A Bayesian inference reliability evaluation on the corrosion-affected underground high voltage power grid

Hang Zhou[a*], Fan Li[b], Michelle Le Blanc[c], Jingzhe Pan[d]

[a]*Institute for Manufacturing, Department of Engineering, University of Cambridge, Cambridge, United Kingdom CB3 0FS*

[b]*Amey plc, London, United Kingdom EC4A 1AB*

[c]*Electricity Transmission, National Grid, Warwick, United Kingdom CV34 6DA*

[d]*Department of Engineering, University of Leicester, Leicester, United Kingdom LE1 7RH*

Abstract:

The underground high-voltage power transmission cables are high value engineering assets that suffer from multiple deteriorations through-out life cycles. Recent studies identified a new failure mode – the pitting corrosion deterioration on the layer of phosphor bronze reinforcing tape, which protects the oil-filled power transmission cables from oil leakage due to deterioration of the lead sheath. Two models estimating the phosphor bronze tape life were established separately in this study. The first model, based on mathematical fitting, is generated using a replacement priority model from the power supply industry. This is considered as an empirical-based model. The second model, based on the corrosion fatigue mechanism, utilizes the information of the pit depth distribution and the concept of pit-to-crack transfer probability. The Bayesian inference approach is the conjunction algorithm to update the existing probability of failure model with the newly identified failure modes. Through this algorithm, the integrated probability of failure model contains a more comprehensive background information while maintaining the empirical knowledge on the engineering assets' performance.

* Corresponding author (hang.zhou@eng.cam.ac.uk).

Keywords: tin-bronze corrosion, underground power transmission cable, failure analysis, engineering asset management, Bayesian Inference, Markov Chain Monte Carlo, Metropolis Hastings algorithm

1. Introduction

The estimation of the remaining useful life (RUL) for engineering assets is a valuable tool for decision making in the majority of engineering industries. The accuracy of RUL estimation, especially when interpreted as the probability of failure (PoF) of engineering assets, enhances the confidence to the decision-making process when performing maintenance and replacing assets at their end of life as it improves the understanding of the assets' reliability. An engineering asset, at its design stage, the initial anticipated service life is estimated considering all the known failure modes. However, during the operation of the asset, new failure modes are at times discovered and identified due to better understanding of the assets and the deteriorations in different operational conditions. These new failure modes are new information and influence the established PoF model of the assets which is the existing empirical-based PoF model of the assets [1, 2]. This model generalises the PoF of all the assets in service. The advantage of an 'one-model-fits-all' solution is that the model can describe the reliability of the asset in general with a simple approach and is easy to be understood for deployment in the industry. The disadvantage, however, is also obvious: such a model cannot account for the variety of asset operational conditions. There is evidence that, as more assets approach the end of the designed life, the predictions fail to be convincing. The reason for the inadequate precision of the predictions has been identified in recent work, a new failure mode is discovered to be one of the major contribution of the research target's deterioration, and new PoF model is developed based on the newly identified failure mode[3-5]. Since the new failure modes follow distinctive failure mechanisms, the new PoF models are therefore named as the mechanism-based model. There is currently a research gap on the updating of the empirical-based PoF model with the identification of new failure modes, where the failure modes generate novel PoF models towards key components of the asset system, but at the same time influences the assets' reliability at a system level. This paper proposes a novel algorithm based on the Bayesian inference method which serves as a conjunction to update the empirical-based model with the mechanism-based model to achieve a more comprehensive asset performance evaluation.

When discussing engineering asset management, there are two major approaches to assist decision making: the RUL estimation and the probability of failure estimation. The approach of RUL estimation has been researched widely in different engineering disciplines. Martin et.al. [6] studied the remaining life of the transformer insulation and concluded that the remaining life is strictly related to paper insulation degradation. Khalifa et.al [7] developed a quantitative model for gas turbines on a risk-based maintenance and remaining life assessment. Segovia et.al [8] created a Cox model based on corrosion studies of power transmission tower remaining life estimation. Ahmadzadeh and Lundberg [9] applied an artificial neural network (ANN) method in the remaining useful life estimation of grinding mills. Animah and Shafiee [10] discussed the remaining life estimation for offshore oil and gas assets, with a proposal for the decision making of life extension.

The other approach – the PoF estimation - is discussed by the following researchers. Su et al. [11] applied a one-dimensional integral approach to the probability of failure estimation for geotechnical structures. Wang et al. [12] discussed the failure probability for an ethylene cracking furnace tube, which is a key component in the petrochemical industry. Zhou et al. [13] analysed the probability of failure of multi-storied reinforcement-concrete structures based on the evaluation of seismic control effect. Liang et al. [14] evaluated the failure rate of a power transformers and proposed a decision strategy for replacement. Seo et al. [15] studied the probability of failure model based on the safety factor in a water supply network; the determination of safety factor was related to the fundamental mechanism of corrosion in water pipes. Furthermore, based on the same fundamental pitting corrosion studies, Kioumarsi et al. [16] discussed the failure probability of a reinforced concrete beam.

Among the proposals of the RUL estimation and the PoF estimations, the following methods have proved to be the popular approaches. The Monte Carlo method is used in near-reality data simulations [12, 17], the neural network algorithm is applied to data processing [17, 18] and the Weiner process is used in remaining-life model building [19, 20]. Currently, the Bayesian Inference approach to update the old-information model with new-information model is not widely used, especially in engineering disciplines. Mosallam et.al [21] applied a Bayesian approach to the RUL estimation, but mainly focused on the algorithm of discrete Bayesian filtering. Wang et.al [20] applied the Bayesian Inference algorithm to model updating. However, with insufficient pre-processing of data, an over-simplified assumption, and the simplification in handling the Bayesian Inference function, the process of this algorithm is not suitable to provide the potential of general application in the engineering field. Further to the statistical approach of the PoF estimation, this research fully adapted the characteristic of the targeted asset and further proposed a novel graph-based PoF evaluation for the assets.

The structure of this paper is as follows: In the second section, the empirical-based PoF model is calculated from the existing model in the power supply industry [1, 22]. The new mechanism-based probability of failure is deduced from the life estimation model, which is concluded from previous research [3, 5]. After obtaining both probability of failure models, the Bayesian Inference method is introduced. Due to the complexity of the model, the conclusive results are obtained by the Markov Chain Monte Carlo (MCMC) algorithm, in which the independence sampling is done by the Metropolis-Hastings algorithm. Section three presented the results of the work and further proposes to establish a 'tailored probability of failure model' for each individual power transmission cable. This is a significant improvement comparing to the current industrial approach on the power cable management and risk assessment. All 'tailored models' are shown in the validation section of the paper.

2. Methodology

This section describes the main components of the proposed algorithm, shown as a flowchart in Figure 1. The algorithm is consisted of empirical model development, the new failure mode PoF model deduction, and the Bayesian inference method to conjunct the empirical PoF model with newly identified PoF models.

2.1 Empirical-based PoF model

The current PoF model in the power supply industry is based on the normal distribution which is shown in Figure 2a. Key assets of the industry - the power transmission cables are initially designed with an anticipated life. Due to the complexity of the operational environment, the anticipated life for underground power transmission cables is estimated with a normal distribution. The normal distribution approach is widely used in the power industry. The mean of the distribution represents the anticipated life for the majority of the cable population. The minimum anticipated life of the cable represents the earliest anticipated failure of cable. An early failure of the cable is due to either a heavy usage or the cable being used in an unfavourable environmental condition or an early life failure mode. The maximum anticipated life of the cable can be achieved by the assets operated in ideal conditions and therefore last longer than the majority of the population. The minimum anticipated life and the maximum anticipated life bonds the 95% confidence of cable life estimation distribution, which is four times the standard deviation of the distribution (4σ), as shown in Figure 2.

In the current industrial definition, cable conditions are assessed periodically and then labelled with replacement priorities. This replacement priority can be regarded as an RUL estimation based on the condition of the cables. The RUL is provided with a time window, this time window varies in association with the criticality conditions of the investigated cables.

Consider a normal distribution with the mean anticipated life of the power cable population as μ, with the probability density function as:

$$PDF_{normal} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} ------ (1)$$

This is the foundation of the empirical based model, with μ (the mean life of cable population) as the main parameter that distinguishes the PoF estimation for cables in service at different geographical locations.

2.2 A graphic-based PoF estimation

A corrosion fatigue mechanism-based model for cable life estimation is the foundation to deduce the graphic based PoF model. The following procedure is a summary:

2.2.1 Processing the cable section data

The data available for power transmission cable provided by our industrial partner consist of two types: the geological data and the oil pressure within the cable, with mean stress and alternating stress. The geological data includes the number of cable sections, the length of each cable section and the elevation of each cable section.

2.2.2 Obtaining the pit depth distribution

Basically, to estimate the life of a target cable, a section of the tin bronze tape must be obtained in order to measure the corrosion pit depths on its surface. Then applying the Monte Carlo simulation and transfer the experimental pit depths distribution into a theoretical GEV distribution [3]. Generally, the pit depth distribution of the targeting cable at any calendar year Y (e.g. current year is 2020), can be written in the format as:

$$F(x) = \left\{ \frac{1}{\sigma} [1 + \xi (\frac{x-\mu}{\sigma})]^{-\frac{\xi+1}{\xi}} e^{-[1+\xi(\frac{x-\mu}{\sigma})]^{-\frac{1}{\xi}}} \right\} \cdot (Y - t_c)^{0.33} ----(2)$$

Where t_c is the commission calendar year of the cable (begins operation), and μ, σ, ξ are location, scale and shape parameter of the pit depth distribution respectively. Equation (2) is a factorized distribution function.

2.2.3 Obtaining the pit-crack transfer pit depth

Existing work [5] introduced a pit-crack transfer probability function which can be expressed in the form of the cumulative distribution function of the Weibull distribution:

$$P(x) = 1 - e^{-(x/109.3)^{6.1}} ----- (3)$$

The pit depth distribution function (2) is then multiplied by this pit-crack transfer probability function (3). The pit depth with the highest probability from the asset function is considered as the transfer pit depth, naming it $x_{transfer}$.

2.2.4 Calculation of service time

From [5], the estimation of cable life can be obtained once the pit-crack transfer pit depth is known. The life estimation function is:

$$t_{cable} = \frac{{x_{transfer}}^{0.774}}{0.108 \times (\sigma_{mean} + \Delta\sigma)^{0.453}} ----- (4)$$

The advantage of this model is that, upon obtaining the transfer pit depth, the only additional information required for cable life estimation are the physics properties of the insulation oil pressure with mean stress σ_{mean} and the alternating stress $\Delta\sigma$.

2.2.5 Graphic-based PoF evaluation

As the focus of this section is to propose a graphic-based PoF evaluation for the power cables, a demonstration is presented with the assistance of Figure 3. The black solid curve plot is the estimated service life along the entire cable length based on the model of equation (4) with the natural spline algorithm for curve smoothing. The anticipated total length of failed cable at any service year is marked in red in Figure 3, as the anticipated life for these sections fall below the threshold. Assume the entire cable length is L, the life estimation function with the spline algorithm is noted as $l(d)$, here d is the length of the cable section. The PoF of the cable at time t is expressed as:

$$PoF(t) = \frac{\sum_{l(d)\leq t} l(d)}{L} - - - - - (5)$$

The solution of equation (5) relies on solving the equation at time t for $l(d) = t$, obtaining the solution of the variable d from the piecewise spline function, in order to assemble the sum of the failure cable length against the total cable length.

2.3 PoF updating with Bayesian inference algorithm

Two types of PoF models are introduced in this paper, the empirical-based model and the mechanism-based model. Both models are continuous functions. The Bayesian inference algorithm is proposed on the integration of the two types of PoFs, in order to form the comprehensive PoF model containing both the existing known knowledge of the engineering assets and the newly identified and learnt knowledge on such assets. The foundation of the model conjunction is the Bayes theorem. Consider one empirical PoF model P_E and a number of new PoF estimations $P_{N_j}, (j = 1,2, \ldots, n)$, based on n newly identified failure modes and associated failure mechanisms. The Bayesian inference base model considering the continuous functions is expressed as [23]:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)} - - - - - -(6)$$

Where

$$p(y) = \int p(\theta)p(y|\theta)\, d\theta - - - - - (7)$$

Substituting function (6) into (5), it can be concluded that under the condition of continuous function, the base Bayesian inference model is:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta)\, d\theta} - - - - - (8)$$

Further explanation of each term from equation (8) is given here:

- $p(\theta)$ represents the probability density function derived from the probability of failure function, concluded from the empirical model in the industry. It is a prior probability where the evidence of data was not observed or not clear.
- $p(y|\theta)$ is the probability density function derived from the probability of failure function, concluded from the mechanism-based model. This conditional probability can be interpreted as follows: given the documented critical underground power cable information θ, the novel probability with new observed information y (specifically, the up to date research results in previous work [3, 5]).
- The rest of the terms can be calculated, the difficulty being to calculate the marginal likelihood $\int p(\theta)p(y|\theta)\, d\theta$, where it is an integration of a complex function. This complicated integration requires a set of algorithms and will be stated in the 'Results' section.

Extending the base model to n newly identified PoF model, the integrated PoF updating model is therefore:

$$P_u = \frac{\prod_{j=1}^{n} P_E \cdot P_{N_j}}{\int_{\underbrace{\Theta\times\Theta\times\ldots\times\Theta}_{n}} \prod_{j=1}^{n} P_E \cdot P_{N_j}\, d\theta} ----(9)$$

2.4 Solution of the n-fold Bayesian inference model

The solution of the model follows a looped Markov Chain Monte Carlo method, to simplify the expressions, from here, the PoF models are noted by:

$$\begin{cases} Empirical - based\ PoF\ model\ PDF = F(t) \\ Mechanism - based\ PoF\ Model\ PDF = G(t) \\ Updated\ PoF\ Model\ PDF = U(t) \end{cases}$$

The Bayesian inference base model stated in Section 2.3 therefore is written as:

$$U(t) = \frac{F(t)G(t)}{\int F(t)G(t)dt} ----- (10)$$

Model expressions $F(t)$ and $G(t)$ are known. The difficulty exists for the calculation of denominator in equation (1): with the multiplication of two functions, $F(t)$ being an exponential Function and $G(t)$ a rational function, the analytical solution is hard to obtain. Instead, a numerical algorithm is applied to solve the problem stated in the following sub-sections.

2.4.1 Markov Chain Monte Carlo (MCMC) method

In this section, the purpose of the algorithm is to calculate the value of the denominator in equation (10):

$$D = \int F(t)G(t)dt - - - - - (11)$$

Under the strong law of large numbers (LLN) and the central limit theorem (CLT), when the numbers of sampling from the domain of t is very large and fulfils the requirement of independent sampling from function $F(t)$, then the value of the integration is approximately the expectation of function $G(t)$, which can be expressed as:

$$D = \int F(t)G(t)dt = E[G(t)] \approx \frac{1}{N} \cdot \sum_{i=1}^{N} G(t_i) - - - - - (12)$$

The algorithm applied here to withdraw independent sampling from the function $F(t)$ is the random walk Metropolis Hastings algorithm [24]. These steps were followed to construct the sequential independent sampling:

1) Initialise state $X_0 = [t_0]$ arbitrarily, the purpose is to construct a Markov chain containing states $X_0, X_1, X_2, \cdots, X_N \in \mathcal{X}$.
2) Let $x, x' \in \mathcal{X}$ be states in the Markov chain, propose an initial distribution $r(x'|x)$, in order to construct the Markov chain. The proposed distribution must satisfy detailed balance, where a factor $l(x'|x)$ can enable that

   $$r(x'|x)f(x)l(x'|x) = r(x|x')f(x') - - - -(13)$$

   In fact, this factor $l(x'|x)$ can be solved as:

   $$l(x'|x) = min\left\{1, \frac{r(x|x')f(x')}{r(x'|x)f(x)}\right\} - - - -(14)$$

   Define a factor in equation (14) as:

   $$\alpha = \frac{r(x|x')f(x')}{r(x'|x)f(x)} - - - -(15)$$

   α is the acceptance ratio for state transfer. The acceptance rate for a successful construction of Markov Chain is around 23.4% [25]. In this current algorithm, the proposed distribution $r(x'|x)$ is set to be a Gaussian distribution which is a symmetric distribution, only with this condition the acceptance ratio α from equation (15) can be simplified as:

$$\alpha = \frac{f(x')}{f(x)} - - - - (16)$$

3) Simulate a random number U from a uniform distribution $U(0,1)$, if $U < l(x'|x) = min\left\{1, \frac{f(x')}{f(x)}\right\} = min\{1, \alpha\}$, then transfer state x to state x', otherwise state x stays at x for the next round.
4) Repeat the steps from initial state $X_0 = [t_0]$ in Step 1) to Step 3) for N rounds, until the samplings become stable and convergent, these states $X_0, X_1, X_2, \cdots, X_N$ form a Markov chain independent sequential sampling from function $F(t)$.

With the above independent sequential samplings from $F(t)$, the integration result in equation (11) can be obtained. The algorithm provides a universal solution for the output on the n-fold conjuncted PoF model.

3. Results

Two working examples are provided in the results section. The two demonstration cable are subject to two different levels of criticality. Cable 1 is a critical cable with a relatively good condition, the timescale for replacing this cable is 5~10 years. Cable 2 is a critical cable with severe corrosion and requires immediate replacement, the timescale for replacing this cable is 0~2 years. These two assumed timescales for cable replacement will be used in the results section as numerical examples for the method and algorithms introduced in this research. The details of Cable 1 and Cable 2 with their assumed conditions are shown in Table 1.

The anticipated life of underground power transmission cable is estimated of between 30 to 70 years depending on the estimation of different power providers and research institutes [26-29]. To demonstrate the algorithm in this paper, the median of the anticipated life is taken as 50 years and the 4σ is bounded by 30 years and 70 years, as seen in Figure 2b. Two sets of geological profile data are assumed for both Cable 1 and Cable 2 to simulate the real situations. For Cable 1, the set of geological profile data is plotted in Figure 4, there are 57 sections of cable in Cable 1, the elevation and horizontal coordinates of the cables are hypothetical to simulate the field work scenario. From the plotted data, the length of each individual section of cable is calculated. All the sections of this cable service under a uniform alternating stress of $\sigma_{mean} = 0.69\ MPa$, but each individual section is different in operating mean stress $\sigma_{mean_i}, i = 1\sim57$. The data of mean stress against the length of cable is plotted in Figure 5. The discrete points in Figure 5 are the average mean stress values, the continuous function describing the relationship between mean stresses and cable length coordinates is simulated by natural cubic splines algorithm. For 57 sections of cable it requires 56 continuous piecewise cubic functions with a smoothing factor of $p = 6 \times 10^{-4}$, covering the entire length of the power cable.

Table 2 shows the information of all the key parameters that controls the corrosion pit depth distributions, the pit-crack transfer pit depths and the alternating stresses of the 2 critical cables. Following the plot in Figure 5, the continuous mean stresses functions are substituted in the life estimation model of equation (4), this iteration creates a further group of 56 piecewise life estimation functions and is plotted in Figure 6.

As discussed in the methodology section on the empirical-based model, the existing model in the industry extracted from [1, 22], is known to follow a normal distribution. Furthermore, the 2 cables being researched, Cable 1 and Cable 2, are assumed to be with a light and severe criticality respectively. When defining the PoF of the empirical-based model used in the industry, it can be explained with the assistance of Figure 2b as: with the increase of cable life, there is an increasing percentage of cable population entering the condition that a replacement is required, the amount is the shaded area in red in Figure 2b. This leads to the conclusion that the older the cables get, the fewer stays in the 'safe zone' which is the unshaded area under the normal distribution bell-shaped curve. Apply the graphic-based PoF estimation in the methodology section, the PoF at each year after the cable commission is defined as the percentage of the length of the estimated failed cable occupying the total length of the cable. In Figure 6, this definition is further explained with the assistance of 40 years and 50 years after cable commission to practically showing the results of the graphic-based PoF estimation on the real datasets. Draw a dashed line at year 40 in Figure 4, it cuts the estimated life plot into two parts, the plot above represents the cable with an estimated life of over 40 years. While the plot below represents the cable with an estimated life of less than 40 years, this part is within the shaded blue area. The projection of the estimated life plot on the horizontal line at 40 years is the length of the cable predicted of failure. The probability of failure at year 40 is then calculated as the total length of blue straight line over the length of the entire power cable.

$$PoF_{40} = \frac{\sum l_{blue}}{L_{Location\ 1}} ------ (24)$$

Here the l_{blue} represent of the length of all predicted failed cable parts at year 40, and $L_{Location\ 1}$ is the entire length of the cable at Cable 1.

With the same definition, at year 50, the predicted failed cable is below the horizontal dashed line at 50 years, under the shaded red areas. The total length of failed cable is the combination of horizontal red straight lines. The probability of failure is then calculated as:

$$PoF_{50} = \frac{\sum l_{red}}{L_{Location\ 1}} ------ (25)$$

l_{red} would be all the length of all predicted failed cable parts at year 50. It is easy to observe that, under this definition, the probability of failure for the cable increases from 0 at the minimum estimated life point of the plot until reaching 100% at the maximum estimated life point of the plot.

The graphic method outputs a set of discrete values representing the PoF value with the increase of service time *t*, the set of output values is noted as $\boldsymbol{p_t}$. It can be seen in Table 1 that, if the evaluated cable is showing obvious sign of deterioration, the cables with an assumed light criticality is with 5~10 years of replacement priority range. While the cables with an assumed severe criticality is with 0~2 years of replacement priority range. The replacement priority range can be interpreted as the upper and lower boundaries for the confidence of remaining life estimation. The lower boundary within the range is considered as a conservative estimation of the cable condition, while the upper boundary within the range is considered as a liberal estimation of the cable condition. With the above definitions, the PoF functions for Cable 1 and Cable 2 in both conservative and liberal estimations are given below. A representation of the PoF functions is shown in Figure 7.

Cable 1 is further taken as the working example to explain the detailed steps for posterior probability of failure calculation. The other location of cable used the same procedure of algorithm, these results are shown in the validation.

3.1 Probability density function of an empirical model

The conservative probability density function for severe criticality of cable evaluation:

$$PDF_{Severe\ Criticality_{conservative}} = \frac{1}{\sqrt{2\pi\sigma_{typical}^2}} e^{-\frac{(x-\mu_{typical})^2}{2\sigma_{typical}^2}} \quad ------(17)$$

The liberal probability density function for severe criticality of cable evaluation:

$$PDF_{Severe\ Criticality_{liberal}} = \frac{1}{\sqrt{2\pi\sigma_{typical}^2}} e^{-\frac{\left((x-2)-\mu_{typical}\right)^2}{2\sigma_{typical}^2}} \quad ------(18)$$

The conservative probability density function for light criticality of cable evaluation:

$$PDF_{Light\ Criticality_{conservative}} = \frac{1}{\sqrt{2\pi\sigma_{typical}^2}} e^{-\frac{\left((x-5)-\mu_{typical}\right)^2}{2\sigma_{typical}^2}} \quad ----(19)$$

The liberal probability density function for light criticality of cable evaluation:

$$PDF_{Light\ Criticality_{liberal}} = \frac{1}{\sqrt{2\pi\sigma_{typical}^2}} e^{-\frac{\left((x-10)-\mu_{typical}\right)^2}{2\sigma_{typical}^2}} \quad ------(20)$$

The conservative PoF density function is the lower limit of the RUL estimation, while the liberal PoF density function is the upper limit of the RUL estimation.

3.2 Probability density function of mechanism-based model

Following the methodology section with the definition of the mechanism -based model PoF estimation, a curve fitting procedure is used for the discrete set of PoF values $\boldsymbol{p_t}$ on obtaining one reasonable function that can describe the data while achieving a close fit. The fact that Weibull distribution is known for its application in failure analysis [30-33] (including mechanical failure analysis), it is also used to fit the probability of failure in this study.

Regarding the two locations as examples in this paper, with the data taken from previous work [3, 5].The fitted function is plotted in Figure 8 with red dashed line in comparison with the actual result. The statistical similarity between the actual result and the fitted function is provided in Table 3. The function for Cable 1 with service life t as variable is:

$$f_{Cable\ 1}(t) = 1 - e^{-\left(\frac{t}{53.52}\right)^{10}} -\!-\!-\!-\!-\ (21)$$

This procedure to deduct the probability of failure is also applied exactly to Cable 2, the result is:

$$f_{Cable\ 2}(t) = 1 - e^{-\left(\frac{t}{46.24}\right)^{15}} -\!-\!-\!-\!-\ (22)$$

The fitted function for probability of failure in Cable 1 is compared with both the conservative and liberal probability of failure function, also shown in Figure 8. Both functions have the format of the cumulative distribution function of the Weibull distribution. Furthermore, the PDF for both locations by mechanism-based model are:

$$\begin{cases} PDF_{Cable\ 1} = 0.187 \times \left(\frac{t}{53.52}\right)^{9} \times e^{-\left(\frac{t}{53.52}\right)^{10}} \\ PDF_{Cable\ 2} = 0.325 \times \left(\frac{t}{46.24}\right)^{14} \times e^{-\left(\frac{t}{46.24}\right)^{15}} \end{cases} -\!-\!-\!-(23)$$

Supplied with the working example from values of parameters in Cable 1, the pit-crack transfer pit depth is $x_{transfer} = 115\ \mu m$. The sequential samplings in the above steps corresponding to $F(t)$ is shown in Figure 9, which is an example of the samplings with 10,000 values, the red points are the accepted samplings for 1 trial. They are then all taken out and form one sequence of independent sampling which is also plotted in Figure 9. Because the independent samplings has a burn-in stage to achieve a stationary stage, it is hard to determine after which amount of values this stage will start [34]. This study takes a conservative approach: the first 10% of the accepted values of samplings are ignored, and the rest of the are imported to Function (21) as the values of t_i. The values of $G(t_i)$ determine the final results of the denominator D. Corresponding to the independent samplings t_i in Figure 9, the results of $G(t_i)$ are shown in Figure 10.

Due to the randomness of the Monte Carlo algorithm, the results appear to be different at small scale with each array of sequential samplings. To balance this randomness and achieve convincing and stable results, the above steps of samplings are carried out 1000 times in order to obtain an average value for the expectation value of $G(t_i)$; the result is: $D \approx 0.0287$.

3.3 Expression of the Bayesian inference updated model

Taking the result of the denominator, substitute the value of D into equation (10), the result is the Bayesian inference PDF. Shown as an example in Figure 11 is the plot of the Bayesian Inferenced probability density function for the liberal estimation. Apply another curve fitting which the result is a Gaussian distribution, the probability density function is written as:

$$U(t)_{Cable\ 1\ conservative} \approx 0.0559 \times e^{-\frac{(t-55)}{2\times 7.16^2}^2} -----(26)$$

This leads to the updated probability of failure function as:

$$PoF_{Update\ Cable\ 1\ conservative} \approx \int 0.0559 \times e^{-\frac{(t-55)}{2\times 7.16^2}^2}\, dt -----(27)$$

The calculation of the updated PoF function for liberal estimation in Cable 1, both the conservative and liberal updated PoF function for Cable 2 are shown in Table 4.

3.4 Discussion on the output result

The two plots shown here, Figure 12 and Figure 13, provide a comparison among the original conservative/liberal estimation and the updated conservative/liberal estimation for both Cable 1 and Cable 2. Take the service age of 65 years as an example, for both locations, with the cable reaching higher service life, the probability of failure increases rapidly and are reaching higher probability values compared to the existing model. While for the relatively short service age, these are of lower probability of failure compared to the existing model. This result shows the updated model after one round of new research-based information being input into the original model and it shows the development trend of such a statistical evaluation model with the learning process. The purpose to introduce this learning model is that, with new research carried out on such engineering assets, better understanding on newly identified failure mechanism are being developed. The new knowledge of the assets focus on specific components of a complex system, creating a PoF model of the specific component while influencing the reliability of the entire system. The method proposed in this research provides a blueprint to a route on integrating all the component PoF model into a complex system PoF model, by maintaining the existing knowledge on failure and reliability and introducing the new influential factors when new knowledge comes into the knowledgebase.

4. Conclusion

This paper introduced the Bayesian inference method in combining two probability of failure models for underground power transmission cables, one being the empirical model used in the power supply industry. Based on prior experience and knowledge, the other one being the mechanism-based model which was recently developed [3, 5] based on the corrosion fatigue mechanism on phosphor bronze reinforcing layers in cables. This combination enables a 'Tailored probability of failure' model for each cable locations. The 'Tailored probability of failure model' is of higher accuracy in estimating the probability of failure for two reasons:

- The updated model has a more complete background information compared to the previous model. With a deeper understanding of the failure mechanism, this model fits the phenomenon of failure cases better.
- The updated model is 'tailored' to separate locations. Instead of using a universal model for all locations, this model enables the distinction of the divergent conditions at the different cable locations and is thus more suitable for estimations.

The Bayesian inference method applied in this paper to solve engineering problems creates the concept of 'Intelligent Assets'. This concept is based on the philosophy that the aging and failure control is dynamic with the continuous updating of knowledge and information from the asset itself. This is the assets' 'self-learning'. With each input of the new information, the decisions on asset management become more accurate.

Acknowledgements:

This is a research work following on a project funded by the National Grid to the University of Leicester (Project number: NIA_NGET0103), which is gratefully acknowledged. Hang Zhou acknowledges financial supports for his PhD study by the Great Britain-China Educational Trust (GBCET) and the Henry Lester Trust.

References:

1. *Statistics: Overhead Lines and Underground Cables*, in *Aging of the System-Impact on Planning*. 2000, CIGRE: Paris, France.
2. Buhari, M., V. Levi, and S.K.E. Awadallah, *Modelling of Ageing Distribution Cable for Replacement Planning.* IEEE Transactions on Power Systems, 2016. **31**(5): p. 3996-4004.
3. Zhou, H., et al., *Measurement and Modeling of Pitting Depth Distribution for Phosphor Bronze Tapes Used in Underground Power Transmission Cables.* CORROSION, 2017. **73**(7): p. 844-852.
4. Zhou, H., *An engineering asset management approach to the evaluation of underground power transmission cables*, in *Department of Engineering*. 2018, University of Leicester: Leicester.
5. Zhou, H., et al., *Life prediction of phosphor bronze reinforcing tape used in underground power cables* CORROSION, 2018. **74**(5): p. 530-542.
6. Martin, D., et al., *An updated model to determine the life remaining of transformer insulation.* IEEE Transactions on Power Delivery, 2015. **30**(1): p. 395-402.
7. Khalifa, M., F. Khan, and J. Thorp, *Risk-based maintenance and remaining life assessment for gas turbines.* Journal of Quality in Maintenance Engineering, 2015. **21**(1): p. 100-111.
8. Segovia, M., et al., *Predicting remaining life of transmission tower steelwork components*, in *Risk, Reliability and Safety*. 2016, CRC Press: London, UK.
9. Ahmadzadeh, F. and J. Lundberg, *Remaining useful life prediction of grinding mill liners using an artificial neural network.* Minerals Engineering, 2013. **53**(Supplement C): p. 1-8.
10. Animah, I. and M. Shafiee, *Condition assessment, remaining useful life prediction and life extension decision making for offshore oil and gas assets.* Journal of Loss Prevention in the Process Industries, 2017.
11. Su, Y., et al., *A one-dimensional integral approach to calculating the failure probability of geotechnical engineering structures.* Computers and Geotechnics, 2017. **90**(Supplement C): p. 85-95.
12. Wang, W., et al., *Comparative analysis of failure probability for ethylene cracking furnace tube using Monte Carlo and API RBI technology.* Engineering Failure Analysis, 2014. **45**(Supplement C): p. 278-282.
13. Zhou, J., W. Zhao, and W. Mao, *Least Favorable Probability of Failure for 5- and 10-story RC Frame Structures with Vertical Irregularities.* Journal of Earthquake Engineering, 2015. **19**(7): p. 1158-1180.
14. Liang, G., et al., *A transformer replacement decision method based on probabilty assessment of failure rate.* Energy and Power Engineering, 2017. **9**: p. 748-755.
15. Seo, J., et al., *A Study on the Probability of Failure Model Based on the Safety Factor for Risk Assessment in a Water Supply Network.* Procedia Engineering, 2015. **119**(Supplement C): p. 206-215.
16. Kioumarsi, M.M., et al., *The effect of interference of corrosion pits on the failure probability of a reinforced concrete beam.* Engineering Structures, 2016. **114**(Supplement C): p. 113-121.
17. Jamali, A., et al., *Probability of failure for uncertain control systems using neural networks and multi-objective uniform-diversity genetic algorithms (MUGA).* Engineering Applications of Artificial Intelligence, 2013. **26**(2): p. 714-723.

18. Qiao, L., J. Shi, and W. An, *An application of systemic prediction evaluation parameters for neural network remaining useful life predictions models.* 2015 IEEE Conference on Prognostics and Health Management (PHM), 2015: p. 1-4.
19. Le Son, K., et al., *Remaining useful life estimation based on stochastic deterioration models: A comparative study.* Reliability Engineering & System Safety, 2013. **112**(Supplement C): p. 165-175.
20. Wang, H.-W., T.-X. Xu, and W.-Y. Wang, *Remaining Life Prediction Based on Wiener Processes with ADT Prior Information.* Quality and Reliability Engineering International, 2016. **32**(3): p. 753-765.
21. A.Mosallam, K.Medjaher, and N.Zerhouni, *Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction.* Journal of Intelligent Manufacturing, 2016. **27**(5): p. 1037-1048.
22. Buhari, M. and S. Awadallah, *Modelling of Ageing Distribution Cable for Replacement Planning*. Vol. 1. 2015.
23. Gelman, A., et al., *Bayesian Data Analysis*, ed. F. Dominici, et al. 2014, New York: Taylor & Francis Group.
24. Chib, S. and E. Greenberg, *Understanding the Metropolis-Hastings Algorithm.* The American Statistician, 1995. **49**(4): p. 327-335.
25. Roberts, G.O. and J.S. Rosenthal, *General state space Markov chains and MCMC algorithms.* Probab. Surveys, 2004. **1**: p. 20-71.
26. *Critical components, Underground cables*. [cited 2018; Available from: https://www.hydro.mb.ca/projects/system_renewal/components.shtml.
27. *Undergrounding high voltage electricity transmission lines*. 2015; Available from: https://www.nationalgrid.com/sites/default/files/documents/39111-Undergrounding_high_voltage_electricity_transmission_lines_The_technical_issues_INT.pdf.
28. *Overview of the Potential for Undergrounding the Electricity Networks in Europe*. 2003, ICF Consulting: London, United Kindom.
29. *Transmission & Distribution Infrastructure*. 2014, Harris Williams & Co.
30. Lihou, D.A. and G.D. Spence, *Proper use of data with the Weibull distribution.* Journal of Loss Prevention in the Process Industries, 1988. **1**(2): p. 110-113.
31. Seal, C.K. and A.H. Sherry, *Weibull distribution of brittle failures in the transition region.* Procedia Structural Integrity, 2016. **2**: p. 1668-1675.
32. Thomason, J.L., *On the application of Weibull analysis to experimentally determined single fibre strength distributions.* Composites Science and Technology, 2013. **77**: p. 74-80.
33. Gourier, C., et al., *Mechanical analysis of elementary flax fibre tensile properties after different thermal cycles.* Composites Part A: Applied Science and Manufacturing, 2014. **64**: p. 159-166.
34. Hobert, J.P. and C.P. Robert, *A Mixture Representation of π with Applications in Markov Chain Monte Carlo and Perfect Sampling.* The Annals of Applied Probability, 2004. **14**(3): p. 1295-1305.

Author Contribution Statement:

All four authors contributed to the completion of this paper. Hang Zhou developed and implemented the mathematical algorithm, wrote the main manuscript text and produced all the results presented in the paper. Michelle Le Blanc provided the industry perspective for this work. Jingzhe Pan and Fan Li offered the fundamental ideas that were used in this paper. To quantify the authors’ contribution towards the completion of this study, this can be estimated as: Hang Zhou 60%, Fan Li 10%, Michelle Le Blanc 10% and Jingzhe Pan 20%.

Competing Interests Statement:

The authors of this paper declare no competing interests both financially and non-financially.

List of tables:

Table 1: Summary of existing cable data for this research

Location	Criticality	Estimated Remaining Life
Cable 1	Light	5~10
Cable 2	Severe	0~2

Table 2: Location parameter μ for cables and corresponding transfer pit depths

Location	Pit depth distribution parameters	Transfer pit depth (μm)	Alternating Stress (MPa)
Cable 1	σ = 0.5, μ = 1.0, k = 0.5	115	0.69
Cable 2	σ = 0.5, μ = 2.2, k = 0.5	115	1

Table 3: Statistical evaluation of similarity between the actual probability of failure and the fitted function describing the probability of failure for the mechanical-based model

Location	Evaluation		Function
Cable 1	$R^2 =$	0.9948	$f(t) = 1 - e^{-\left(\frac{x}{53.52}\right)^{10}}$
	$SSE =$	0.3271	
Cable 2	$R^2 =$	0.9947	$f(t) = 1 - e^{-\left(\frac{x}{46.24}\right)^{15}}$
	$SSE =$	0.2298	

Table 4: 'Tailored probability of failure function' for all critical locations

Location	'Tailored probability of failure function'
Cable 1 (Conservative)	$PoF_{Update\ Cable\ 1\ conservative} \approx \int 0.0559 \times e^{-\frac{(t-55)}{2\times7.16^2}^2}\ dt$
Cable 1 (Liberal)	$PoF_{Update\ Cable\ 1\ liberal} \approx \int 0.0616 \times e^{-\frac{(t-60)}{2\times6.49^2}^2}\ dt$
Cable 2 (Conservative)	$PoF_{Update\ Cable\ 2\ conservative} \approx \int 0.0548 \times e^{-\frac{(t-50)^2}{2\times7.3^2}}\ dt$
Cable 2(Liberal)	$PoF_{Update\ Cable\ 2\ liberal} \approx \int 0.0571 \times e^{-\frac{(t-52)^2}{2\times7^2}}\ dt$

List of figures:

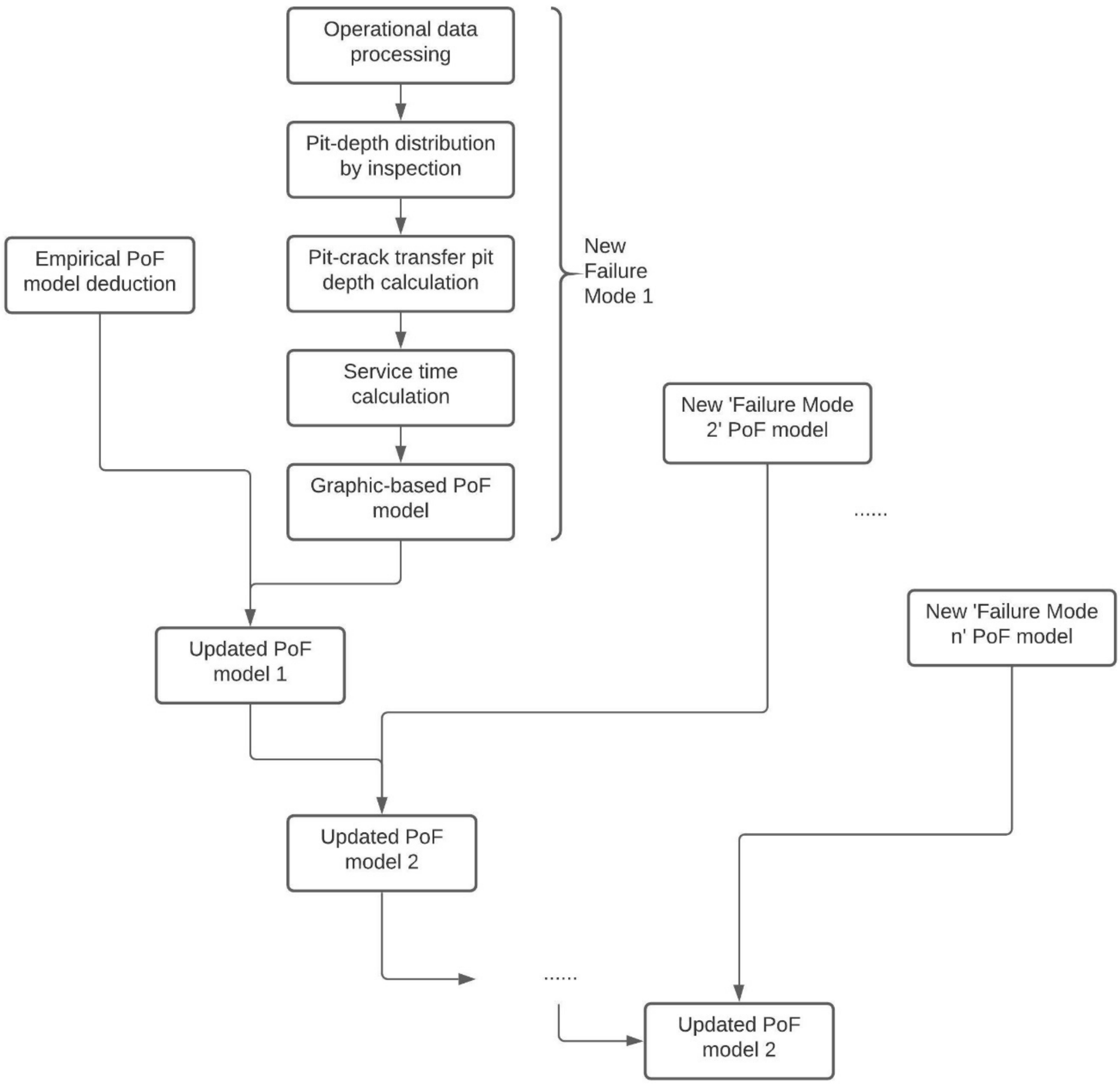

Figure 1: Flowchart of Bayesian inference algorithm for PoF model updating

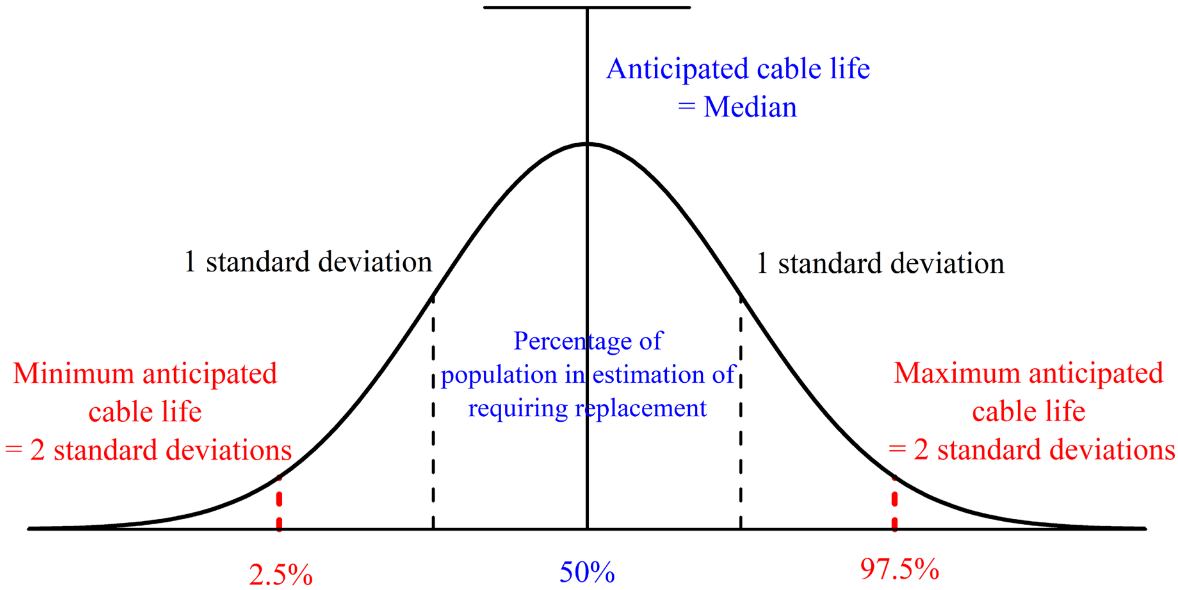

Figure 2a: Pictorial representation of the cable replacement model in the power supplement industry

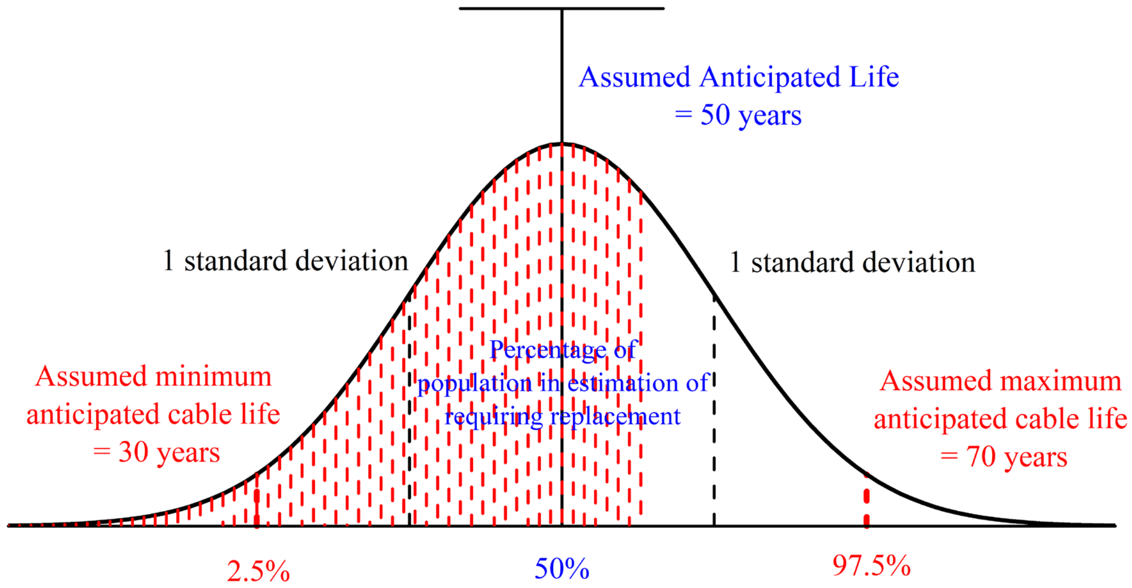

Figure 2b: Assumed most endangered cable with the life anticipation distribution

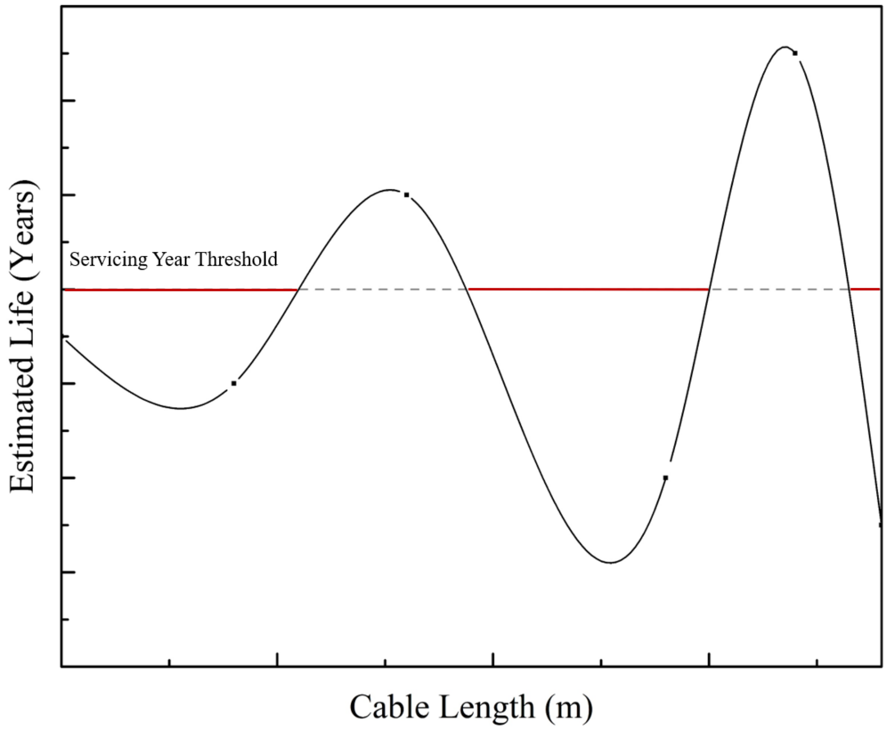

Figure 3: Demonstration graphic-based PoF estimation for power cables

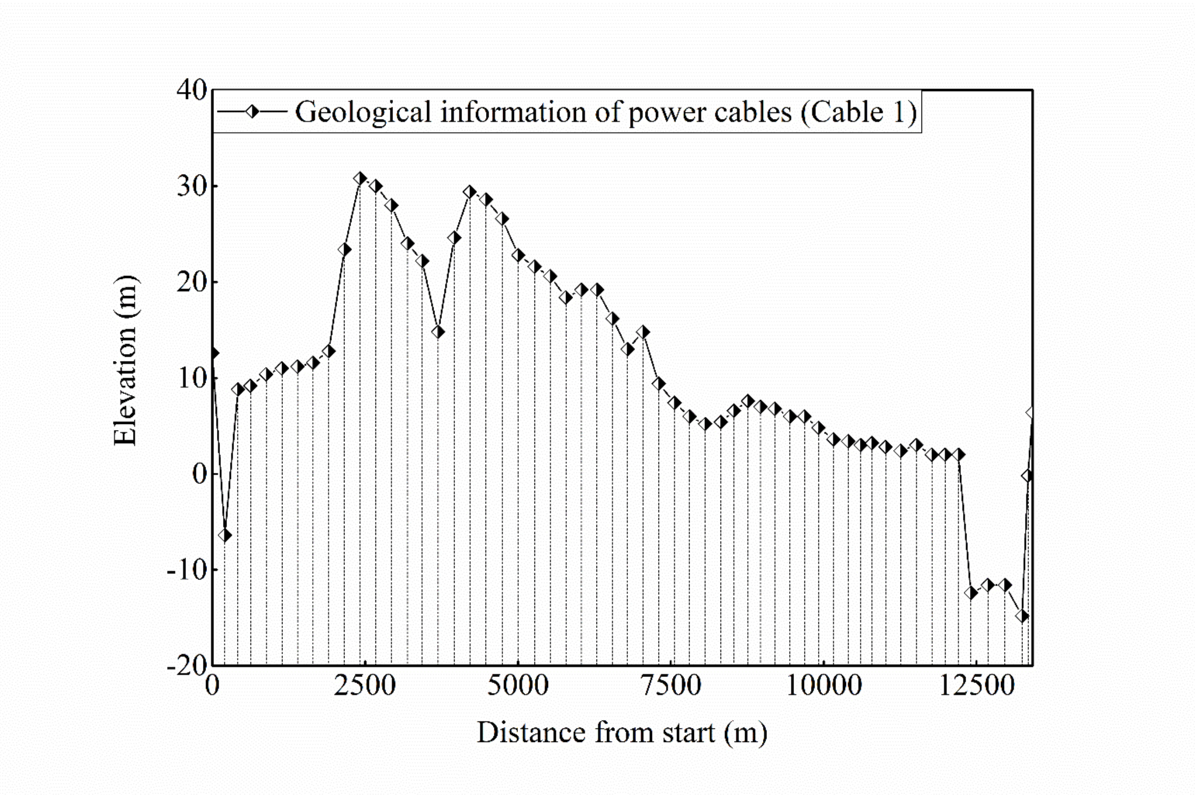

Figure 4: Geological data for Cable 1

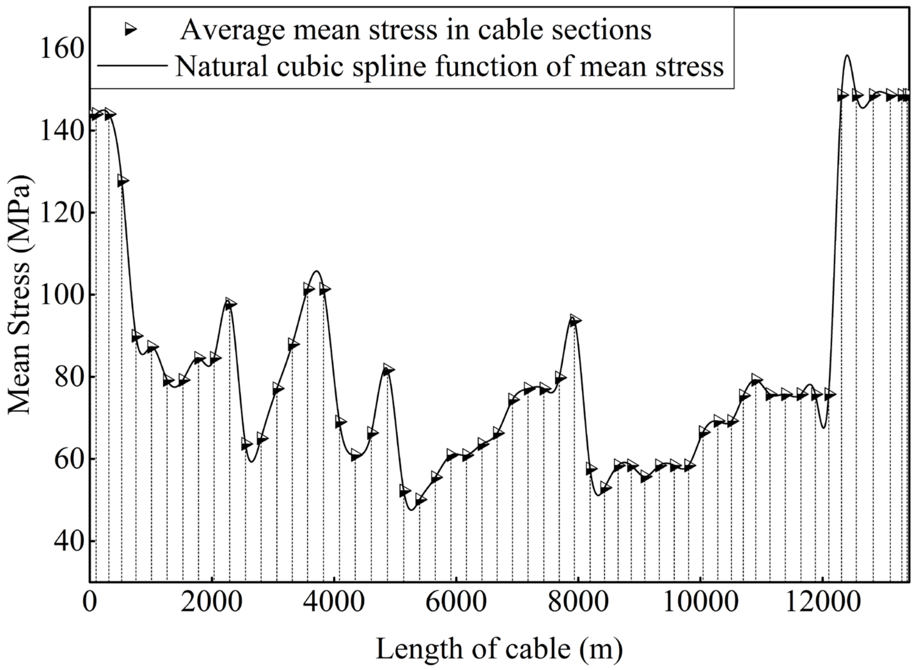

Figure 5: Mean stress vs Length of cable for Cable 1 (curves connecting mean stresses are simulated by natural cubic splines algorithm)

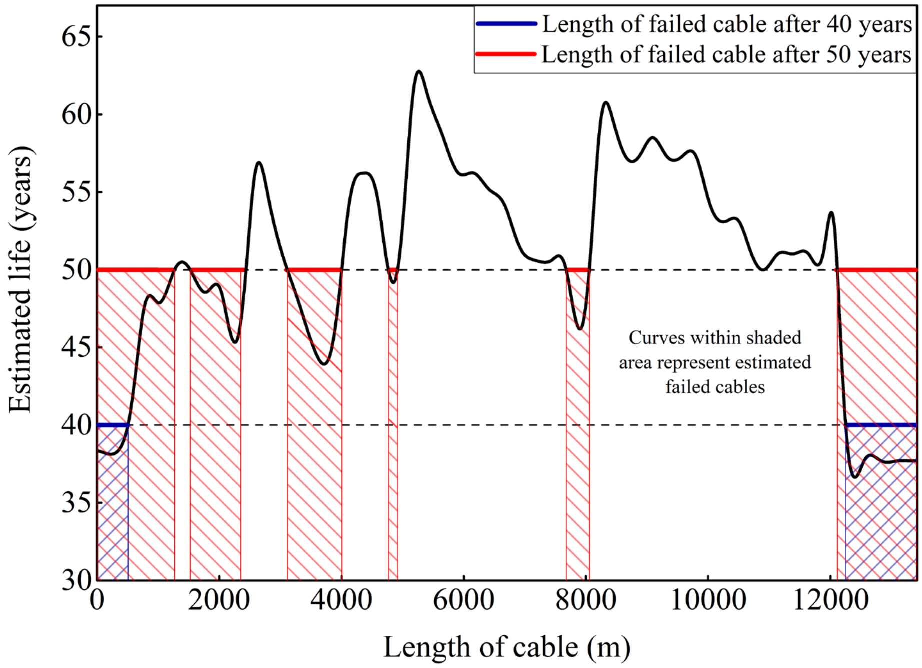

Figure 6: Estimated life vs Length of cable for Cable 1 with representation for the definition of the probability of failure

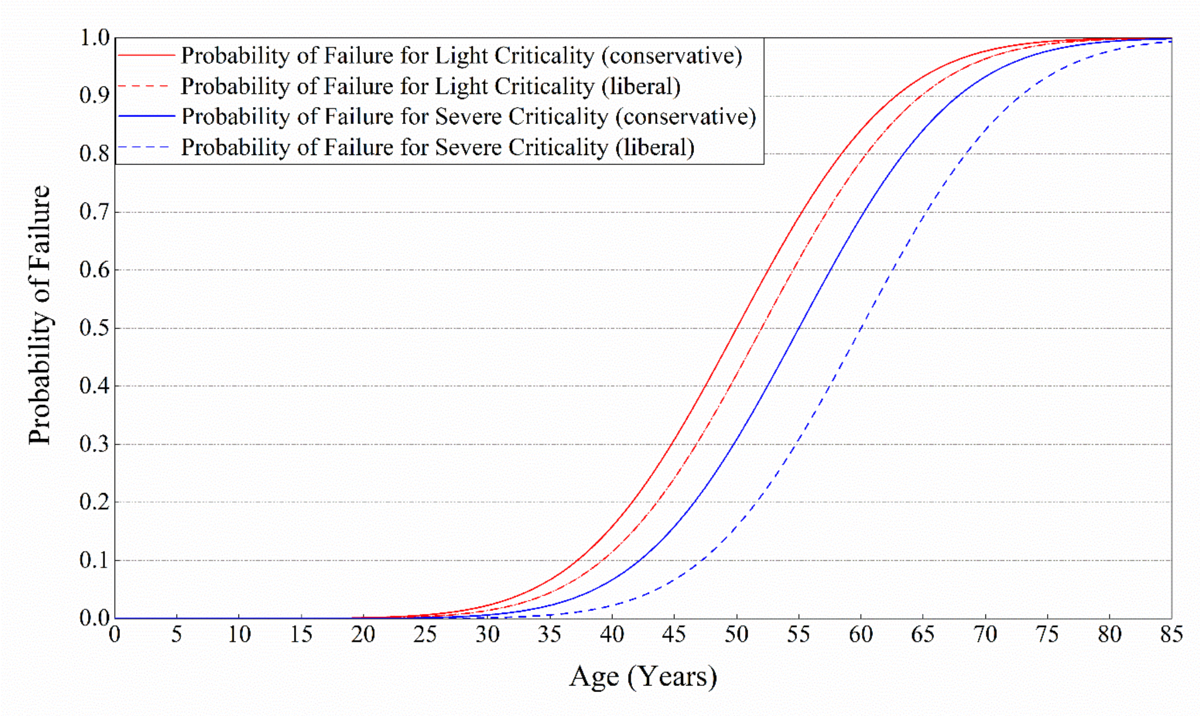

Figure 7: Empirical model on probability of failure for cables with light and severe criticality evaluation interpreted from industry

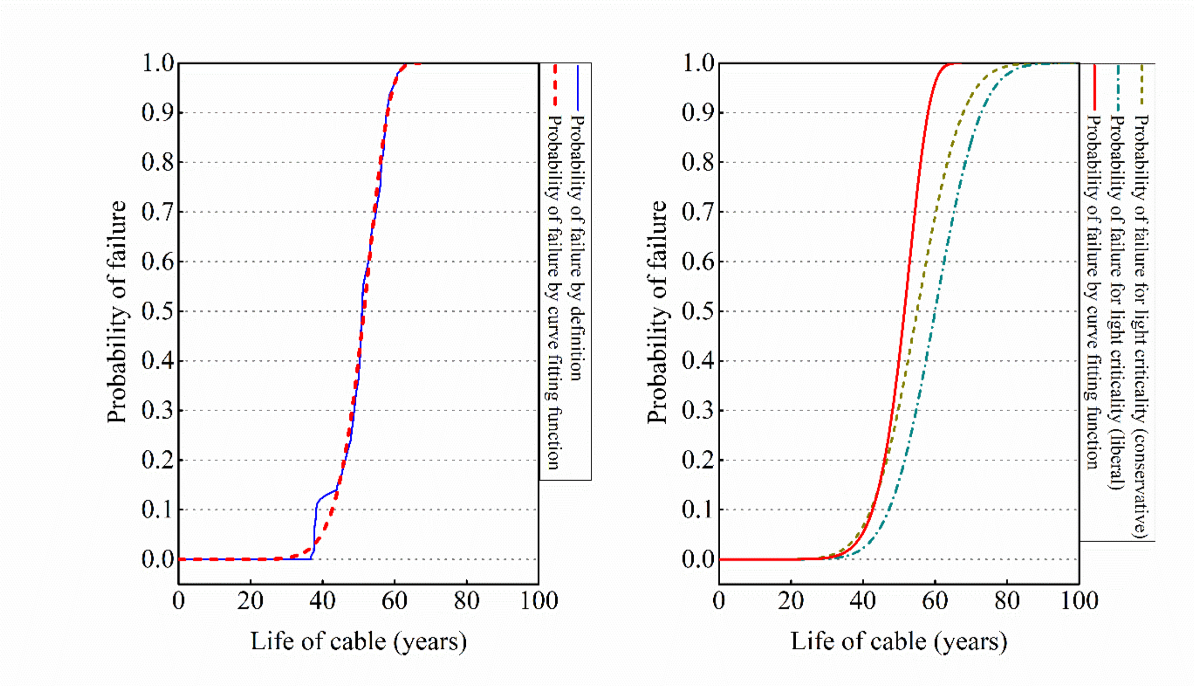

Figure 8: The probability of failure for Cable 1 under the definition from the mechanism-based model (Left)

The fitted function comparing to the conservative and liberal probability of failure function by empirical-based model (Right)

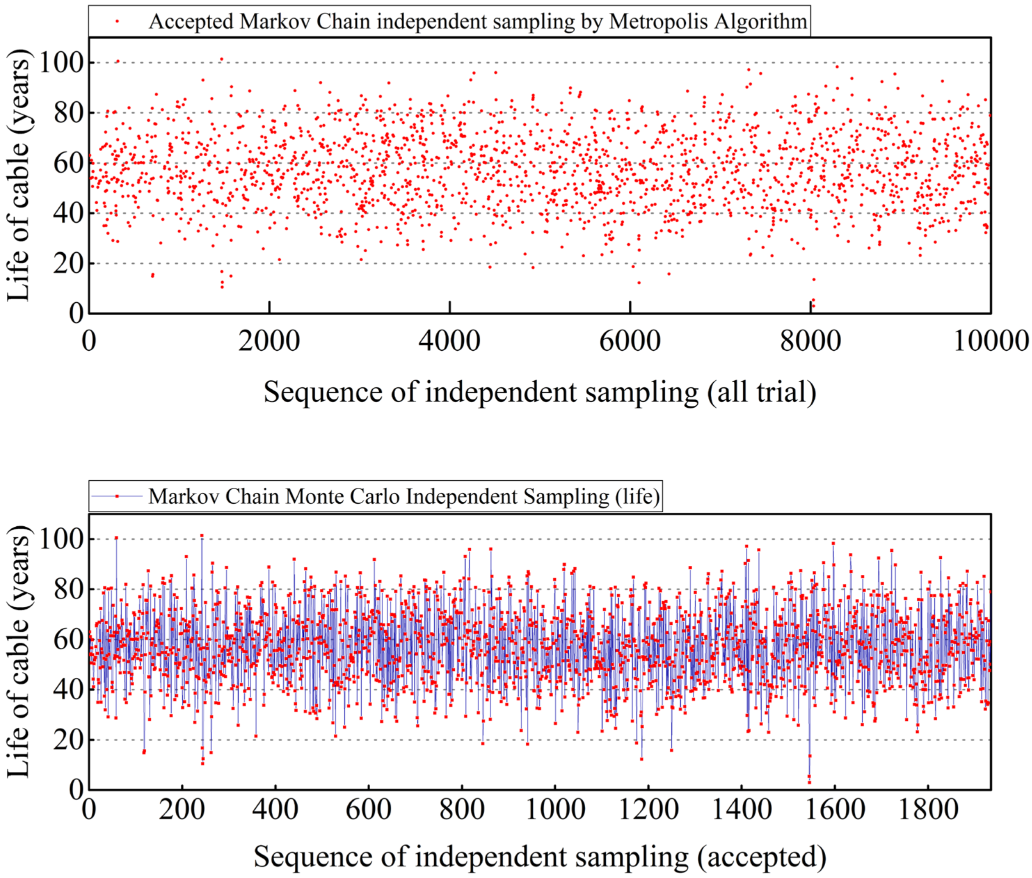

Figure 9: One trial of Markov Chain construction by Metropolis Hastings algorithm sampling with trial elements of 10000, among which only the accepted values are plotted (upper)

Fluctuation of all the accepted values (bottom)

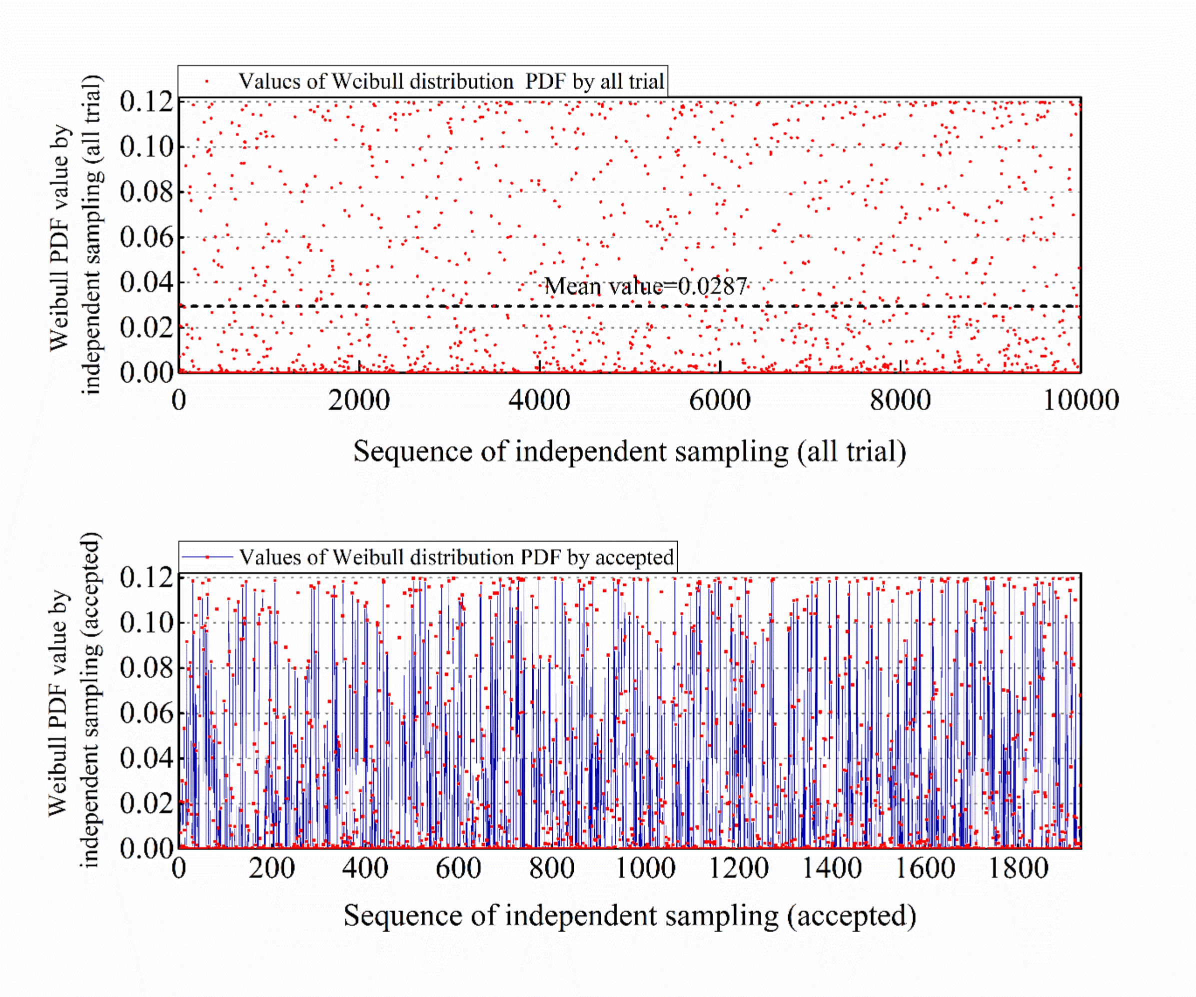

Figure 10: Values of $G(t_i)$ corresponding to the samplings shown in Figure 7

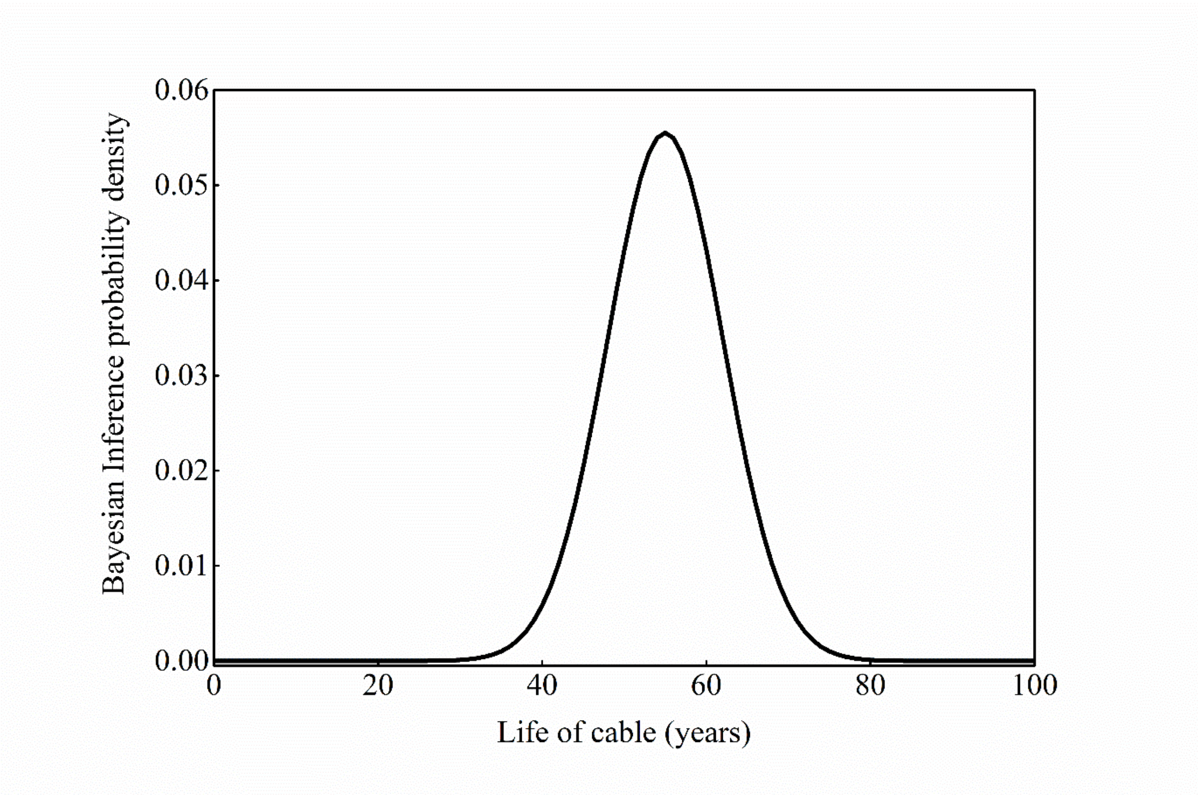

Figure 11: The plot of the Bayesian Inferenced probability density function of the liberal estimated model of Cable 1

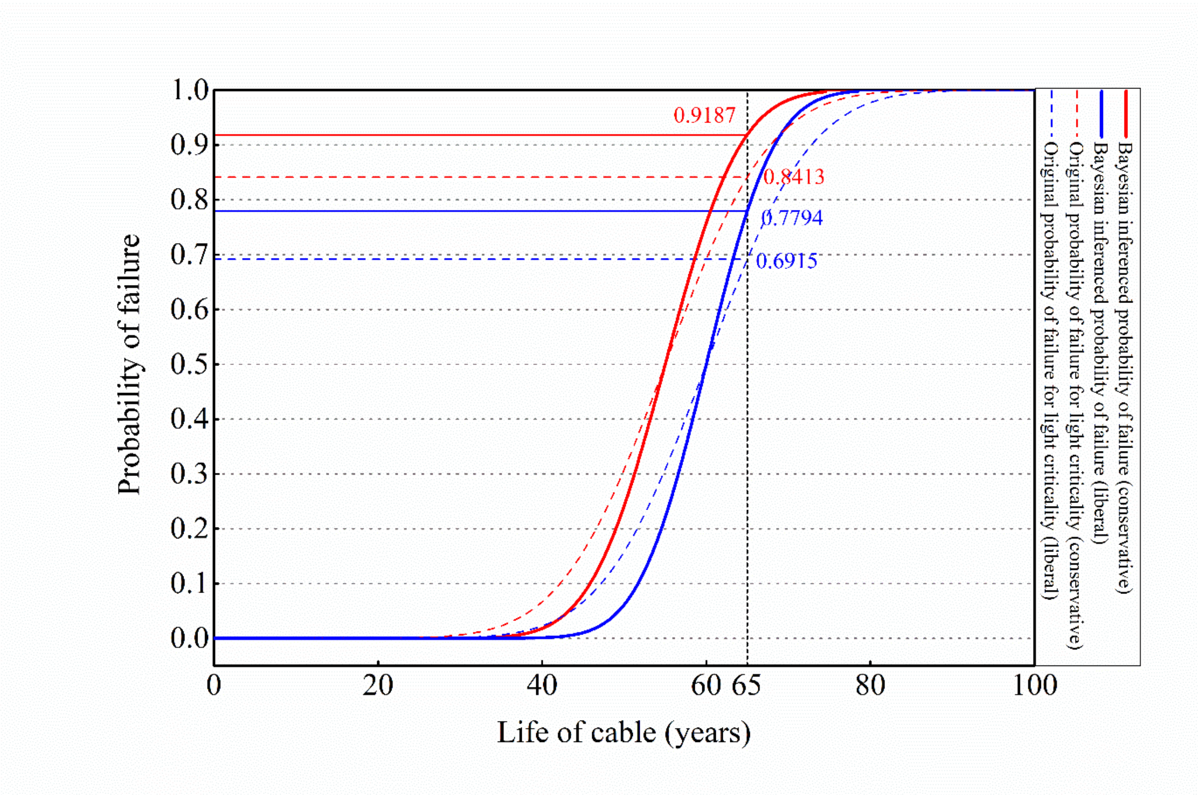

Figure 12: Comparison of Bayesian Inferenced model and original industrial replacement priority model for probability of failure estimation (both conservative and liberal of Cable 1)

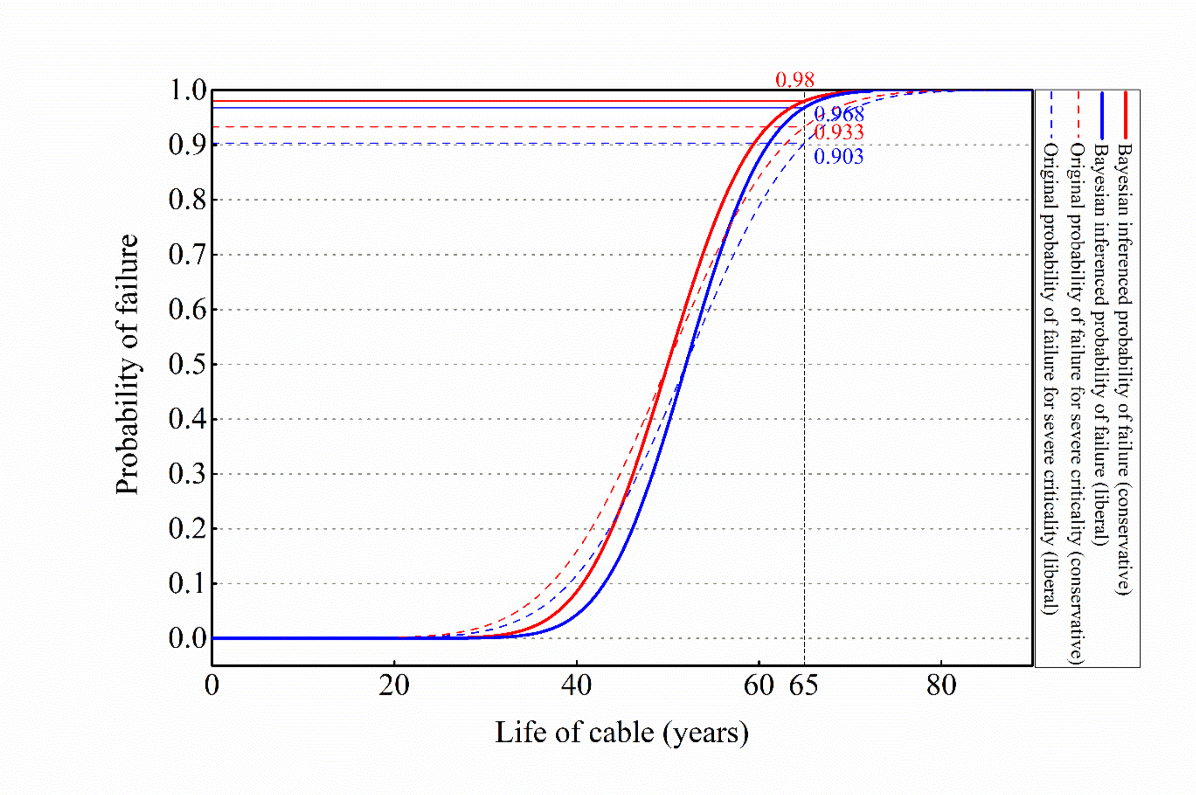

Figure 13: Comparison of Bayesian Inferenced model and original industrial replacement priority model for probability of failure estimation (both conservative and liberal of Cable 2)

Author Bio:

Hang Zhou holds a PhD degree and is currently a Postdoctoral Research Associate at the University of Cambridge. His research interests include system & reliability engineering, industrial risk assessment, and engineering asset management. His research has close ties with the industries. His clients and partners include Rolls Royce, BAE Systems, Siemens, National Grid, Port of Felixstowe, etc. from the industrial sectors of aerospace, transportation, energy and infrastructure.

Fan Li holds a PhD degree in mathematical modelling and has over 15 years industry experience. He is currently working as a digital transformation leader dedicating bespoke industrial solutions enabled by cutting edge technologies and practical approach. Fan is passionate about fit-for-purpose innovation and creating real-life social-economic impacts. Fan is renowned in UK infrastructure sector for his expertise and consistently delivering the highest quality projects. Fan has built and maintained long-term relationships with high-profile clients including National Grid UK, Network Rail, Ofgem (UK), European Commission, as well as a number of UK universities.

Michelle Le Blanc BEng CEng MIET has worked for National Grid since 1997, predominantly working in the field of management of electricity transmission assets, specialising in understanding the mechanisms that contribute to equipment end of life and asset replacement planning. She is a chartered engineer.

Jingzhe Pan is a Professor in Mechanics of Materials at the School of Engineering of the University of Leicester since 2006. His main research interest is mathematical modelling of material behaviours such as degradation, corrosion, sintering and creep of materials. He is a fellow of IoM3.