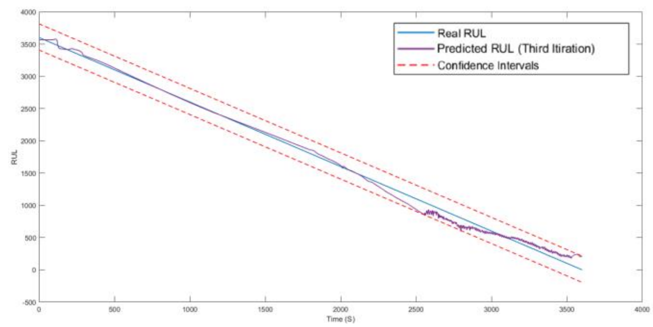

Figure 10: Predicted RUL with Stepwise regression within 95% confidence interval after the third iteration.

Difference between Figure 8 and Figure 10 is clearly visible, where the prediction in Figure 10 is more accurate which depicts the usefulness of the proposed dynamical approach.

5. CONCLUSION AND FUTURE WORK

This paper presented an original performance enhancement method after a comparative study of three linear regression models, and applied them to predict the remaining useful life of the DHSV to reduce the non-productive time that the oil and gas industry faces in case of completion tools failures.

Two tests were performed in order to select the best linear model for our application, and found that the Stepwise Linear Regression model was the best performing model in both tests.

The proposed performance enhancement method is applied to the Stepwise Linear Regression model to optimize its performance. Positive effect was recorded; *RMSE* was significantly reduced, confidence interval was stretched to 95%.

Further studies will aim to improve the proposed feedback mechanism in terms of performance and reliability, as well as speed of convergence. On the other hand, future work will aim to work with real data. In addition, RUL estimation techniques like similarity and degradation models can be implemented in order to give a real time estimate for the RUL.

ACKNOWLEDGEMENT

The authors would like to thank the Petrobras team from the Unidade de Negócios-Espírito Santo for making the 3 W dataset publicly available.

REFERENCES

Banjevic, D. & Jardine, A.K.S. (2005), *Calculation of reliability function and remaining useful life for a Markov failure time process*. IMA Journal of Management Mathematics. doi:10.1093/imaman/dpi029.

Brown, K. E. (1984). *The Technology of Artificial Lift Methods*. United States: PPC Books.

Chinnam, R.B. & Baruah, P. (2003). *Autonomous diagnostics and prognostics through competitive learning driven HMM-based clustering.* Proceedings of the International Joint Conference on Neural Networks, Vols. 1-4, New York, pp. 2466-2471.

Derbal, M. & Houari, T. (2018). *Early Fault Diagnosis in Exciting Capacitors of Self-Excited Induction Generator for Wind Energy Applications*. International Conference on Communications and Electrical Engineering (ICCEE), pp. 1-5, doi: 10.1109/CCEE.2018.8634495.

Emhanna, S. (2018). *Analysis of Non-Productive Time (NPT) in Drilling Operations- A Case Study of the Ghadames Basin*. Second Scientific Conference of Oil and Gas Feb 5-7 / 2018 Ajdabiya – Libya.

Fahrmeir, L., & Kneib, T., & Lang, S. (2009). *Lineare Regressionsmodelle. In: Regression*. Statistik und ihre Anwendungen. Springer, Berlin, Heidelberg. Doi:10.1007/978-3-642-01837-4_3.

Garcia, P., & Jacinto, C. M., & Lima, B.D.S.L. & Pires, & Droguett, E. (2006). *Optimizing downhole safety valve test scheduling using a multiobjective genetic algorithm*. Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management, PSAM 2006.

Hong, T., & Gui, M., & Baran, M. E., & Willis, H. L. (2010). *Modeling and forecasting hourly electric load by multiple linear regression with interactions*. IEEE PES General Meeting, pp. 1-8, doi: 10.1109/PES.2010.5589959.

Houari T., Sayed-Mouchaweh M., Benmiloud M., Defoort M., Djemai M. (2021).Self adaptive learning scheme for early diagnosis of simple and multiple switch faults in multicellular power converters. ISA Transactions, Volume 113, Pages 222-231, https://doi.org/10.1016/j.isatra.2020.03.025.

Krygier, N., & Solarin, A., & Ivanka O. B. (2020). *A Drilling Company's Perspective on Non-Productive Time NPT Due to Well Stability Issues*. Paper presented at the SPE Norway Subsurface Conference, Virtual. doi: 10.2118/200732-MS.

Kwan, C., & Zhang, X., & Xu, R., & Haynes, L. (2003). *A novel approach to fault diagnostics and prognostics.* Proceedings of the IEEE International Conference on Robotics and Automation, Vols. 1-3, New York, pp. 604-609.

Li, Y., & Billington, S., & Zhang, C., & Kurfess, T., & Danyluk, S., & Liang, S. (1999). *Adaptive prognostics for rolling element bearing condition*. Mechanical Systems and Signal Processing, Vol. 13, pp. 103-113.

Li, Y., & Kurfess, T.R., & Liang, S.Y. (2000). *Stochastic prognostics for rolling element bearings*. Mechanical Systems and Signal Processing, Vol. 14, pp. 747-762.

Lin, D. & Makis, V. (2004). *Filters and parameter estimation for a partially observable system subject to random failure with continuous range observations.* Advances in Applied Probability, Vol. 36, pp. 1212-1230.

Ning, H., & Hao, H. Z. (2017). *A Note on High-Dimensional Linear Regression with Interactions*. The American Statistician, 71:4, 291-297, Doi: 10.1080/00031305.2016.1264311.

Phelps, E., & Willett, P., & Kirubarajan, T., (2001), *A statistical approach to prognostics*. Component and Systems
Diagnostics, Prognosis and Health Management, Vol. 4389, pp. 23-34.

Purser, P. E. (1997). *Review of Reliability and Performance of Subsurface Safety Valves*. Paper presented at the Offshore Technology Conference, Houston, Texas. doi:10.4043/2770-MS.

Qiubing, R., Mingchao, L. & Shuai, H. (2019), *Tectonic discrimination of olivine in basalt using data mining techniques based on major elements: a comparative study from multiple perspectives*. Big Earth Data, 3:1, 8-25. doi: 10.1080/20964471.2019.1572452.

Resato International (2020). *Down Hole Safety Valve. Controlling Objects, Oil and Gas Industry*. Duitslandlaan, Netherlands.

Salem, H., Sayed-Mouchaweh, M. (2020). *A Semi-supervised and Online Learning Approach for Non-Intrusive Load Monitoring*. In: Brefeld, U., Fromont, E., Hotho, A., Knobbe, A., Maathuis, M., Robardet, C. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2019. Lecture Notes in Computer Science(), vol 11908. Springer, Cham. https://doi.org/10.1007/978-3-030-46133-1_35

Saxena, A. & Celaya, J. & Balaban, E. & Goebel, K. & Saha, B. & Saha, S. & Schwabacher, M. (2008). *Metrics for evaluating performance of prognostic techniques*. International Conference on Prognostics and Health Management. pp. 1-17. doi: 10.1109/PHM.2008.4711436.

Schneider, A., & Hommel, G., & Blettner, M. (2010). *Linear regression analysis: part 14 of a series on evaluation of* *scientific publications*. Deutsches Arzteblatt international, 107(44), 776–782. Doi:10.3238/arztebl.2010.0776.

Tran, V. & Yang, B. S. (2009). *Machine Fault Diagnosis and Prognosis: The State of the Art*. The International Journal of Fluid Machinery and Systems. 2. 61-71. Doi:10.5293/IJFMS.2009.2.1.061.

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

Vlok, P.J., & Wnek, M., & Zygmunt, M. (2004). *Utilising statistical residual life estimates of bearings to quantify the influence of preventive maintenance actions.* Mechanical Systems and Signal Processing, Vol. 18, pp. 833-847.

Wang, W., & Scarf, P.A., & Smith, M.A.J. (2000). *On the application of a model of condition-based maintenance.* Journal of the Operational Research Society, Vol. 51, pp. 1218-1227.

Wang, W. (2002). *A model to predict the residual life of rolling element bearings given monitored condition information to date*. IMA Journal of Management Mathematics, Vol. 13, pp. 3-16.

Xiao-Sheng, S., & Wenbin W., & Chang-Hua H., & Dong-Hua Z. (2011). *Remaining useful life estimation – A review on the statistical data driven approaches*. European Journal of Operational Research, Volume 213, Issue 1, Pages 1-14. Doi: 10.1016/j.ejor.2010.11.018.

XTO Energy (2019). *Benefits of Natural Gas and Oil.* Online article. Energy and Environment. Unconventional Recourses Development.

Yan, J., Koc, M., and Lee, J. (2004). *A prognostic algorithm for machine performance assessment and its application.* Production Planning and Control, Vol. 15, pp. 796-801.

Zhang, S. & Ganesan, R. (1997). *Multivariable trend analysis using neural networks for intelligent diagnostics of rotating machinery*. Trans. ASME Journal of Engineering for Gas Turbines and Power, Vol. 119, pp. 378-384.