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

Harrouz Aymen1, Toubakh Houari1, and Kafi Mohamed Redouane1, Moamar Sayed-Mouchaweh2, Hajer Salem3

1University of Kasdi Merbah Ouargla, Ouargla, 30000, Algeria. 2Institute Mines-Telecom Lille Douai, Douai, France. 3Audensiel 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).

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 a few of them: Tubing Retrieve Safety Valve (TRSV), Surface Controlled Subsurface Safety Valve (SCSSV), Wireline Retrieve 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).

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), (Banjевич 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.
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](image)

Figure 2: Down Hole Safety Valve Vertical Cut and Components (Garcia, Jacinto, Lima, Pires and Drogutt, 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.

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>Real</th>
<th>Simulated</th>
<th>Sketched</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal</td>
<td>597</td>
<td>0</td>
<td>0</td>
<td>597</td>
</tr>
<tr>
<td>1</td>
<td>Abrupt BSW Increase</td>
<td>5</td>
<td>114</td>
<td>10</td>
<td>129</td>
</tr>
<tr>
<td>2</td>
<td>Spurious DHSV Closure</td>
<td>22</td>
<td>16</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>Severe Slugging</td>
<td>32</td>
<td>74</td>
<td>0</td>
<td>106</td>
</tr>
<tr>
<td>4</td>
<td>Flow Instability</td>
<td>344</td>
<td>0</td>
<td>0</td>
<td>344</td>
</tr>
<tr>
<td>5</td>
<td>Rapid Productivity Loss</td>
<td>12</td>
<td>439</td>
<td>0</td>
<td>451</td>
</tr>
<tr>
<td>6</td>
<td>Quick PCK Restriction</td>
<td>6</td>
<td>215</td>
<td>0</td>
<td>221</td>
</tr>
<tr>
<td>7</td>
<td>Scaling in PCK</td>
<td>4</td>
<td>0</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>Hydrate in Prod. Line</td>
<td>3</td>
<td>81</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1025</td>
<td>939</td>
<td>20</td>
<td>1984</td>
</tr>
</tbody>
</table>

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 3W database, including tag names, descriptions, and measuring units.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-PDG</td>
<td>Pressure at permanent downhole gauge (PDG)</td>
<td>Pa</td>
</tr>
<tr>
<td>P-TPT</td>
<td>Pressure at temperature/pressure transducer (TPT)</td>
<td>Pa</td>
</tr>
<tr>
<td>T-TPT</td>
<td>Temperature at temperature/pressure transducer (TPT)</td>
<td>°C</td>
</tr>
<tr>
<td>P-MON-CKP</td>
<td>Pressure upstream of production choke (CKP)</td>
<td>Pa</td>
</tr>
<tr>
<td>T-JUS-CKP</td>
<td>Temperature downstream of production choke (CKP)</td>
<td>°C</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td>Unit</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>P-JUS-CKGL</td>
<td>Pressure downstream of gas lift choke (CKGL)</td>
<td>Pa</td>
</tr>
<tr>
<td>T-JUS-CKGL</td>
<td>Temperature downstream of gas lift choke (CKGL)</td>
<td>°C</td>
</tr>
<tr>
<td>QGL</td>
<td>Gas lift flow rate</td>
<td>m³/s</td>
</tr>
</tbody>
</table>

### 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:

\[ Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n \]  (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:

\[ RUL = a + b_{1PDG} + b_{2TPT} + b_{3TPT} + b_{4JUSCKP} \]  (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:

\[ Y = a + b_1X_1 + b_2X_2 \]  (3)

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

\[ Y = a + b_1X_1 + b_2X_2 + b_{12}X_1X_2 \]  (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:

\[ \text{Number of terms} = 2^{\text{interactions order}} \]  (5)

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

\[ RUL = a + b_{1PPT} + b_{2TPT} + b_{3PPTTPT} \]  (6)

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

\[ \text{Number of terms} = 2^4 = 16 \]  (7)
3.4. System Overview

3.4.1. Comparative Study System

The proposed system is represented in Figure 3. The system flow starts 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:
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.

\[
Y_1 = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + 0 \\
Y_2 = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \hat{Y}_1 \\
Y_3 = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \hat{Y}_2 \\
\vdots \\
Y_n = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \hat{Y}_{n-1}
\]

Or as

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

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 \( RMSE \):

\[
RMSE = \sqrt{\frac{1}{n} \sum (\hat{y} - y)^2}
\]

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).

\[ MSE = (\bar{y} - y)^2 \]  \hspace{1cm} (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:

\[ R^2 = 1 - \frac{\sum(y - \bar{y})^2}{\sum(y - \bar{y})^2} \]  \hspace{1cm} (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:

\[ MAE = |\bar{y} - y| \]  \hspace{1cm} (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 to 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.

<table>
<thead>
<tr>
<th>Feature</th>
<th>P-PDG</th>
<th>P-TPT</th>
<th>T-TPT</th>
<th>T-JUS-CRP</th>
<th>RUL</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-PDG</td>
<td>1</td>
<td>-0.21</td>
<td>-0.25</td>
<td>0.05</td>
<td>-0.12</td>
</tr>
<tr>
<td>P-TPT</td>
<td>-0.21</td>
<td>1</td>
<td>0.65</td>
<td>0.11</td>
<td>0.52</td>
</tr>
<tr>
<td>T-TPT</td>
<td>-0.25</td>
<td>0.65</td>
<td>1</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>T-JUS-CRP</td>
<td>0.05</td>
<td>0.11</td>
<td>0.73</td>
<td>1</td>
<td>0.79</td>
</tr>
<tr>
<td>RUL</td>
<td>-0.12</td>
<td>0.52</td>
<td>0.80</td>
<td>0.79</td>
<td>1</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Performance Metrics</th>
<th>Linear Regression</th>
<th>Linear Interaction Regression</th>
<th>Stepwise Linear Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>425</td>
<td>384</td>
<td>286</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.83</td>
<td>0.86</td>
<td>0.92</td>
</tr>
<tr>
<td>MSE</td>
<td>$1.8 \times 10^5$</td>
<td>$1.5 \times 10^5$</td>
<td>81550</td>
</tr>
<tr>
<td>MAE</td>
<td>326</td>
<td>283</td>
<td>212</td>
</tr>
</tbody>
</table>

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

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>
<thead>
<tr>
<th>Performance Metrics</th>
<th>Linear Regression</th>
<th>Linear Interaction Regression</th>
<th>Stepwise Linear Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>398</td>
<td>380</td>
<td>304</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.85</td>
<td>0.87</td>
<td>0.91</td>
</tr>
<tr>
<td>MSE</td>
<td>$1.6 \times 10^5$</td>
<td>$1.4 \times 10^5$</td>
<td>92695</td>
</tr>
<tr>
<td>MAE</td>
<td>304</td>
<td>305</td>
<td>230</td>
</tr>
</tbody>
</table>

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

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>
<thead>
<tr>
<th>Performance Metrics</th>
<th>Linear Regression</th>
<th>Linear Interaction Regression</th>
<th>Stepwise Linear Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>278</td>
<td>215</td>
<td>172</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.93</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>MSE</td>
<td>77800</td>
<td>46368</td>
<td>29802</td>
</tr>
<tr>
<td>MAE</td>
<td>221</td>
<td>178</td>
<td>131</td>
</tr>
</tbody>
</table>

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

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 noiser. One may also notice that the results in the second experiment were better than the first one, due to the clean sensor measurement.
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.

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>
<thead>
<tr>
<th>Iterations</th>
<th>RMSE (validation)</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>268</td>
<td>85%</td>
</tr>
<tr>
<td>1</td>
<td>113</td>
<td>92%</td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>93%</td>
</tr>
<tr>
<td>3</td>
<td>71</td>
<td>95%</td>
</tr>
</tbody>
</table>

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


Schneider, A., & Hommel, G., & Blettner, M. (2010). Linear regression analysis: part 14 of a series on evaluation of


