Estimating residual life of equipment using subjective covariates

J.S. Schoeman[a] and P.J. Vlok[b]

Received 06 February 2015, in revised form 1 July 2015 and accepted 15 October 2015

The acquisition of historical failure data and condition monitoring data in practice has proven to be very difficult. This data is often used in prognostic survival models which allow reliability estimates and eventually residual life predictions to be made. The predictions form an important part of proactive maintenance. This study then aims to verify whether subjective covariates obtained from experts can be used to populate the prognostic survival models in place of the classical objective covariates, this study is comprised of mixed methods. Six different prognostic survival models were considered, the six models most found in literature, and a case study was conducted. The case study was used to validate the theory of using the subjective covariates by comparing results obtained from the subjective covariates to results obtained when using objective covariates. The results obtained from the subjective covariates proved to be less conservative in all cases. Thus it is possible to use subjective covariates to predict the residual life of physical assets but it is inherently more risky.

Additional keywords: Survival analysis, Proportional Hazards Model, Accelerated Failure Time Model, Proportional Covariate Model, Prognostics

Nomenclature

Roman

x Discrete local event time
z Vector of covariate values

Greek

$\boldsymbol{\beta}$ Vector of regression coefficients
Ψ Functional term of PHM
ρ Intensity function for PWP

Subscripts

k Total number of covariates used

1 Introduction

Physical Asset Management (PAM) has become increasingly relevant when addressing the main pressure areas in the modern asset intensive industries. The three main pressure fields are financial, environmental and the health and safety of the employees[1,2,3]. PAM is becoming an increasingly popular field to optimize because of its potential to positively affect all three of the pressure fields, this is emphasized by ISO[4] and BSi[5]. An important aspect of PAM includes the maintenance of the assets, it might sound simple enough but there are many aspects that should be considered when choosing a maintenance strategy for assets[4].

Gorjian *et al*.[1] state the reliability of assets can be monitored and managed by maintaining them according to specific guidelines. Various maintenance guidelines have emerged over time from the different maintenance strategies. The three main maintenance strategies include *life improvement maintenance*, *reactive maintenance* and *proactive maintenance*. This paper focuses on the proactive maintenance strategy to utilize its advantages of minimizing the down time of assets and preventing unexpected failures. The proactive maintenance strategy can be split further into two different tactics, namely the predictive and preventative tactics.

Preventative maintenance blindly replace or maintain assets, no matter what their current state is, thus incurring unnecessary costs. The predictive maintenance tactic on the other hand only considers the current state of an asset by monitoring relevant characteristics of the assets. The maintenance schedules for preventative maintenance are created by using statistical models based on the historical failure times of the relevant assets. Predictive maintenance is dependent on a process known as Condition Monitoring (CM), which collects the data describing the current state of assets from various sensors. A middle ground between the two proactive maintenance tactics, known as prognostics, addresses the drawbacks of both procedures and is capable of providing more accurate reliability estimates by combining the two procedures.

Prognostics is an engineering discipline which makes use of survival models to estimate asset reliability. The survival models use both historical failure data from the preventative maintenance tactic and the CM data from the predictive maintenance tactic. The two data types are used as special variables known as covariates in the survival models.

One might now be asking: "If this strategy has tactics in place and have been used until now what is the problem?'' The answer to this is that the prognostic survival models can often not be used due to the lack of the data required. Sun[6] states that historical failure data need to be gleaned in industry as a result of poor record keeping. CM data is often not trustworthy or not available at all[7]. The data is generally seen as objective data since it is obtained from historical records and various sensor readings. Availability and integrity of this data has proven to be problematic and the main obstacle to overcome in the prognostics field. The industry has realized that some alternative need to be found for this objective data.

In an attempt to find an alternative to the objective data, subjective data will be considered for this study. Subjective data are subject to being biased and is effected by personal opinion. The data for this study is obtained from various people considered to be experts in a specified field, therefore seen as subjective data. A research question for this paper can

a. Master student. Stellenbosch University. P.O. Box 104, Kareeboom, Willowmore, 6445, South Africa. E-mail: jaco_schoeman@hotmail.com
b. Study leader and professor. Stellenbosch University. Private bag X1, Matieland, Stellenbosch, 7602, South Africa. E-mail: pjvlok@sun.ac.za

then be formulated as follow: "Can subjective covariates obtained from experts be used as covariates to populate prognostic survival models thus allowing the prediction of equipment residual life?"

Subjective covariates have been used in survival models in the past, but no study has been done to verify the results obtained. The aim of this paper is to establish whether subjective covariates obtained from experts can be used as a valid alternative to the normal objective covariates in prognostic survival models for physical assets. This will grant organizations the opportunity to utilize the data collected by their employees over the years to predict equipment reliability and ultimately the residual life. Following the given research question the null hypothesis for this study can be stated as: "Subjective covariates obtained from experts cannot be used to populate prognostic survival models to allow the prediction of the residual life of equipment." To be able to answer the research question and thereby either rejecting or not rejecting the null hypothesis several research objectives must be met.

The objectives listed below must be met in a systematic manner to ultimately allow the research question to be answered. The objectives are:

- Gain a thorough understanding of survival analysis and survival models.
- Determine suitable CM data to use as covariates.
- Evaluate applicability of various survival models.
- Establish a guide for selecting experts to obtain data from.
- Elicit subjective data from the selected experts.
- Deliver estimates of equipment reliability using the subjective data.
- Validate the theory tested by making use of a case study and objective data.

Systematically achieving these objectives will result in the research question being answered and therefore the deliverable of this study. An important part of this study is comparing the results obtained from the subjective data to that of results obtained from using objective data. This paper continues by first introducing the reader to survival analysis and then providing more detail on the survival models considered in this study. A case study then follows and lastly the results yielded by the case study.

2 Survival Models

Survival analyses make use of survival models to estimate equipment reliability. The data describing the equipment or systems can be classified into two different categories, namely repairable and non-repairable systems. With non-repairable systems the hazard, also known as Force of Mortality (FOM), is used to model the system. The Rate of Occurrence of Failure (ROCOF) would be used in the case of a repairable system. A trend test is done on the failure times of the data set to determine whether the data is from a repairable or non-repairable system, the Laplace trend test is a very popular test method and also used in this study.

There are various survival models used in the field of PAM. The models that are used in this study are the models which were found to be the most popular survival models, they constantly reappeared in the relevant literature[6,8]. Selected models were all applied to the data in this study, the models were selected according to how well they fit the data sets used. Only the model which could recreate the data fed into it the most accurately was used to calculate residual life predictions in the end. The models considered in this study are now reviewed only to give the reader a basic understanding of the principles which each models operates on.

2.1 Proportional Hazards Model

The Proportional Hazards Model (PHM) is also known as the Cox regression model and was originally developed in 1972 for analysing survival data in the medical discipline[8]. The basic principle of the PHM is the assumption that the hazard function of the subject/equipment considered, change in proportion with a specific underlying hazard function. Mathematically the PHM is represented by equation 1,

$$h(x|\boldsymbol{z}) = h_0(x)\Psi(\boldsymbol{z}(x);\boldsymbol{\beta}), \quad (1)$$

where $\Psi(\boldsymbol{z}(x);\boldsymbol{\beta})$ is a positive functional term that incorporates the effect of the covariates on the equipment. A popular form for the functional term is the exponential form and is also utilized in this study. The functional form of Ψ can thus be written as $\exp(\beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_k z_k)$, where k is the number of covariates used, $\boldsymbol{\beta}$ is a vector of regression coefficients and $\mathbf{z}$ is a vector containing covariate values. The underlying hazard function is then specified by h_0, in the parametric case the Weibull distribution is often used to describe the underlying hazard function. PHM's are widely used and very popular, various extensions of the model have also been developed over the years.

2.2 Prentice, Williams and Peterson Model

One such extension of the PHM was created and presented by Prentice *et al.*[9], this model is a generalization of the PHM to a proportional intensity function ($\rho(t|\boldsymbol{z})$). This intensity is also known as the ROCOF, therefore this Prentice, Williams and Peterson (PWP) model is an adaptation of the PHM to be able to accommodate repairable system data and still operates on the same assumption. The PWP model is written mathematically as equation 2.

$$\rho\big(t\big|N(t),\boldsymbol{z}(t)\big) = \rho_0(t - t_{S-1})\,\Psi(\boldsymbol{z}(t);\boldsymbol{\beta}) \quad (2)$$

Here t is the continuous global time and $N(t)$ is the number of failures at the instant t. The number of failures divide the system into strata, thus stratum S can be written as $S = N(t) - 1$. An alternative to the PHM, and therefore also the PWP, is a model known as the Accelerated Failure Time Model (AFTM).

2.3 Accelerated Failure Time Model

An alternative to the popular PHM can be found in the form of the AFTM. Komárek *et al.*[10] explain how the AFTM operates on a different assumption of the effect of the covariates on the system. The AFTM establishes a direct relationship between the time to failure and the covariates, equation 3 clearly shows this.

$$Log\ T = \mu + \boldsymbol{\beta}'\boldsymbol{z} + \sigma\epsilon \quad (3)$$

The time of failure is given by T, while μ denotes the intercept value. A scale parameter is denoted by σ and the residual error by ϵ. The distribution of the residual error will determine the distribution of the failure times. A table of

corresponding distributions is given by Qi[11]. These models are often used to model electronic components.

2.4 Additive Hazards Model

Another model that can be used is the Additive Hazards Model (AHM). As for PHM the name suggests what the underlying assumption for this model is. The AHM assumes that covariates have an additive effect on the system hazard or FOM. In its most basic form the AHM is given by,

$$h(x|\boldsymbol{z}_i) = h_0(x) + \alpha(\boldsymbol{z}_i), \quad (4)$$

where h_0, is the baseline FOM of the system considered. The additive coefficient can have different functional forms. The simplest polynomial form is utilized in this study, $\alpha(\boldsymbol{z}) = \boldsymbol{\beta}'\boldsymbol{z} = \sum_{i=1}^{k} \beta_i z_i$.

This additive effect of the covariates on the FOM enables the system to have a non-zero value for its FOM at the start of operation[12]. The AHM can reveal information about the maintenance done on the system by analyzing the value of $\alpha(z_i)$. A $\alpha(z_i)$ value greater than zero indicates that the hazard function immediately after the failure is higher than at the time of failure, a zero indicates that it is the same as before the failure and less than zero indicates that the FOM immediately after the failure is lower than at the time of failure. This model is often considered when the system considered is better off after the maintenance actions but not as good as new.

2.5 Proportional Covariate Model

The Proportional Covariate Model (PCM) is also an extension of the PHM and is largely based on the PHM[13]. The PCM and the PHM both assume proportionality of the system FOM. A PCM assumes that the covariates, or a function of the covariates, is proportional to the system FOM instead of using a baseline hazard (or FOM) function like the PHM. Sun developed the PCM in order to adress some of the shortcomings of the PHM listed by Sun[6].

The PCM can be represented in its most basic mathematical form as equation 5, where only one covariate is used therefore not a function of covariates.

$$Z_r(x) = C(x)h(x) \quad (5)$$

The simplest form is represented here, $Z_r(x)$ is the value of a single covariate, more complicated models could have this as a function of the covariates. A baseline covariate function is denoted by $C(x)$ and the FOM is denoted by $h(x)$. The PCM can be used for both repairable and non-repairable systems and is capable of being made use of even when historical failure data of the specific piece of equipment considered is not available[6].

2.6 Proportional Odds Model

Initially the Proportional Odds Model (POM) was also considered to be an alternative to be used as a survival model. However this model is used to model systems where the effect of the covariates diminish over time[8,14,15]. The covariates used to monitor the health of physical assets portray the degradation of the assets. Their effect on the assets increase over time, contrary to what the POM is suitable for. The POM was therefore not further considered to be a valid alternative.

All these models relate to the reliability of equipment to the effect of covariates monitored in some manner. Reliability monitoring allows the prediction of equipment residual life. Applicability of each model must be tested for each data set they are applied to. Various goodness of fit methods are available for each model and reveals whether the model can be considered as a valid model to use or not. This study aims to confirm whether subjective data can be used to populate these survival models instead of using the traditional objective data which is so rarely available.

3 Method of Experiment

To establish whether subjective data can be used as an alternative to objective data or not, the results obtained from the subjective data must be compared to that of the classical objective data. This section describes the structure followed when conducting the study. The first objective is to select people whom are considered as experts in the field of interest to obtain the subjective data from.

3.1 Selecting Experts

There is no set standard as to when a person can be considered as an expert, thus the researchers have a large influence on which people are to be used as experts. In an attempt to provide a rough guideline of when a person is considered to be an expert, Schoeman[16] provides a table. This table shows when a person is generally considered to be an expert based on their experience, education and occupation. The bottom line amounts to that when a person is considered as an expert through the consensus of others they can be classified as an expert in their specific field. So when selecting the experts it is important to consult co-workers and other colleagues that operate in the same field or one very similar. After the experts have been selected the covariates to be used in the survival models need to be considered.

3.2 Determine Covariates

Different systems require different characteristics to be monitored in order to be able to deduce the current health or state of the system. Aspects such as the practicality of measurement and operating conditions must be considered as well as what parameter will deliver characteristic signs. The parameters to be considered differ for different machine types.

A standard[17] created to aid the CM process published by the International Organization for Standardization (ISO) in 2002 provides a table that relates different types of machines to several general parameters to be considered. This table reveals eight parameters that relate to all the machine types covered in the table, these parameters are:

- Temperature,
- Noise,
- Vibration,
- Acoustic techniques,
- Oil pressure,
- Oil consumption,
- Oil (tribology),
- Speed.

The table provided by ISO[17] offers general guidelines and will direct the researcher in the right direction. Asking the people selected as experts will however be the most effective

manner of establishing which parameters to use as covariates. Once the experts and the covariates have been selected the next step is to extract the necessary data from the experts and obtain objective data to have something to compare the subjective data with.

3.3 Data Requirements and Elicitation

It is important to know exactly what data is needed before going to extract it from the selected experts. The prognostic survival models, as mentioned before, use both historical failure data and CM data. It is however also necessary to know whether the data was recorded at a failure or if it was simply during an inspection.

3.3.1 Required data

The data to be extracted from the experts can best be tabulated to show what all the necessary information is. In table 1 all the necessary data is displayed in a typical set-up, here X_i is the time of the event while C_i is an event indicator where 1 indicates a failure and 0 an inspection. A matrix of covariate values (**Z**) is split into individual vectors Z_1, Z_2, …, Z_k, where k is the total number of covariates considered. A value of the covariates at the start and end of each operating period is required.

Table 1: Typical format for table of necessary data.

X_i	C_i	Z_1		Z_2		…	Z_k	
		Start	End	Start	End	…	Start	End
…	…	…	…	…	…	…	…	…

When it is known what data is required and the experts to elicit the data from have been identified the next step can be taken. The following step is to obtain the data, eliciting subjective technical data can be difficult and cause confusion. Thus it is necessary to have a process that is straight forward.

3.3.2 Elicitation of Data

Asking experts to deliver a dataset from scratch with no guidelines given is not a viable option. Therefore three different techniques are considered when extracting data. The first technique is to provide the experts with failure times and have them deliver the trends of the various covariates over those periods of operation. Covariate values are needed from the start of operation up to the time of the event.

The second technique is then to provide the experts with a partial data set again. This time they are given the progression of the covariate values from an initial point in time up to a certain point. They are then required to provide the event times for the covariate values given.

The third technique is simply a combination of the first two techniques. One way of doing this might be to have a data set and elicit the first half of the data set with the fist technique and the second half with the second technique. A random selection between the two methods can also be applied if desired.

With all three of the techniques it is required that a consensus is reached about the values provided. The values provided by the individual experts are reviewed by all the other experts taking part in the study and a consensus must be reached on the values provided before moving on. Should it happen that the experts cannot agree on a certain value a new scenario will be presented to them until they can provide some value which they all agree on. A flow chart is given of this elicitation process by Schoeman[16]. It is advised to base the data given to the experts on an actual objective data set, thus the objective data can be used in parallel to the subjective data allowing validation of the theory in the end.

3.3.3 Manipulation Required

Before populating the survival models with the data, subjective or objective data, it might be necessary to slightly manipulate the data to fit each survival model's requirements. Slight modifications is to be done to all the data necessary before starting the process populating the survival models and selecting the most appropriate one.

3.4 Model Selection

The most appropriate prognostic survival model will differ from data set to data set. It is therefore important to test which model is most appropriate each time a new data set is being considered. To do this each model has to be populated with the data and then a Goodness of Fit (GOF) test is to be done on each.

A GOF test is done for each model, if the test reveals that the model can be used for the specific data set it is then used to recreate the data set. All models considered to be a viable model for the data set recreates the subjective data set, the model that most accurately recreates the data set is then considered to be the most appropriate survival model for that specific case.

3.5 Validation of Theory

Validation of this theory is very important, as it is the core focus point of this study. The prognostic survival models used in this study have all been validated in prior studies, this study therefore only aims to determine whether they can be used with subjective data. Residual Life (RL) predictions can be made from the reliability predictions obtained from the survival models.

To validate the theory of using subjective data as a substitute for objective data both data types are necessary as mentioned before. The most appropriate model is chosen as explained before. Once the model has been chosen it can be used to obtain reliability estimates, therefore allowing RL predictions to be made. This has to be done with the most appropriate models for the subjective and objective data. The results from the subjective data can then be compared to the results from the objective data. Since the models have all been validated and used in industry with objective data the comparison between the results can be used as the validation. This will be conducted in the form of a case study, covered in the following section.

4 Case Study

Validation of the theory is done in the form of a case study, this is necessary to obtain all the relevant data from industry and not solely from an academic environment. The case study was conducted at a heavy minerals processing company in South Africa. This organization uses smelting furnaces which require large power transformers.

It was decided to use two of the large power transformers for this case study since both of the transformers considered are approaching their end-of-life phase. CM data is available for both transformers and there are several experts on the

plant who were glad to take part in the study. The experts that took part in the study all gave consent to take part in the study and they themselves can be considered as experts in the field of CM on power transformers. There is uncertainty about the time span of the end-of-life phase of the power transformers and it is therefore a good opportunity to add value to an organization in industry as well as the academic field.

4.1 Experts and Covariates

The experts were chosen; only five experts could participate in the study. They were consulted about which parameters to use as covariates. Various parameters were considered, but it was eventually decided to use only one covariate, the Degree of Polymerization (DP) of the insulating paper around the coils of the transformer which is submerged in a mineral oil. This indicates the tensile strength of the insulating paper by considering the number of glucose rings in the cellular composite of the oil molecules. The DP is determined by taking an oil sample from the transformer and analysing the small paper bits within the oil. Only the DP was used because the other parameters which are measured only fluctuate outside of a specified range when there are irregularities or definite failures or break downs.

When new, the DP of a transformer ranges anywhere from 1000-1200 number of glucose units per cellulose chain[18]. The number of glucose units per cellulose chain is simply known as the DP, where the DP has the value of the average number of glucose units per cellulose chain. In the case of this study a new transformer is considered to have a DP of 1000. A power transformer is considered to fail as soon as its DP value dips to 200 or below, as it then simply becomes too dangerous to continue using the transformer.

There is a set standard which has been accepted by the industry to calculate how the DP of a transformer degrades over time depending on its loading factor. This standard is used by Zhong[19] to simulate the ageing of power transformers. A data set can therefore be created from the subjective opinions of the experts while another is created that can be seen as industry standard objective data by making use of the method used by Zhong[19]. These two data sets can then be used to populate the most appropriate survival model and RL predictions can be made from both data set estimates and compared.

4.2 Data and Most Appropriate Model

The subjective data was elicited from the experts and proved to belong to a non-repairable system. Survival models therefore operate with the FOM and not the ROCOF, the PWP is therefore also not an applicable survival model. Experts were given ten different scenarios (operating periods), for half of the scenarios the failure times were provided and the experts had to provide the regression of the DP. The other half of the scenarios are the opposite, where the experts provide the failure times from DP values. The industry standard data set was created by interpolating the failure times and DP values provided by the experts. A final data set of ten failure times and the corresponding covariate values is used as the subjective data set. Where the data elicited is obtain through the two methods described.

Survival models that were considered then included the PHM, AFTM, AHM and the PCM. After the GOF test for each model was done it was discovered that the AHM is also not an applicable model because a negative hazard was yielded. The other three models were populated with the subjective data and the data set was then recreated by using the models. The results yielded are summarized in table 2.

Table 2: Error when recreating the data set.

Survival model	Error (years)
PHM	2.037
AFTM	3.930
PCM	0.065

Here it can be seen that the model that recreated the subjective data set the most accurately is the PCM. The PCM is thus the most appropriate survival model in this specific case. Therefore the PCM is the only model used to estimate the reliability allowing RL predictions in this study.

4.3 Results

Initial estimates of the PCM parameters are made by only considering the failure times of the association events. Thus to obtain the initial estimates the first half of the data set is used to derive the initial estimates and the second half is then used to update these estimates. Doing this also ensures that the older data does not bear as much weight as the new data. This is done to ensure that any error in the estimates do not propagate and grow along with time.

The initial estimate of the PCM parameters for all five of the experts is therefore the same. The updated parameter values for each expert then differs according to their data. Parameters that are estimated include the shape (γ) and scale (η) parameters of the Weibull distribution, because that is the functional form chosen for the transformer's failure distribution. The change in the parameter values is easily observed when examining the reliability and/or the FOM curves of the system. The initial reliability estimation and the updated curves of each expert are displayed in figure 1 Considering this, it can be seen that the shape parameter γ, decreases while the scale parameter η increases for all of the experts when compared to the initial estimates.

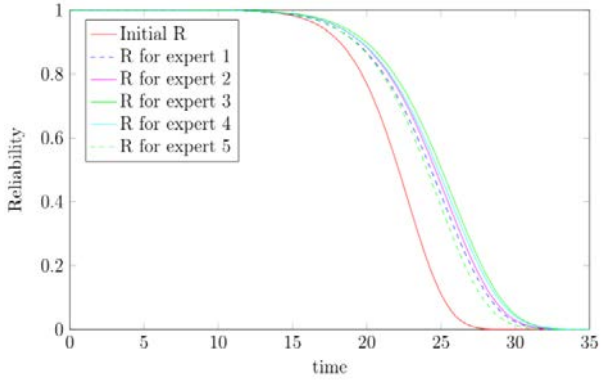

Figure 1: Initial vs. updated reliability with the time in years.

The parameter values can also be used to create the FOM curves of the initial estimates as well as that of the updated values for each expert, see figure 2.

Here it can clearly be seen what the difference is between the initial estimate, calculated only from historical failure times, and the updated estimates. Two industry standard data sets were created and compared to the subjective data set, both were created using the method given by Zhong[19]. The

"industry standard one" data set was created by calculated the degraded DP values up to the same event time as that of the subjective data set. The "industry standard two" data set was created by calculating the DP value to the same level of degradation as the subjective average, no matter in what time those values are reached. The effect of the change in γ and η is clearer when plotting the FOM, it can be seen that after the parameter estimates are updated the FOM is considerably lower than the initial estimate.

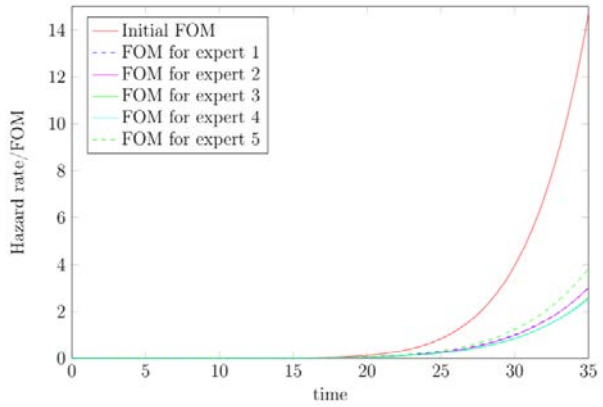

Figure 2: Initial vs. updated FOM with the time in years.

The expert opinions are all relatively near one another, an average of their parameter values is taken to be the average expert opinion. The average of the reliability obtained from the expert's opinions as well as the FOM are compared to the industry, shown in figure 3 and 4 respectively. These two figures illustrate the effect of the experts' opinion having a larger shape parameter and a smaller scale parameter.

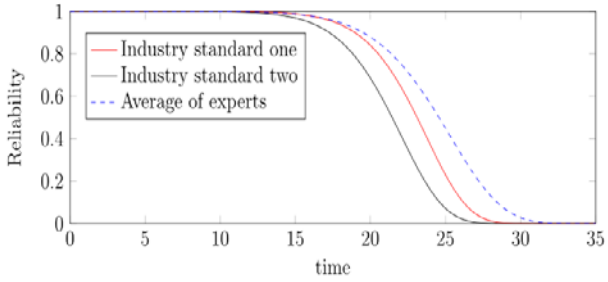

Figure 3: Reliability of expert opinion and industry standard with the time in years.

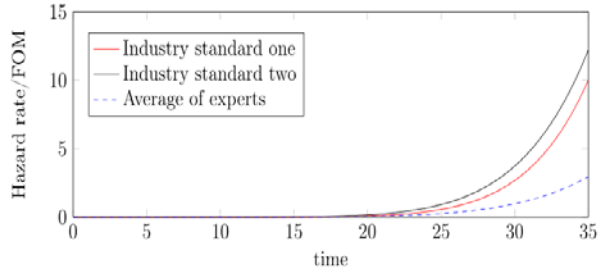

Figure 4: FOM of expert opinion and industry standard with the time in years.

The plot in figure 3 spans to 35 years. When extending the plot to up to 80 years it can be seen that the difference between the industry standard's FOM and that of the experts become exponentially larger. Figure 5 has a worrying trait; the FOM based on the expert opinions is considerably lower than those based on industry standard data.

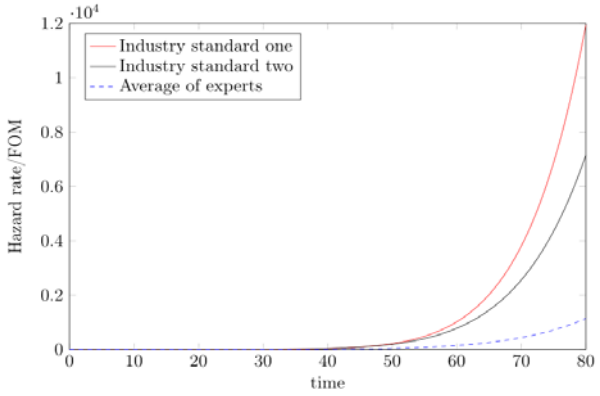

Figure 5: FOM of experts and industry over 80 years.

The data set with the ten different scenarios which was provided to the experts was used to make RL predictions of the transformer for every scenario. A RL prediction was made by making use of the parameter values from the average expert opinions, the industry standard one and two data sets. RL predictions obtained are graphically displayed for each scenario in figure 6. This figure illustrates how the RL estimate based on the expert opinions is always greater than those based on the industry standard.

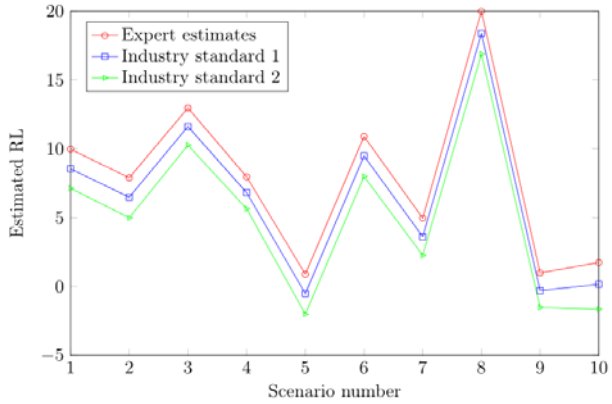

Figure 6 RL predictions.

Examining the figures reveal important information. This information is discussed and elaborated on in the following section.

5 Discussion

The results obtained from the case study revealed traits that need to be taken notice of, especially when considering the use of subjective data in survival analysis. It is seen that using the updated parameter estimates are very important because it decreases the change of the accumulation of error over time by using the CM data and the new failure time to update the parameter values. This emphasizes the applicability and relevance of prognostics in the PAM field.

The reliability and FOM curves of the average expert opinion were then compared to that of both industry standards. Since the scale parameter of the experts' opinion is larger than that of both industry standards, the curves show how the system reliability is stretched out over a longer time span before settling to zero. The increase in the expected life is also longer than that of the industry standards because the shape parameter of the experts is smaller than both standards.

This means that the rate at which the reliability degrades is slower than that of both standards.

All of these aspects have an effect on the RL estimates, the predictions from figure 6 show that the industry standard predictions are always smaller than the expert's predictions. In some cases the industry standard predictions are negative; this indicates that according to that standard the DP level should already be below 200.

The expert knowledge is proven to be less conservative than the industry standard data on all occasions. There are also cases where the RL prediction calculated by using the industry standard data is negative while the experts still think the transformer can continue operation. The negative value indicates that according to the industry standard data the transformer should have failed already and cannot continue operation.

In conclusion, the experts' knowledge can be used as covariates to populate the survival models. The opinions which the experts delivered are less conservative than both of the industry standard data sets, this is an alarming result because many organizations still only use their employees to monitor their physical assets. It is concerning in this case because it shows that the experts always think that the transformers can continue operation for longer than what they are supposed to. This is why designing with a safety factor is so important. The research question can therefore be answered with a tentative "yes", because even though the experts' knowledge delivered results that are not far off compared to the industry standard, they insinuate that the experts are underestimating the rate of degradation of the transformers. This can be dangerous if no CM is applied to the transformers and the experts decide on how long to operate them at which levels. When the answer to the research question is yes it also means that the null hypothesis can be rejected.

6 Future Research

During the execution of this study areas emerged which will allow for the improvement of this study even though all the objects were met and the null hypothesis rejected. Questions and ideas that were noted should be addressed and has the potential to reassure the conclusion reached in this study. The questions and ideas are formulated into recommendations as provided below:

- It is not certain if the knowledge of experts are less conservative concerning the RL estimates for all types of physical assets. Therefore it would provide valuable insight if this study could be conducted on different asset types. The results can then be compared to see if the experts' knowledge reveal the same characteristic for all of the asset types.
- This study could possibly deliver more accurate results with survival models that were not reviewed. It would therefore be suitable to assess the applicability of more survival models.
- In industry generally anything somebody asks for they would have already liked to have by the time they ask, therefore the speed at which RL estimates are made should be as quick as possible. A toolbox which combine all the different survival models and automatically runs the applicability tests and delivering real time estimates would prove favourable in industry.
- The data collection process is a tedious one, especially when asking each expert to provide a data set based on their knowledge. A way to ease and quicken this process will be of great aid when conducting a similar study.

All of the recommendations stated above are to improve research conducted in the future.

References

1. *Gorjian N, Ma L, Mittinty M, Yarlagadda P and Sun Y*, A review on degradation models in reliabiltiy analysis, In *Engineering Asset Lifecycle Management*, 369-397, Springer, London, 2010.
2. *Hamann R*, Mining companies' role in sustainable development: the 'why' and 'how' of corporate social responsibility from a business perspective, *Development Southern Africa*, 20(2), 237-254, 2003.
3. *Campbel JD, Jardine AK and McGlynn J*, Asset management excellence: Optimizing equipment life-cycle decisions, CRC Press, 2011.
4. ISO 55000, Asset management – overview, principles and terminology, 2014.
5. PAS 55-1, Asset management part 1: Specification for the optimized management of physical assets, British Standards Institution, 2008.
6. *Sun Y*, Reliability prediction of complex repairable systems: an engineering approach, PhD Thesis, Queensland University of Technology, 2006.
7. *Mann L, Saxena A and Knapp GM*, Statistical-based or condition-based preventive maintenance?, *Journal of Quality in Maintenance Engineering*, 1(1), 46-59, 1995.
8. *Carstens WA*, Regression analysis of Catepillar 793D haul truck engine failure data and through-life diagnostic information using the proportional hazards model, Masters Thesis, Stellenbosch University, 2012.
9. *Prentice RL, Williams BJ and Peterson AV*, On regression analysis of mutlivariate failure time data, *Biometrika*, 68(2) 373-379, 1981.
10. *Komárek A, Lesaffre E and Hilton JF*, Accelerated failure time model for arbitrarily censored data with smoothed error distribution, *Journal of Computational and Graphical Statistics*, 14(3), 726-745, 2005.
11. *Qi J*, Comparison of proportionan hazards and accelerated failure time models, PhD Thesis, University of Saskatchewan, 2009.
12. *Gorjian N, Ma L, Mittinty M, Yarlagadda P and Sun Y*, A review on reliability models with covariates, In *Engineering Asset Life Cycle Management*, 385-397, Springer, London, 2010.
13. *Sun Y, Ma L, Mathew J, Wang W and Zhang S*, Mechanical systems hazard estimation using condition monitoring, *Mechanical systems and signal processing*, 20(5), 1189-1201, 2006.
14. *Murphy SA, Rossini AJ and Van der Walt AW*, Maximum likelihood estimation in the proportional odds model, *Journal of the American Statistical Association*, 92(439), 968-976, 1997.
15. *Royston P and Parmar MKB*, Flexible parametric proportional hazards and proportional odds models for censored survival data, with application to prognostics

modelling and estimation of treatment effects, *Statistics in Medicine*, 21(15), 2175-2197, 2002.
16. *Schoeman JS*, Estimating residual life of equipment using subjective covariates, Masters Thesis, Stellenbosch University, 2015.
17. ISO 13380, Condition monitoring and diagnostics of machines - General guidelines on using performance parameters, 2002.
18. *Gray IAR*, Evaluation of Transformer Solid Insulation, Technical Report, Transformer Chemistry Services, 2006.
19. *Zhong Q*, Power transformer end-of-life modelling: Linking statistics with physical ageing, PhD Thesis, University of Manchester, 2011.

 © SAIMechE All rights reserved.