Combined principal component analysis and proportional hazard model for optimizing condition-based maintenance

T Bankole-Oye1, *, I El-Thalji1 and J Zec2

1Department of Mechanical and Structural Engineering and Materials Science, University of Stavanger, Norway
2Maersk Drilling AS, 4604, Kristiansand, Norway

* Contact author: tosinbankoleoye@gmail.com

Abstract: Large companies are investing heavily in digitalization to be more competitive and economically viable. Hence, physical assets and maintenance operations have been digitally transformed to transmit a high volume of data, e.g., condition monitoring data. Such high-volume data can be useful to optimize maintenance operations and minimize maintenance and replacement costs. A tool to optimize maintenance using condition monitoring data is the Proportional hazard model (PHM). However, it is challenging to implement PHM for industrial complex systems that generate big data. Therefore, machine learning algorithms shall support PHM method to handle such a high volume of data. Thus, the purpose of this paper is to explore how to support PHM with Principal Component Analysis (PCA) to maintenance optimization of complex industrial systems. A case study of hydraulic power unit was purposefully selected to apply and validate the proposed analytical approach. The results show that PCA supported PHM optimizes and extends the preventive maintenance interval by 79.27% which might lead to maintenance cost reductions. This model enables PHM to handle complex systems where big data is collected.

1. Introduction

The Oil and Gas industry is facing significant problems such as low oil prices, stricter emission targets and recently the health crisis which has crippled global energy demand. New fields are being installed in offshore locations at greater depths while seeking cost-effective alternatives to extend the life of old fields. There are challenges regarding optimizing productivity, improving the life cycle of equipment and real-time monitoring of assets [1]. Companies are eagerly seeking to adopt maintenance as a revenue-generating business component to increase profit margins [2]. The need for maintenance was further collaborated by [3] who reported that 27% of all offshore injuries are associated with maintenance and about 10% of hydrocarbon emissions are associated with maintenance activities. Despite large investments in maintenance, equipment failure remains the principal cause of production shortfalls [4]. DuPont [5] had stated that maintenance is the single most influential expenditure in the plant with maintenance budget responsible for 67% of the total yearly profit. This led to a consensus in adopting new monitoring techniques which can reduce downtime and maintenance costs. Embracing asset management will propel a transition from reactive and scheduled maintenance to condition-based maintenance (CBM). For systems subject to CBM, maintenance tasks are done usually when required
based on observed conditions. CBM is carried out on the premise that components suffer a degradation process before failure which can be observed by process indicators such as temperature, pressure, oil analysis and vibration.

The system failure occurs when the degradation state exceeds a pre-determined failure threshold [6, 7]. However, actual system performance has shown that it is challenging to determine failure threshold as factors such as environmental conditions and product characteristics influence the degradation process which increases the likelihood of system failure but should not lead directly to failure. Both cumulative degradation process and internal aging mechanism enhance failure [8]. Industrial assets exhibit a joint effect of internal wear and environmental conditions [9-11]. A suitable and standardized method of incorporating aging and degradation effect into system failure is by a proportional hazards model [12]. Proportional hazard model (PHM) integrates a baseline hazard rate that accounts for the aging effect and environmental factors known as covariates such as temperature, vibration, running speed and pressure. PHM has been successfully employed in various disciplines including medicine, manufacturing plants, finance and energy generators [13]. Numerous researches have been conducted in optimizing maintenance using PHM. A control-limit optimization policy for assets subject to scheduled inspection routine [14]. PHM utilized for systems with imperfect information and the state of equipment is unknown [15]. Continuous-time Markov chain is used to describe the state of the system while investigating the value of perfect monitoring using PHM [16]. A CBM policy was adopted for multi-component systems using PHM [17]. Decision policy suggested for inspection scheduling using PHM [18]. An optimization model was proposed for a multi-unit system using PHM [19]. PHM was also proposed to develop predictive maintenance of industrial systems [20].

However, it is challenging to implement PHM for industrial complex systems that generate big data, i.e., high-frequency data, with high resolution from multiple sources with multiple formats (numeric, text, images, curves). Therefore, the purpose of this paper is to explore how to support PHM with Principal Component Analysis (PCA) to enhance maintenance optimization of complex industrial systems. A case study of hydraulic power unit (HPU) was purposefully selected to apply and validate the proposed analytical approach (PCA & PHM).

This research aims to propose an innovative and integrated approach to optimize the maintenance operation for drilling HPU based on big amount of condition monitoring data. The goal is to reduce maintenance costs while increasing the availability level. Different covariates have been examined to identify the most significant covariates. Oil analysis results are considered as a covariate because they are environmental factors and affect degradation mechanisms. Moreover, Principal Component Analysis (PCA) has been employed to extract the most important information from the multivariable dataset and address “the curse of dimensionality” associated with a large dataset as reported in [21]. The Weibull distribution function is used to model the baseline hazard rate. The model provides decision support to planning preventive maintenance tasks.

The paper is organized as follows. The applied methodology of eight steps and associated methods are discussed, together with the case study in section 2. In section 3, the results of the following methods are presented and discussed: system analysis, Pareto chart, PCA, Cox modeling, Schoenfeld test, Weibull analysis, Survival rate estimation and maintenance optimization. In section 4, the conclusions regarding the proposed methodology and the obtained results are provided.

2. The analytical approach and theories
In this section, the analytical approach and associated theories will be explained. Moreover, the case study of hydraulic power unit in offshore drilling rig will be presented. The analytical approach in this paper consists of eight connected analysis steps, as shown in Figure 1. This approach starts with systems analysis to understand the stakeholders’ needs, system boundary, system functions, technical hierarchy, and failure modes and explore and evaluate potential conceptual solutions. The conceptual solution
evaluation helps us to select the best maintenance optimization model, i.e., PHM, Markov-chain. The second step is to analyze historical maintenance data to determine the most critical components and recognize failure patterns or trends. The third step is important to select the most relevant parameters for optimizing the condition monitoring and maintenance operations. This step can be done by several feature selection or dimension reduction algorithms, however, PCA is selected and applied in this paper. It is important to reduce the data dimensions without losing the important data. In the fourth step, the proportional hazard analysis using Cox modeling is proposed to determine what is the impact of each covariate (parameter or dimension) on operational performance. It deals with limited parameters, therefore, PCA is required. Before the most relevant covariate(s) are selected, a proportionality test is required to avoid any violated results. Thus, the fifth step is the proportionality test, where Schoenfeld residuals are applied. Once the proportionality is valid, the sixth step will take place, where the most relevant covariates can be selected using Akaike Information Criterion (AIC). In the seventh step, the proportional hazard can be modeled with help of Weibull distribution. Weibull distribution is frequently used to model the failure times of mechanical systems and is the most popular distribution in PHM for modeling baseline survival rate. Finally, the maintenance operations can be optimized based on the PHM in the eight-step.

In summary, these eight steps involve several analytical methods. System analysis, Failure Mode Effect Criticality Analysis (FMECA), Pareto chart, PCA, Cox modeling, Schoenfeld test, Weibull analysis, Survival rate estimation and maintenance optimization. In the following paragraphs, these methods are defined and explained.

The objectives of PCA [22] are to (1) Extract the most important information from the data matrix, (2) Compress the size of the matrix and retain the most important data, (3) Simplify the description of the data matrix, (4) Investigate the structure of the variables.
For a multivariate $\mathbf{X}_n = (x_1, x_2, \ldots, x_p)^T$ where $x_1, \ldots, x_p$ are $p$ variables. It is impractical to use a variable to represent the original multivariate. Moreover, reducing the dimension of the multivariate will lose the underlying information of the data structure and will not describe the nature of the multivariate. Standardized linear combination (SLC) is a flexible method to reduce the dimension and give weights to the variables of the multivariate as seen in equation (1).

$$\delta^T \mathbf{X} = \sum_{j=1}^{p} \delta_j x_j \text{ and } \sum_{j=1}^{p} \delta_j^2 = 1 \quad (1)$$

The direction of the unit vector $\delta$ is provided by the eigenvector $\gamma_1$ which connotes the largest eigenvalue $\lambda_1$ of the covariance matrix $\Sigma = \text{Var} (\mathbf{X})$. The SLC with the largest variance is the first principal component (PC): $y_1 = \gamma_1^T \mathbf{X}$. The SLC with the second-largest variance is the second PC $y_2 = \gamma_2^T \mathbf{X}$ (2)

The eigenvalue of a component is the sum of the squared factor scores for this component. The contribution of the observation is calculated as the ratio of the square factor score by the eigenvalue. The contribution $\text{ctr}_{i, \ell}$ is expressed as

$$\text{ctr}_{i, \ell} = \frac{f_{i, \ell}^2}{\sum_i f_{i, \ell}^2} = \frac{f_{i, \ell}^2}{\lambda_\ell} \quad (3)$$

where: $\lambda_\ell$ is the eigenvalue of the $\ell$-th component.

The value of a contribution is between 0 and 1 and the addition of all contributions should be equal to 1. The bigger the value of the contribution, the more it contributes to the component.

PCA has been used to improve reliability and maintenance studies. It is useful in maintenance optimization to deal with covariates as they grow exponentially which creates a problem of dimensionality. PCA is used to reduce dimension and eliminate collinearity among covariate datasets. [23] utilized PCA to oil analysis data and vibration data. PCA was used in oil analysis to reduce covariates from 20 to 6 metal elements while in the vibration analysis, 49 covariates were reduced to one. The PCs obtained were used to develop an optimal maintenance policy. PCA analysis was applied to an oil dataset to obtain the principal components capturing the most information of the oil data. Seven-dimensional oil monitoring data was reduced to four PCs and used to determine the remaining useful life oil of the lubricating oil[24].

Cox modeling of each selected covariate aims to determine their impact on operational performance. Hypothesis tests will be carried out to determine the level of significance of the covariates. When the null hypothesis is greater than the threshold, the result will be discarded [25]. Then, the proportionality test will be performed to determine whether the proportionality assumption is satisfied. This assumption states that the hazard ratio (HR) for two observations should be constant with time. The goodness of fit (GOF) tests are used for checking proportionality assumptions. They are useful because they provide a test statistic and p-value for model checking which gives a more robust conclusion. Common GOF tests include Schoenfeld residuals, martingale residuals and Cox-Snell residual.

The assumption is obtained from

$$HR = \frac{h_1(t \mid P_1)}{h_2(t \mid P_2)} = \frac{h_0(t) \exp(\beta^T P_1)}{h_0(t) \exp(\beta^T P_2)} = \exp[\beta^T (P_1 - P_2)] \quad (4)$$

where,

- $P_1$ and $P_2$ are covariates for two observations while the term
- $HR = \exp [\beta T (P_1 - P_2)]$ gives the hazard ratio and gets a constant value over time.
Models that violate the proportional hazard assumption can be modified either by using the stratification approach or time-dependent covariate (extended cox model) [26]. The covariate selection process aims to determine the most suitable covariates for the optimization model. Fitting a model with too many covariates may lead to contributing predictors with little or no significance. There are different approaches for model selection, including the Stepwise method, Akaike Information Criterion (AIC) [27] and Bayes Information Criteria [28].

The baseline hazard modeling can be performed, through parametric and non-parametric models, to determine the survival rate. Non-parametric methods can be used to describe the features of the distribution of time without specifying the functional form of the baseline hazard rate. The coefficients of regression analysis are obtained by maximizing the partial likelihood function. The most common non-parametric method is the product-limit estimator known as Kaplan-Meier (KM) estimator [29]. The parametric model consists of specific distributions, including exponential, Weibull, normal and lognormal distributions. The Weibull distribution is frequently used to model the failure times of mechanical systems and is the most popular distribution in PHM for modeling baseline survival rate [30]. The general form for the proportional hazard model is:

\[
    h(t \mid P) = h_0(t) \exp(\beta^T P) \tag{5}
\]

where,

- \( h_0 \) is the baseline hazard function,
- \( P \) is a vector of covariates while
- \( \beta^T \) is the vector of regression coefficients that describes the effects of the covariates on the hazard function.

The hazard rate for the two-parameter Weibull distribution is written as

\[
    h_0(t) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta - 1} \tag{6}
\]

The form of the Weibull Proportional hazard model (WPHM) is

\[
    h(t, z) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta - 1} \exp(\beta^T P) \tag{7}
\]

where: \( \beta > 0 \) and \( \eta > 0 \) are the shape and scale parameters of the Weibull distribution. The parameters are obtained by Maximum Log-Likelihood estimation.

The optimization of maintenance is performed to ensure that preventive maintenance is carried only when required. PHM has been successfully utilized to optimize maintenance tasks using principal component analysis [31]. An illustrative case study was the oil analysis data set obtained from a mining company. The optimal maintenance for bearings in the food industry can be modeled using PHM [32]. Weibull proportional hazard model was proposed to model the hazard rate of systems subjected to failure. Different criteria can be considered for optimizing maintenance including cost, availability, safety and profit [33]. According to [34] lowest-cost optimization is often the maintenance goal. A procedure is to utilize the cost-time equation as seen in equation 8

\[
    CT(t) = \frac{C_P R(t) + C_U [1 - R(t)]}{\int_0^t R(s) ds} \tag{8}
\]

where,

- \( CT \) is the cost per time
- \( C_P \) is the cost for preventive maintenance
- \( C_U \) is the cost for corrective maintenance.
- \( R(t) \) is a reliability function that can be obtained from the Weibull distribution. The optimal maintenance time is the time that minimizes \( CT(t) \).
2.1 Offshore drilling rigs
Drilling rigs are usually equipped with several hydraulic power unit (HPU) which provides hydraulic pressure to offshore drilling systems. The HPU is designed with high reliability and redundancy to ensure continuous operation in case of failure in critical parts of the system. Moreover, the HPU has low workloads except when high pressure is required to the offshore drilling system. However, they are subjected to a high level of inspections and calendar-based preventive maintenance. The cost spent on maintenance including inspections and replacements tasks is deemed too high and there is the possibility that the system is overly maintained. The interest is in using data-driven CBM that will incorporate the age of the HPU and monitoring variables in describing an optimum maintenance policy.

Maintenance records for drilling HPU had been collected from an industrial case company. A summary of the maintenance work order was collected from July 2017 till October 2020. The work orders were carefully analyzed to determine the number of preventive and corrective maintenance tasks taken. Oil analysis results were also collected between the same period. A typical oil test result includes additive, elemental, wear and contaminant analysis of an oil sample reflecting the level of degradation. Running hours of the pumps were obtained to determine the working age of the system at failure and maintenance. The current maintenance plan was critically studied to determine the need for optimization.

3. Results and discussions
In this section, the results of the eight analysis steps are presented and discussed in the following order: Systems analysis, Failure Mode Effect Critical Analysis (FMECA), Pareto chart, PCA, Cox modeling, Schoenfeld test, Weibull analysis, Survival rate estimation, maintenance optimization

3.1 System Analysis
An initial system analysis was performed. This analysis began with the system design of the HPU which includes the development of the system functionalities and technical hierarchy. The outcome was used in the functional design phase. Failure analysis was performed using Failure Mode Effect Criticality Analysis (FMECA). In the first instance, a hierarchical structure of the system was established to understand the complexity within the system. The drilling HPU is connected to other systems such as top drive, Blowout preventer (BOP), drawworks and cranes. It also consists of sub-components such as pumps, filters, tank, instrumentation and values. FMECA is performed to eliminate operation breakdown leading to downtimes and poor service delivery in firms at the failure mode level. Identification of the failure modes at the planning stages of conceptual design enables asset managers to invest in predictive monitoring through condition-based maintenance to improve the efficiency of the systems. The HPU system has inherent failure modes which require proper problem identification and mitigation mechanism to eliminate instances of malfunctioning in the operation. The pumps are the most critical component and have also be divided into the component level. There are different sources of failure on the HPU system including oil leakage, imbalance and misalignment. Oil leakage through the pump has the highest Risk Priority Number (RPN). Important maintenance tasks were also developed to prevent these failures.
3.2. Maintenance data analysis

The lifetime data was collected from the maintenance history register. The register has entries of all corrective and preventive maintenance tasks. The register contained information such as failure date and time, description of the failure mode, failure phenomenon and repair time. Maintenance data collected contains time to event data where the event is the failure of the HPU pump. However, data is also collected when events (failure) did not occur either through replacement or maintenance, which is known as censored data.

![Figure 2. Snapshot of the FMECA for drilling HPU](image2)

### Table 1. Failure Mode Effects and Criticality Analysis (FMECA)

| Component | Function | Potential Failure Mode | Potential Failure Cause | Consequence | Freq | Det | Sev | RPN | Corrective actions |
|-----------|----------|------------------------|-------------------------|-------------|------|-----|-----|-----|-------------------|
| Pump      | Generates the required pressure | Oil leakage | High temperature | Loss of pressure | 3 | 4 | 2 | 24 | Oil analysis program |
|           |          | Cavitation | Low pressure | Overheating | 2 | 3 | 2 | 12 | Monitor pump pressure |
|           |          | Distorted piston | Wear | Reduced pressure | 2 | 2 | 3 | 12 | Vibration monitoring |
| Bearing   | Reduce friction and support directional load | Leakage | Wear | Reduced pressure | 3 | 4 | 4 | 12 | Oil analysis program |
|           |          | Structural imbalance | Wear | Structural imbalance | 3 | 4 | 3 | 18 | Vibration monitoring |
| Seal      | To airtight the system and avoid loss of pressure | Leakage | Wear | Leaking | 3 | 3 | 2 | 12 | Vibration monitoring |
|           |          | Deficiency | Wear | Failure | 2 | 4 | 2 | 10 | Improved design |
|           |          | Failure | Friction wear | Failure | 2 | 3 | 2 | 12 | Vibration monitoring |
| Motor     | Generates mechanical power to do work | Imbalance | Vibration | Leakage | 2 | 3 | 3 | 18 | Scheduled Inspections |
|           |          | Misalignment | Vibration | Leakage | 2 | 3 | 2 | 8 | Oil analysis program |
|           |          | Bent shaft | High temperature | Bent shaft | 2 | 3 | 3 | 6 | Preventive maintenance |

**Figure 3. Historical Maintenance Records**

**Figure 4. Pareto chart for drilling HPU**

The lifetime data were divided into sub-components of the drilling HPU based on the technical hierarchy previously developed. Every failure of the HPU system is categorized into pumps, filters, sensors, circulation filters, valves, hoses and motors based on the main components of the system. The pump is the most critical component of the HPU based on the failure frequency data. Running hours were collected for the main pumps for one year of operation. The functionalities of the pumps were
controlled by the control logic to ensure even operation and wear. As seen from Figure 5, pumps 1-3 have worked similar hours with little variations in pumps 4 and 5.

Trend analysis is critical to determine the parametric form for fitting the model. A cumulative failure plot against time has been used to determine the existence of a trend as shown in Figure 6. The graph produced a straight line which means the data is identically distributed and free of trends. The absence of trends in the data enables parametric models to be deployed in modeling baseline hazards.

3.3 Principal component analysis

Oil analysis is the selected covariate that is deemed to have the highest environmental impact on the system based on the risk map. The oil analysis report dataset obtained for this project contains 30 variables. Out of these, 21 variables are categorical and have a predefined threshold value such as <1, <10. The remaining are numerical variables that will be considered for PCA analysis. The variables included in the analysis are viscosity, acid number, 4µm, 6µm, 14µm, calcium, phosphorous, zinc and sulphur.

PCA analysis was performed using R Software. A scree plot is used to determine the number of principal components (pcs) by drawing the eigenvalues from largest to smallest as seen in Figure 7. From the plot, it is acceptable to stop at the fourth pcs. This is because 86% of the information contained in the data set is retained by the first four pcs.

The contributions of variables in determining the variables in a pc can be expressed in bar plots. Variables most correlated with each pc best explain the variability in the data set. A variable with a larger contribution than a set threshold (10%) is considered to contribute to the component. The contribution of the most important variables is shown in Figure 8. The red dashed line on the graph indicates the expected average contribution. The variables obtained include 4µm, 6µm, 14µm, viscosity, zinc, total acid number. These variables will be used to develop the cox model.
3.4. Cox modelling
PHM was developed based on the selected covariates from PCA. Analysis was performed to illustrate and describe the effect of each covariate. Hypothesis testing is used to assess whether the covariates have a significant effect on the failures. The multivariable analysis depicted in Figure 11 shows that viscosity is the most significant covariate.

3.5 Test for proportionality
It is important to assess how our data fits the proportional hazard assumption. Schoenfeld residuals are a common statistical method used to check this assumption. It is independent of time and a plot that shows a linear relationship with time is proof of the violation of proportionality of hazard. The plots of residuals for each of the covariates are plotted against time. The solid line is a smoothing spline fit to the plot. The dashed lines represent a +/-2-standard error band around the plot which represents a 95% confidence interval. Figure 10 illustrates that the residuals for the covariates do not have any definite pattern with time suggesting that there is no violation of the proportional hazard.

3.6. Model selection and analysis
An approach for selecting the most important covariates is to use the stepwise approach (backward elimination) for model selection and Akaike Information Criterion (AIC) for hypothesis testing. Table 1 shows the result of the analysis and can be observed that viscosity and size are the most important covariate.

| Covariates | AIC     | Comment |
|------------|---------|---------|
| -Zinc      | 115.98  | Discard |
| -TAN       | 116.25  | Discard |
| -Size      | 116.62  | Keep    |
| Viscosity  | 118.96  | Keep    |

3.7. Baseline hazard modelling
The baseline hazard is modeled using parametric baseline distribution. The Weibull distribution is the most popular distribution modeling baseline survival because it can model the failure times of mechanical systems [30].

The data used for fitting the Weibull distribution is the time to fail (TTF) event histories of the pumps which include both censored and uncensored data. The coefficients of the Weibull distribution are obtained by Maximum Log-Likelihood estimation which is
\[ L = \prod_{i=1}^{n} \left( \frac{\beta}{\eta} \frac{t_i}{\eta} \right)^{\beta - 1} \exp \left[ - \left( \frac{t_i}{\eta} \right)^{\beta} \right] = \left( \frac{\beta}{\eta} \right)^{n} \prod_{i=1}^{n} \left( \frac{t_i}{\eta} \right)^{\beta - 1} \exp \left[ - \left( \frac{t_i}{\eta} \right)^{\beta} \right] \]  

(9)

Table 2. Coefficients of weibull hazard modelling

| Parameters  | Coefficients | Significance test (p-value) |
|-------------|--------------|-----------------------------|
| Viscosity   | 0.707        | 0.0549                      |
| Size        | 1.000        | 0.2751                      |
| Shape (\(\beta\)) | 2.606       | 0.0000                      |
| Scale (\(\eta\))  | 807.65      | 0.0000                      |

Viscosity is seen to have a p-value close to 0.05 which makes it statistically relevant. Size is not relevant due to its high p-value. The shape (\(\beta\)) has a value greater than one which signifies that the hazard rate increases with time. Viscosity is the most important covariate and will be used in the final model development. The complete equation becomes

\[
h(t, z) = \frac{2.606}{807.65} \cdot 1.606 \cdot e^{0.346x}
\]

where \(x\) is the viscosity covariate and \(t\) is the time that can be used for hazard prediction which is usually the time at inspection.

The reliability and hazard rate values were predicted for the developed Weibull hazard model. The reliability function is the probability of surviving to a time interval. Figure 11(left) shows the probability that the HPU pump will survive more than 850 hours is

\[ R(t = 1) = P[T > 1] = 0.5. \]

This means that 50% of the pump will survive past 850 hours and the reliability of the HPU at 850 hours is approximately 0.5

Figure 11. (left) Reliability function and (right) Hazard rate for HPU

The hazard function is the probability that the equipment will fail in \((t, t + \delta t]\) for small \(\delta t\) given it had not failed before \(t\). Figure 11 (right) demonstrates the hazard rate for the PHM. The hazard rate is shown to be monotonically increasing during the study period. The hazard function plotted in this report describes the failure rate of the HPU pump based on the age and the rate of degradation of the oil analysis covariates. Hazard rate consists of the baseline (Weibull) hazard and an exponential function which is the effect of viscosity (monitored variable). The result shows a rapid increase in the hazard rate of the equipment. Specifically, the hazard rate rose from 0 to 3 between 0 to 400 hours. However, the rate of increase in hazard rate slowed from 400 till 1,000 with only a unit increase in hazard rate.

The parameters of the Weibull distribution are used to develop the full Weibull proportional hazard model. The baseline hazard represented by the Weibull distribution illustrates an increasing hazard function with a shape factor of 2.606. Covariates can be described as diagnostic or monitoring variables that affect (increase or decrease) the hazard rate. Viscosity has been described as the most significant covariate based on the covariate selection process and hypothesis testing (p-values). This has been supported by relevant literature. Pirro and Wessol [35] stated that the single most important physical
characteristic of hydraulic fluid is its viscosity. It is also acknowledged in [36, 37] that viscosity is one of the most important properties of lubricating oil.

Viscosity is a measure of fluid flow at a given temperature. Viscosity index (VI) is a dimensionless quantity that describes the behaviour of viscosity with changing temperature. The VI test (ASTM D 2270) measures the kinematic viscosity of the oil at 40°C and 100°C. Lubricating oil with high VI will show minimal change in viscosity with temperature change while the oil with low VI will demonstrate a significant change in temperature with temperature change. The viscosity as a covariate is an environmental factor that contributes to the failure rate of the HPU system. A unit increase in viscosity will reduce the hazard rate by 29.3%. Standard methods and techniques such as viscometer must be used to measure and monitor the viscosity index of the hydraulic oil since it is the main source of degradation [38].

3.8. Maintenance optimization

Maintenance optimization ensures that preventive maintenance is carried only when required. Different criteria can be considered for optimizing maintenance including cost, availability, safety and profit. According to [34] lowest-cost optimization is often the maintenance goal. The HPU is a highly reliable system with in-built redundancy. Hence, cost optimization is required to determine the frequency of preventive maintenance as seen in Figure 12.

![Figure 12. Lowest-cost optimization for HPU pump](image)

The current maintenance policy for the drilling HPU is to replace the unit at a predetermined interval (410 hours) or failure (block replacement policy). The preventive maintenance interval is based on the recommendation provided by the OEM. The monthly working age for the HPU in a month is 410.4 hours while the monthly maintenance plan includes an oil change and regular maintenance which could improve the quality of viscosity (covariate) and reduce the hazard rate. Moreover, the use of condition monitoring is to optimize maintenance tasks which involve cost minimization [39]. The optimal point obtained from the graph is 735 hours (One month and three weeks). The cost per unit time for the normal maintenance schedule is 0.25 while the cost for optimal maintenance is 0.20. The preventive maintenance interval is therefore extended by 79.27% which will reduce the maintenance costs. This means that the HPU is overly maintained and the current maintenance policy which states that preventive maintenance should be done every 410 hours should be changed to 735 hours.

4. Conclusion

The purpose of this paper is to explore how to support PHM with Principal Component Analysis (PCA) to enhance maintenance optimization of complex industrial systems. Eight steps have been proposed to enable PHM to handle big technical data and provide benefits by extending the maintenance interval
which will lead to a reduction in maintenance costs. The results show that PCA-supported PHM optimizes and extends the preventive maintenance interval by 79.27% which might lead to maintenance cost reduction.

References

[1] Giunta G, Bernasconi G, Giro R A and Cesari S 2020 Digital Transformation of Historical Data for Advanced Predictive Maintenance. In: Abu Dhabi International Petroleum Exhibition & Conference,
[2] Sharma A, Yadava G and Deshmukh S 2011 A literature review and future perspectives on maintenance optimization, J Qual Maint Eng 17(1), 5-25
[3] Arthur N and Lusher J 2015 Asset Integrity: Managing the Maintenance Burden. In: SPE Offshore Europe Conference and Exhibition, Sept. 8 - 11, Aberdeen, Scotaland, Paper Number: SPE-175505-MS
[4] Mobley R K 2002 An Introduction to Predictive Maintenance (Second Edition), ed R K Mobley (Burlington: Butterworth-Heinemann), 202-16
[5] O’Brien L 2006 Providing the business value proposition for automation, Special Report Automation Systems, ABB Review
[6] Caballé N C, Castro I T, Pérez C J and Lanza-Gutiérrez J M 2015 A condition-based maintenance of a dependent degradation-threshold-shock model in a system with multiple degradation processes Reliab Eng Syst Saf 134, 98-109.
[7] Peng H and van Houtum G-J 2016 Joint optimization of condition-based maintenance and production lot-sizing, Eur J Oper Res 253 94-107
[8] Liu B, Liang Z, Parlikad A K, Xie M and Kuo W 2017 Condition-based maintenance for systems with aging and cumulative damage based on proportional hazards model, Reliab Eng Syst Saf 168 200-9
[9] Wardhana K and Hadipriono F 2003 Analysis of recent bridge failures in the united states, J Perform Constr Facil 17(3), 144-50.
[10] Wu X and Ryan S M 2011 Optimal replacement in the proportional hazards model with semi-Markovian covariate process and continuous monitoring, IEEE Trans Reliab 60 580-9
[11] Mo H, Xie M 2016 A Dynamic Approach to Performance Analysis and Reliability Improvement of Control Systems With Degraded Components 46 1404-14.
[12] Lin D Y and Wei L J 1989 The robust inference for the Cox proportional hazards model, J Am Stat Assoc 84 1074-8.
[13] Jardine A K S and Tsang A 2013 Maintenance, Replacement, and Reliability: Theory and Applications, Second Edition.
[14] Banjevic D, Jardine A, Makis V, and Ennis M 2001 A control-limit policy and software for condition-based maintenance optimization, INFOR: Inform Syst Oper Res 39(1), 32 - 50.
[15] Ghasemi A, Yacout S and Ouali M 2007 Optimal condition based maintenance with imperfect information and the proportional hazards model, Int J Prod Res 45, 1012 - 989.
[16] Wu X and Ryan S M 2010 Value of condition monitoring for optimal replacement in the proportional hazards model with continuous degradation, IIE Trans 42, 553-63.
[17] Tian Z and Liao H 2011 Condition based maintenance optimization for multi-component systems using proportional hazards model, Reliab Eng Syst Saf 96, 581-9.
[18] Lam J Y J and Banjevic D 2015 A myopic policy for optimal inspection scheduling for condition based maintenance, Reliab Eng Syst Saf 144, 1-11
[19] Hu J and Chen P 2020 Predictive maintenance of systems subject to hard failure based on proportional hazards model, Reliab Eng Syst Saf 196, 106707
[20] Jafari L, Naderkhani F and Makis V 2017 Joint optimization of maintenance policy and inspection interval for a multi-unit series system using proportional hazards model, J Oper Res Soc 69, 36-48
[21] Lin D, Banjevic D and Jardine A K S 2017 Using principal components in a proportional hazards model with applications in condition-based maintenance, J Oper Res Soc 57, 910-9
[22] Abdi H and Williams L J 2010 Principal component analysis, Wiley Interdisciplinary Reviews: Comput Stat 2, 433-59.
[23] Lin D, Banjevic D and Jardine A K S 2006 Using principal components in a proportional hazards
model with applications in condition-based maintenance J Oper Res Soc 57, 910-9
[24] Du Y, Wu T, Zhou S and Makis V 2020 Remaining useful life prediction of lubricating oil with
dynamic principal component analysis and proportional hazards model, J Eng Tribol 234, 964 -
71.
[25] Tang Z, Zhou C, Jiang W, Zhou W, Jing X, Yu J, Alkali B and Sheng B 2014 Analysis of significant
factors on cable failure using the Cox proportional hazard model, IEEE Trans Power Delivery 29
951-7.
[26] Thijssens O W M and Verhagen W J C 2020 Application of extended Cox regression model to time-
on-wing data of aircraft repairables, Reliab Eng Syst Saf 204, 107136
[27] Portet S 2020 A primer on model selection using the Akaike information criterion, Infect dis modell
5, 111-28.
[28] Konishi S and Kitagawa G 2008 Information Criteria and Statistical Modeling: Springer Science &
Business Media)
[29] Kaplan E L and Meier P 1958 Nonparametric estimation from incomplete observations, J Am Stat
Assoc 53, 457-81
[30] Wang F, Chen X, Dun B, Wang B, Yan D and Zhu H 2017 Rolling bearing reliability assessment via
kernel principal component analysis and Weibull proportional hazard model, Shock Vib 2017, 1-
11.
[31] Lin D, Banjevic D and Jardine A K S 2006 Using principal components in a proportional hazards
model with applications in condition-based maintenance, J Oper Res Soc 57 910-9
[32] Jardine A K S, Joseph T and Banjevic D 2008 Optimizing condition-based maintenance decisions for
vibration monitored bearings, J Q Maint Eng 5, 192-202.
[33] Hu J and Chen P 2020 Predictive maintenance of systems subject to hard failure based on proportional
hazards model, Reliab Eng Syst Saf 196 106707
[34] Campbell J D, Jardine A K and McGlynn J 2016 Asset Management Excellence: Optimizing
Equipment Life-Cycle Decisions: CRC Press)
[35] Pirro D M and Wessol A A 2001 Lubrication Fundamentals (New York: Marcel Dekker)
[36] Mobley R K, Higgins L R and Wikoff D J 2014 Maintenance Engineering Handbook: McGraw-Hill
Education)
[37] Kučera M, Köpčanová S and Sejkorová M 2020 Lubricant analysis as the most useful tool in the
proactive maintenance philosophies of machinery and its components, J Manage Syst Prod Eng
28, 196-201.
[38] Palub 2020 Understanding the viscosity index of a lubricant. In: Q8Oils, ed V I VI
[39] Holme A 2006 Experiences and Lessons Learned From Utilizing Automated Reliability-Centered
Maintenance on Drill-Floor Equipment To Optimize Operation and Maintenance Planning. In:
IADC/SPE Drilling Conference,