On Estimation of Equipment Failures in Electric Distribution Systems Using Bayesian Inference

Aaqib Peerzada
Texas A&M University
College Station, TX
peerzada@tamu.edu

Miroslav Begovic
Texas A&M University
College Station, TX
begovic@ece.tamu.edu

Wesam Rohouma
REAPERLab
Texas A&M-Qatar
wesam.rohouma@qatar.tamu.edu

Robert Balog
REAPERLab
Texas A&M-Qatar
robert.balog@qatar.tamu.edu

Abstract

This paper presents a new statistical parametric model to predict the times-to-failure of broad classes of identical devices such as on-load tap changers, switched capacitors, breakers, etc. A two-parameter Weibull distribution with scale parameter given by the inverse power law is employed to model the survivor functions and hazard rates of on-load tap changers. The resulting three-parameter distribution, referred to as IPL-Weibull, is flexible enough to assume right, left, and even symmetrical modal distribution. In this work, we propose an inferential method based on Bayes' rule to derive the point estimates of model parameters from the past right-censored failure data. Using the Monte Carlo integration technique, it is possible to obtain such parameter estimates with high accuracy.

1. Introduction

For electric utilities, the problems of accurately identifying the service times of equipment and scheduling preventive maintenance are of critical importance as answers to these questions represent substantial savings to the utility [1],[2]. The service life of power equipment is affected by several factors that include but are not limited to insulation strength, thermal and non-thermal stress, moisture, etc. Many factors that impact the lifetime of a device are not practical for monitoring since failure databases with a complete list of all the failure-inducing factors rarely exist. This is further exacerbated by the fact that the impact of a number of failure-causing factors is neither well documented nor is the failure mechanism well understood. Given the limitations of the accurate logging of equipment failure, a probabilistic aging model that captures the most salient features of the aging process, is the most practical for predicting the times-to-failure.

In reliability literature, the degradation of equipment, in particular, degradation induced by the aging of the device, is consistent with the Weibull distribution [3],[4]. The density function and the hazard function of the Weibull distribution have many interesting properties. In particular, the hazard function can assume a variety of shapes. For most aging-related failures, the hazard increases with time, thus increasing the probability of failure given that the device has survived until the present time instant.

Previous work in this area focused either chiefly on failure models derived from simple distributions or entirely ignored the impact of external stress on device aging. For example, in [5], an exponential failure model characterized by a single parameter is presented. The authors utilize Bayesian learning to estimate failure-times based on historical failure data. The exponential distribution solves the problem of analytical tractability since it permits the use of a conjugate prior distribution for the parameter of interest. However, the exponential model lacks the flexibility that a Weibull distribution offers which makes it less suitable to model device lifetimes. In [6], a distribution based on Perks Hazard function is presented. The model, however, does not incorporate the impact of external stress on equipment aging.

In this work, we present a failure model that is obtained by combining the inverse power law and the Weibull distribution. This enables accurate modeling of non-thermal stress-related failures in devices like on-load tap changers given a high penetration of intermittent non-scheduled generation. Since the failure database is often only partially available, we assume that the failure data contains only the following information: year of installation, the total number of assets, censoring number, and retirement history without replacement. An inferential technique based on Bayes' rule is developed to obtain the point estimates of model parameters. This enables us to predict the future performance of the assets that have survived based on the past failure history of similar devices, given similar failure mechanisms.

2. Problem Statement

In this paper, we consider the problem of accelerated aging of devices like distribution

URI: https://hdl.handle.net/10125/70996
978-0-9981331-4-0
(CC BY-NC-ND 4.0)

HICSS

transformers equipped with load tap-changers and switchable capacitor banks, given a high penetration of intermittent non-scheduled generation. The variable generation interferes with the regular operation of the tap-changers, causing them to operate much more frequently, usually outside of the design limits. Due to the mechanical nature of the tap-changing devices, the increased frequency of operation leads to accelerated degradation of the device, which often results in premature failure of the equipment.

Consider a substation transformer or a voltage regulator with $\lambda(t_i)$ as the tap-ratio at the time instant, t_i. Let us consider a planning horizon, τ with K number of discrete time instants. If h is a fixed time step, the number of tap operations in a discrete interval, $[t_i, t_{i-h}]$ with ΔV as the step change in voltage in per unit is

$$\delta_j = \frac{|\lambda(t_i) - \lambda(t_{i-h})|}{\Delta V}; i = 1, \ldots, \mathrm{K}$$

The cumulative tap operations over the planning period, τ is then given by

$$\zeta = \sum_{j=0}^{\mathrm{K}-1} \delta_j \tag{1}$$

The interaction of the intermittent non-scheduled generation with the on-load tap changers results in heavy operational stresses being imposed on the tap-changing devices. Since direct measurement of such time-dependent stress is difficult, if not entirely impossible, the change in the number of cumulative operations is a highly reliable indicator of such operational stress.

In this work, we formulate a relationship between the mechanical stress imposed on tap-changing equipment and the resultant change in the number of cumulative operations. We use this relationship to develop an inverse power law-Weibull failure probability model of on-load tap-changers (OLTCs) and switchable capacitor banks. The parametric failure model can be used to forecast the remaining useful life and probability of failure of equipment given a high penetration of non-scheduled generation. It is important to note that in this work, we only consider the mechanical stresses imposed on the OLTCs and switched capacitors.

If $p(t)$ is the probability density function of the time to failure, t of a device, then the probability of that device failing before time t is given by

$$F_T(t) = \Pr\{T \leq t\} = \int_0^t f(u)\, du \tag{2}$$

In the succeeding sections, we will examine the form of the function, $f(u)$. In general, a parametric failure model takes the form [7]

$$\mathfrak{F} = \{f(t; \bar{\theta}) : \bar{\theta} \in \Theta, \Theta \subset \mathbb{R}^k\} \tag{3}$$

In equation (3), $\bar{\theta}$ is the vector of model parameters that can take values in the parameter space, Θ . The problem then reduces to one of estimating the parameters that characterize the failure model.

3. Stress Ratio Factor

The development of the stress ratio factor is predicated on the understanding that an OLTC changes taps under the application of a force and hence stress on the contacts. Tap failure can either happen due to the asynchronous operation of the switches, usually caused by a broken axis or due to the carbon formation and oxidation of contacts. While a broken axis may be a sudden event, the carbon formation on the contacts represents gradual aging, exacerbated by the intermittent non-scheduled generation [8].

The carbon formation and oxidation of contacts are direct results of operational stresses imposed by the varying power flow conditions. It is possible to encode the information about the stresses and hence the gradual wear and tear of OLTCs in terms of the number of cumulative tap operations over the length of the planning period. If m is the force (=stress) imposed on the contacts of the tap mechanism per tap operation, we can write for the total stress over the device lifetime, assuming ζ cumulative operations

$$M_0 = \zeta m \tag{4}$$

where, M_0 represents the total baseline mechanical stress over the device lifetime. The baseline stress is indicative of wear and tear of the device under normal conditions when accelerated aging of the device can be ignored. If ζ_{PV} is the cumulative operations of the tap-changer in the presence of solar generation, then

$$M_{PV} = \zeta_{PV} m \tag{5}$$

The stress ratio factor is

$$\gamma = \frac{M_{PV}}{M_0} = \frac{\zeta_{PV}}{\zeta} \tag{6}$$

At time instant, t, the stress ration factor as a function of time is

$$\gamma(t) = \frac{\zeta_{PV}(t)}{\zeta(t)} \tag{7}$$

where, $\zeta_{PV}(t)$ is the cumulative tap operations till time instant, t given a high penetration of solar generation and $\zeta(t)$is the cumulative number of operations till the time, t in the absence of solar generation.

4. Parametric Aging Model

The accurate prediction of equipment failure due to accelerated aging is usually a matter of critical importance in asset management. Since it is not completely known when an equipment will fail in the foreseeable future, the service life of a newly-installed or an in-service equipment must only be a prediction.

However, if the service history (failure times) of a similar population is considered, then the error in the prediction can be minimized. Due to a scarcity of statistical information about the failure rates of power equipment, it becomes necessary to resort to probabilistic parametric models to derive an estimate about imminent and long-term failures. Note that we are restricting ourselves only to the failures caused by gradual or accelerated aging of equipment. In this work, we do not consider random failures due to external influences. Although the methods we present in this work have been applied to lifetime estimation of mechanical assets, like distribution transformers equipped with a load tap-changer mechanism and switched capacitors, the generality of the theory is preserved. As such, the parametric models developed here can be used to conduct failure estimation studies in any setting, where the gradual or accelerated aging due to the application of mechanical stress is the main cause of equipment failure.

In the case of distribution system transformers, aging-related failures contribute over 70% of the total failures, while random failures account for roughly 30% of the total failures [9]. Among the aging-related failures, the wearing out of On-load tap changers (OLTCs), has been singled out as a significant cause of transformer failure [10]. To address the problem, we consider the lifetime estimation of OLTCs and switched capacitor banks in a way that will help formulate replacement strategies.

The lifetime of a device under non-thermal stress has been shown to follow the inverse-power law (IPL). It is a parametric equation characterized by parameter n, referred to as the stress endurance coefficient. The IPL can be expressed mathematically as

$$L(M) = \left(\frac{M}{M_0}\right)^{-n} L_0 \quad (8)$$

where, $L(M)$ is the life as a function of applied stress M, L_0 is the device lifetime corresponding to baseline stress, M_0 and n is the stress-endurance coefficient. The baseline stress is the amount of stress below, which accelerated the aging of a device can be neglected.

Weibull distribution has been extensively used to model the distribution of time to failure. A random variable T has a three-parameter Weibull distribution with parameters α, η, and β, if its density function is given by [11]

$$f_T(t|\alpha,\eta,\beta) = \frac{\beta}{\eta}\left(\frac{t-\alpha}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t-\alpha}{\eta}\right)^{\beta}\right] \; ; t \geq \alpha \quad (9)$$

Where α is called the delay or minimum life in the context of T representing device lifetime. So, the support of $f_T(t|\alpha,\eta,\beta)$ is $t \geq \alpha$. For t being a duration which is nonnegative, and α the minimum duration, the domain of α does not encompass $\Re$ but rather a smaller interval $[0.\infty)$. The second parameter, η is called the characteristic life or scale parameter from a statistical point of view. It is called characteristic life for the reason that for the same α, η and varying β the cumulative density functions ($F_T(t)$), of all Weibull variates, intersect at a point with coordinates $t = \alpha + \eta$ and $F_T(\alpha+\eta|\alpha,\eta,\beta) \approx 0.6321$. In other words, the scale parameter, η is the time at which 63.2% population has failed. The domain of η is $(0,\infty)$ and it is measured in the same units as the random variable T. The third parameter, β is the Weibull-slope, called the form or shape parameter of the Weibull distribution. It is a dimensionless quantity with domain $(0,\infty)$.

To model the device lifetime under the application of stress, the inverse power law (8) can be combined with a two-parameter Weibull distribution. The two-parameter Weibull distribution is obtained by setting the delay parameter to zero in (9), since items typically start to fail after the age of $t = 0$. To get the modified IPL-Weibull distribution, we replace the scale parameter, η in (9) by $L(M)$, and set $\alpha = 0$. This results in the probability density function of the form as

$$f_T(t|\beta, L_0, n) = \begin{cases} \frac{\beta}{L_0^{\beta}}(t)^{\beta-1}\left(\frac{M}{M_0}\right)^{\beta n} \exp\left[-\left(\frac{M}{M_0}\right)^{\beta n}\left(\frac{t}{L_0}\right)^{\beta}\right]; t \geq 0 \\ 0 \qquad ; \; t < 0 \end{cases} \quad (10)$$

The ratio $\left(\frac{M}{M_0}\right)$ is the stress-ratio factor. Substituting (6), the density function with $\psi = [\beta, L_0, n]$; the set of model parameters, takes the form

$$f_T(t|\psi) = \begin{cases} \frac{\beta}{L_0^{\beta}} t^{\beta-1}\gamma^{\beta n} \exp\left(-\gamma^{\beta n}\frac{t^{\beta}}{L_0^{\beta}}\right); t \geq 0 \\ 0 \; ; \; t < 0 \end{cases} \quad (11)$$

5. Statistical Properties of IPL-Weibull

The probability density function of the IPL-Weibull distribution has several interesting statistical properties. The critical functions that completely specify the distribution of the random variable T, are hazard rate function, also known as failure rate function, survivor or reliability function, and the failure probability or the cumulative density function. In this section, we will examine the closed-form solutions of these functions. In addition, we will also derive the closed-form representations of the expectation, median and mode of the IPL-Weibull distribution.

5.1 Survivor Function or Reliability Function

The survivor function is the complement of the failure probability function. In simple terms, it is the probability of survival beyond time, t. Mathematically, it is given by

$$S_T(t) = \Pr\{T > t|\psi\} = \int_t^{\infty} f(u)\,du \qquad (12)$$

For the IPL-Weibull model, we can get the survivor function by substituting (10) in (12) and using integration by parts. The result evaluates to an upper incomplete gamma function which can be further solved using exponential integral transformation. This results in for the survivor function of the IPL model as

$$S_T(t) = \exp\left[-\gamma^{\beta n}\left(\frac{t}{L_0}\right)^{\beta}\right] \qquad (13)$$

5.2 Failure Probability Function

The survivor function and failure probability function are related by the expression

$$S_T(t) = 1 - F_T(t) \qquad (14)$$

Hence, the failure probability or cumulative density is given by

$$F_T(t) = 1 - \exp\left[-\gamma^{\beta n}\left(\frac{t}{L_0}\right)^{\beta}\right] \qquad (15)$$

5.2 Hazard Rate Function

The hazard function is technically not a probability measure but rather an assessment of risk. Hazard function can be thought of as the probability of failure in the small infinitesimal interval of time $[t, t + dt]$ given that the equipment has survived till the time, t. A typical hazard function usually comprises of three parts. The first part represents early failures and is characterized by a decreasing failure rate. The second part is representative of random failures. Random failures are caused by external influences and are independent of the aging of the equipment. The second part of the hazard function is thus characterized by a constant failure rate. The third part is indicative of wear-out failures caused by decreasing mechanical or electrical strength of the materials. These failures share a strong correlation with the aging of devices and are characterized by increasing failure rates. The work presented in this paper is focused on the third part of the bath-tub failure rate curve.

The Hazard function of the IPL-Weibull can be obtained from the definition of the hazard function

$$h_T(t) = \lim_{dt \to 0} \frac{\Pr[t < T \le t + dt | T > t]}{dt}$$

Using this definition, we can write for the Hazard function of IPL-Weibull

$$h_T(t) = \frac{f_T(t|\beta, L_0, n)}{S_T(t)} = \frac{\beta}{L_0}\left(\frac{t}{L_0}\right)^{\beta-1} \gamma^{\beta n}(t) \qquad (16)$$

For the IPL-Weibull, the hazard function depends on the parameter, β. With $\beta = 1$, the IPL-Weibull reduces to IPL-Exponential with a constant failure rate. This is given by the second part of the bath-tub curve characterized by random failures. With $\beta < 1$, the failure rate is decreasing and $\beta > 1$ represents an increasing failure rate. Thus, the hazard function of IPL-Weibull is very flexible and can assume a variety of forms, unlike a Gamma distribution whose density function is always right-modal. On the other hand, the Log-normal hazard function can only model a decreasing failure rate, which is inconsistent with aging-related failures.

5.3 Expectation, Median and Mode

The expectation or mean time to failure (MTTF) of an IPL-Weibull distributed random variable, T is given by

$$E[T] = L_o\gamma^{-n}\,\Gamma\left(1 + \frac{1}{\beta}\right) \qquad (17)$$

Where, $\Gamma\left(1 + \frac{1}{\beta}\right)$ is the gamma function evaluated at $\left(1 + \frac{1}{\beta}\right)$. The median of the IPL-Weibull random variable T can be found by setting, $F_T(t) = \frac{1}{2}$ and solving for t.

$$1 - \exp\left[-\gamma^{\beta n}\left(\frac{t}{L_0}\right)^{\beta}\right] = \frac{1}{2} \Longrightarrow \gamma^{\beta n}\left(\frac{t}{L_0}\right)^{\beta} = \ln(2)$$

This yields for the median

$$m = \frac{L_0}{\gamma^n}[\ln(2)]^{\frac{1}{\beta}} \qquad (18)$$

The mode of the distribution is the value of the argument at which the density function has a local maximum. This can be easily found by taking the derivative of the density function and setting it equal to zero. Hence, for the IPL-Weibull distributed random variable, T

$$\frac{df_T(t|\beta, L_0, n)}{dt} = \frac{\beta\gamma^{\beta n}}{L_0^{\beta}}\left[t^{\beta-1}\exp\left(-\gamma^{\beta n}\left(\frac{t}{L_0}\right)^{\beta}\right)\frac{-\gamma^{\beta n}}{L_0^{\beta}}\beta t^{\beta-1} + \exp\left(-\gamma^{\beta n}\left(\frac{t}{L_0}\right)^{\beta}\right)(\beta - 1)t^{\beta-2}\right] = 0$$

Upon simplification, this results in for the mode of the IPL-Weibull

$$f_T(t_{max}) = \frac{L_0}{\gamma^n}\left[1 - \frac{1}{\beta}\right]^{\frac{1}{\beta}} \qquad (19)$$

6. Bayesian Method Applied to Failure Estimation

The time-to-failure of a device can be predicted from the cumulative failure density function given by (15). However, such a prediction is only possible if the point estimates of the parameters that characterize the failure model are available. The Weibull reliability model with the scale parameter given by the inverse-power law is a three-parametric failure distribution model. The model parameters are the shape parameter, β, the scale parameter, L_0 and the stress parameter, n. To get the point estimates of the model parameters, we propose the use of Bayesian inference since the Bayesian method allows for the incorporation of expert knowledge on the device wear and tear in terms of the prior distribution of the model parameters. Although the degree of subjectivity involved in selecting the prior distribution has often led the proponents of the frequentist statistics to criticize the theory as lacking objectivity. However, in reliability studies, the ability to incorporate subjective knowledge in a failure model is a desirable feature. Also, the use of prior information can help mitigate the effect of a small sample size. In that regard, Bayesian inference is a preferred failure estimation tool of power equipment since the failure data of distribution transformers with on-load tap changers are usually very scarce and not readily available. Besides, Bayesian inference provides a more intuitive interpretation of the results in terms of probabilities that satisfy the likelihood principle.

Bayesian inference is based on the rule of conditional probability, also known as the Bayes rule. If $\boldsymbol{\mathcal{D}}$ is the data vector of failure times that are independent and identically distributed, and ψ is the set of model parameters, then the joint posterior distribution of ψ conditioned on the data in $\boldsymbol{\mathcal{D}}$ is given by the Bayes rule as

$$g(\psi|\boldsymbol{\mathcal{D}}) = \frac{f(\boldsymbol{\mathcal{D}}|\psi)\pi(\psi)}{f(\boldsymbol{\mathcal{D}})} \tag{20}$$

Where the marginal probability distribution, $f(\boldsymbol{\mathcal{D}})$ of the failure data, $\boldsymbol{\mathcal{D}}$ is given by

$$f(\boldsymbol{\mathcal{D}}) = \int f(\boldsymbol{\mathcal{D}}|\psi)\pi(\psi)d\psi \tag{21}$$

Since the marginal probability distribution, $f(\boldsymbol{\mathcal{D}})$ is obtained by integrating out all the model parameters, it is often treated as a normalization constant and (20) reduces to

$$g(\psi|\boldsymbol{\mathcal{D}}) \propto f(\boldsymbol{\mathcal{D}}|\psi)\pi(\psi) \tag{22}$$

The failure times in $\boldsymbol{\mathcal{D}}$ are assumed to be a sample from a multi-parameter failure probability model. Assuming we have N identical devices with similar failure mechanism, the vector $\boldsymbol{\mathcal{D}}$ is a collection of failure times of N devices.

$$\boldsymbol{\mathcal{D}} = \{t_i | t_i \leq C\} \;\; i = 1,2,\dots,N \tag{23}$$

Where t_i is the failure/retirement time of ith device, and C is the present time instant. A data set given by (23) is a complete data set with no censoring. In the absence of censoring, the failure model is represented by the likelihood function

$$f(\boldsymbol{\mathcal{D}}|\psi) = \mathcal{L}(\psi;\boldsymbol{\mathcal{D}}) = \prod_{i=1}^{N} f(t_i|\psi) \tag{24}$$

In (24), $f(t_i|\psi)$ is the failure density function of t_i given ψ. The Bayesian method is concerned with inferring the properties of ψ; the set of model parameters, based on the data in $\boldsymbol{\mathcal{D}}$. In that sense, the Bayesian method treats the model parameters as random variables with a joint prior distribution given by $\pi(\psi)$.

However, the actual observed failure data of power equipment is rarely complete and is almost always censored. In the case of distribution transformers with on-load tap changers, the observed failure data will have two sets; a set of retired/failed OLTCs and a set of in-service OLTCs. Since the test duration is given by the fixed number of r failures of the N number of assets, such a failure data set is referred to as Type-II censored, and the censoring is on the right. With right-censored data, a specimen not censored by the present time instant has survived till the present time, whereas a subject that is censored has retired/failed by the current time instant. For right-censored data, the likelihood function has the general form[11]

$$\mathcal{L}(\psi|\boldsymbol{\mathcal{D}}) = \prod_{i\in\Omega} f(t_i) \prod_{j\in\mathcal{R}} S(t_j) \tag{25}$$

Where Ω is the set of failed/retirement times, and $\mathcal{R}$ is the set of censored times. Assuming N total assets with r of the N assets failed/retired by the present time instant, C the likelihood function of the failure model is

$$\mathcal{L}(\psi|\boldsymbol{\mathcal{D}},N,r) = \prod_{i=1}^{r} f(t_i|\psi) \prod_{j=r+1}^{N} \left[1 - F(s_j|\psi)\right] \tag{26}$$

The probability that j^{th} OLTC will last at least s_j years is

$$\Pr\left[C_j > s_j|\psi\right] = S(s_j|\psi) = 1 - F(s_j|\psi) \tag{27}$$

With the likelihood function given by (26), the conditional joint posterior distribution of the model parameters is

$$g(\psi|t_1,t_2,\dots,t_r,s_{r+1},\dots,s_N,C,N,r) = K\prod_{i=1}^{r} f(t_i|\psi) \prod_{j=r+1}^{N} \left[1 - F(s_j|\psi)\right] \pi(\psi) \tag{28}$$

Where $t_1, t_2, \ldots, t_r$ are the failure times of the devices that have failed/retired by the present time instant, $s_{r+1}, \ldots, s_N$ are the survival times of the remaining $N-r$ devices, C is the present time instant, N is the total number of assets, and r is the censoring number. The normalization constant K is

$$K^{-1} = \int \prod_{i=1}^{r} f(t_i|\psi) \prod_{j=r+1}^{N} \left[1 - F(s_j|\psi)\right] \pi(\psi) d\psi$$

6.1 Point Estimates of Model Parameters

In this work, we are interested in deriving the point estimates of the parameters that characterize the failure model. Consider a failure model, $f(t|\theta)$ parameterized by θ in the parameter space, Θ. In deriving a point estimate of the parameter, θ referred to as $\hat{\theta}$, the discrepancy between θ and $\hat{\theta}$ is measured by the loss function, $L(\theta, \hat{\theta})$. A loss function is a mapping from the parameter space to real space. To measure the risk associated with the point estimator, we consider a quadratic loss function [7]

$$L(\theta, \hat{\theta}) = \left(\theta - \hat{\theta}\right)^2 \quad (29)$$

The posterior risk of the estimator, $\hat{\theta}$ is given by

$$R(\hat{\theta}|\boldsymbol{\mathcal{D}}) = \int L(\theta, \hat{\theta}) f(\theta|\boldsymbol{\mathcal{D}}) d\theta \quad (30)$$

Where $f(\theta|\boldsymbol{\mathcal{D}})$ is the marginal conditional posterior distribution of the parameter, θ given the data in $\boldsymbol{\mathcal{D}}$. With a squared loss function, the point estimate or the Bayes estimator, $\hat{\theta}$ of the parameter, θ is the expectation of θ.

$$\hat{\theta}(\boldsymbol{\mathcal{D}}) = \int \theta f(\theta|\boldsymbol{\mathcal{D}}) d\theta = \mathrm{E}(\theta|\boldsymbol{\mathcal{D}}) \quad (31)$$

6.2 Bayes Estimators of IPL Weibull Model

The Bayes estimators of the IPL Weibull model can be obtained from the joint conditional posterior distribution of the model parameters. The joint conditional posterior distribution of model parameters is proportional to the product of the likelihood function and the joint prior distribution of parameters. To get the joint conditional posterior we re-parameterize the IPL Weibull model as

$$f(t|\beta, \theta, \alpha) = \beta\theta t^{\beta-1} \gamma^{\alpha} \exp(-\gamma^{\alpha}\theta t^{\beta}) \quad (32)$$

Where $\theta = L_0^{-\beta}$ and $\alpha = \beta n$. With little information known a priori about β and θ,[12] proposes the use of Jeffrey's vague prior for the scale and shape parameter. Jeffrey's prior is an uninformative prior and is invariant to parameter transformation. For the stress parameter we assume a uniform prior.

$$\pi(\beta) = \frac{1}{\beta}; \pi(\theta) = \frac{1}{\theta} \quad (33)$$

$$\pi(\alpha|A, B) = \begin{cases} \frac{1}{B-A}, & A \leq \alpha \leq B, A, B > 0 \\ 0 & \end{cases} \quad (34)$$

Assuming the right-censored failure data in $\boldsymbol{\mathcal{D}}$ to be a sample from the Weibull distribution with scale parameter given by the inverse power law, the likelihood function of the IPL Weibull model can be obtained from (26). Substituting (13) and (32) in (26), we get for the right-censored IPL Weibull likelihood function

$$\mathcal{L}(\beta, \theta, \alpha|t_1, \ldots, t_r, s_{r+1}, \ldots, s_N, C, N, r) = \beta^r \theta^r \prod_{i=1}^{r} t_i^{\beta-1} \gamma^{\alpha r} \exp(-\gamma^{\alpha}\theta P) \quad (35)$$

Where P is the rescaled test time given by

$$P = \sum_{i=1}^{r} t_i^{\beta} + (N-r)C^{\beta} \quad (36)$$

With the prior distributions of model parameters defined in (33) and (34), the joint conditional posterior distribution of the parameters of IPL-Weibull model is

$$g(\beta, \theta, \alpha|t_1, \ldots, t_r, s_{r+1}, \ldots, s_N, C, N, r) \propto \beta^{r-1}\theta^{r-1} \prod_{i=1}^{r} t_i^{\beta-1} \gamma^{\alpha r} \exp(-\gamma^{\alpha}\theta P) \frac{1}{B-A} \quad (37)$$

Since the Bayes estimators of parameters are nothing but conditional expectations of parameters, it follows from (31), that evaluating such expectations requires the knowledge of marginal conditional posterior of model parameters. The marginal conditional posterior of a parameter can be obtained by integrating out all model parameters, except the parameter of interest from the joint conditional posterior of (37). Once the marginal conditional posterior of a parameter is obtained, the expectation or the Bayes estimator of the parameter can be calculated by evaluating the integral in (31).

For the IPL Weibull model, the Bayes estimator of the shape parameter, β is obtained by integrating out θ and α from the joint conditional posterior in (37). The Bayes estimator or the expectation of β is

$$\hat{\beta} = K_{\beta,r} \int_0^{\infty} \beta^r \prod_{i=1}^{r} t_i^{\beta-1} P^{-(r)} d\beta \quad (38)$$

Where $K_{\beta,r}$ is the normalization constant and is given by

$$K_{\beta,r}^{-1} = \frac{1}{B-A} \int_0^{\infty}\int_A^{B} \beta^{r-1} \prod_{i=1}^{r} t_i^{\beta-1} P^{-(r)} d\alpha d\beta \quad (39)$$

The integrals in (38) and (39) do not have analytical solutions but can be solved via numerical integration. Similarly, the Bayes estimator of the re-parametrized scale parameter, θ can be obtained by integrating out

β and α from (37). This gives the Bayes estimator of the scale parameter as

$$\hat{\theta} = \frac{1}{B-A}\frac{\gamma^A - \gamma^B}{\gamma^{A+B}\ln\gamma}\frac{I_1}{I_2} \quad (40)$$

Where

$$I_1 = \int_0^\infty \beta^{r-1}\prod_{i=1}^{r} t_i^{\beta-1}\, P^{-(r+1)} d\beta \quad (41)$$

$$I_2 = \int_0^\infty \beta^{r-1}\prod_{i=1}^{r} t_i^{\beta-1}\, P^{-(r)} d\beta \quad (42)$$

The integrals in (41) and (42) can be evaluated by numerical integration and as such do not have closed-form solutions. Finally, the Bayes estimator of the parameter α can be obtained by integrating out β and θ from (37). This results in for the Bayes estimator of α

$$\hat{\alpha} = \frac{K_{\alpha,r}}{B-A}\iint_{0A}^{\infty B} \beta^{r-1}\prod_{i=1}^{r} t_i^{\beta-1}\, P^{-(r)}\alpha d\alpha d\beta \quad (43)$$

Where

$$K_{\alpha.r}^{-1} = \int_0^\infty \beta^{r-1}\prod_{i=1}^{r} t_i^{\beta-1}\, P^{-(r)} d\beta$$

The Bayesian updating algorithm can be summarized in the following steps

- Formulate the joint likelihood function of the failure model considering the censored failure data. This formulation with respect to IPL-Weibull model in given in (35).
- Choose prior distributions of parameters to be estimated and obtain the conditional joint posterior of model parameters. For the IPL Weibull model, this is given in (37).
- Obtain marginal conditional posterior of parameters by integrating out the all the model parameters from the joint conditional posterior, except the parameter of interest.
- The Bayes estimators of parameters can then be obtained by evaluating the integral in (31).

7. Model Validation

The analysis in the previous section assumes that the censored failure data is a sample from the Weibull distribution in which the scale parameter shares an inverse relationship with the applied stress. In order to check the validity of the assumption, we compute the Kolmogorov-Smirnov (KS) distance between the empirical distribution function and the fitted distribution function. KS test is a nonparametric test that is used to test the hypothesis that the sample (failure data) comes from a particular distribution. In our case, we hypothesize that the times to failure come from an IPL-Weibull distribution. The sample failure data is given in . The failure data assumes a total of 40 assets installed ten years ago with a retirement history, as shown in Table 1.

Figure 1 shows the result of the KS test. It is clear that the fitted IPL-Weibull provides an excellent fit to the failure data. Note that the fitted IPL-Weibull CDF uses point estimates of the parameters given by (38),(40) and (43). Figure 1 confirms that the KS test accepts the hypothesis that the failure data of is a sample from the Weibull distribution with scale parameter given by the inverse power law.

Table 1 Sample Data Set for Transformer from Single Vintage [6]

Year	Age	Retirements	Survivors	Survivor Rate
2009	0	0	40	100%
2010	1	1	39	98%
2011	2	0	39	98%
2012	3	0	39	98%
2013	4	0	39	98%
2014	5	1	38	95%
2015	6	2	36	90%
2016	7	1	35	88%
2017	8	5	30	75%
2018	9	4	26	65%
2019	10	6	20	50%

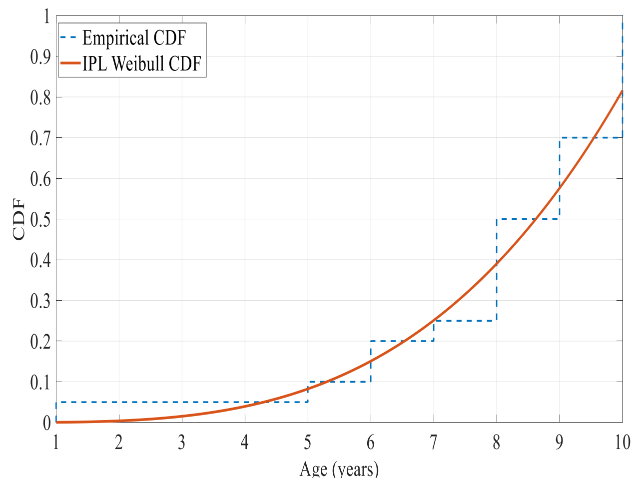

Figure 1 Empirical and Fitted CDF from KS Test

8. Illustrative Example

The proposed Bayesian method is applied to a model of the IEEE-34 Bus test system to estimate the lifetimes of voltage regulators with different penetration levels of solar generation. The test system

has a nominal voltage of 24.9 kV. The feeder is characterized by long lines and light loads and requires two voltage regulators to keep the voltage within ANSI limits. Both the voltage regulators use line drop compensation (LDC) to measure the drop in voltage between the regulator and load center. To study the impact of solar generation on tap-changers, a full three-phase model of the circuit consisting of all circuit lines (single-phase and three-phase lines), regulators, customer loads, capacitor banks, substation, and in-line transformers, control elements of capacitors and voltage regulators are developed in OpenDSS. For the solar generation, a proportionally distributed configuration is chosen with the rated power of the PV systems proportional to the loads. Each PV system is interfaced with an inverter with a rating 10% higher than the PV panel. It is expected that an increase in the solar capacity will negatively impact the device lifetime due to a corresponding increase in the device wear and tear.

This engineering judgment allows us to assign a prior distribution to the model scale parameter such that the inverse proportionality between device lifetime and applied stress is preserved. As for the stress parameter, since we assigned a uniform prior, the length of the interval, $[A, B]$ chosen will have an impact on the model predictions. Obviously, an expert with a priori information about the stress parameter will be able to assign more accurate values for a better forecast. This shows the proposed model is flexible and can incorporate subjective information, if available. Figure 2 shows the survivor data of and the

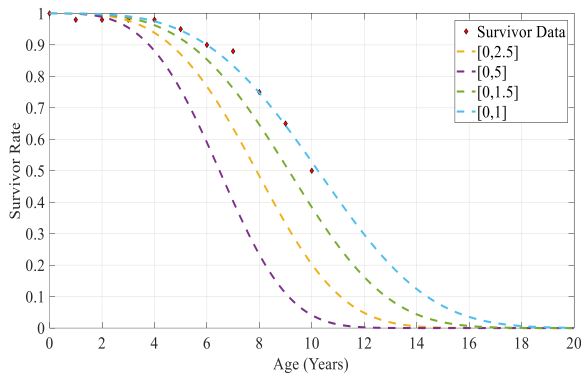

Figure 2 Survivor Functions of IPL-Weibull with different intervals for stress parameter

fitted IPL Weibull survivor functions with different lengths of the interval chosen for the stress parameter. For our purpose, given the sample failure data, the interval $[0,1]$ of the stress parameter results in the best fit of the given survivor rate.

To observe the impact of non-thermal stress or equivalently solar generation on the OLTC lifetimes, we designed annualized experiments over a 10-year planning horizon. The results of the tests are categorized into three scenarios which are 1) Load growth over the next ten years with no solar generation, 2) Load growth with low penetration of solar generation (30%) and 3) Load growth with high penetration of solar generation (90%).

Note that the photovoltaic (PV) penetration level is defined as the ratio of aggregate peak capacity of all PV systems and the total peak active load of the feeder. We consider a load growth of 3% for the first four years, followed by 5% and a 7% growth distributed equally for the remaining six years. This is within the conservative estimate of 3%-7% growth in feeder loads at the distribution level. We consider two different penetration levels for the solar generation over the ten-year planning horizon. Note that the accumulated stress on the OLTCs is proportional to the cumulative number of tap operations over the planning horizon.

Since the failure data is censored on the right, we have two sets of observed lifetimes. One set contains the failed/retirement times $(t_i \leq C); i = 1, \dots, r$ and the second set includes the survival times of the OLTCs not failed or retired by the present time instant.

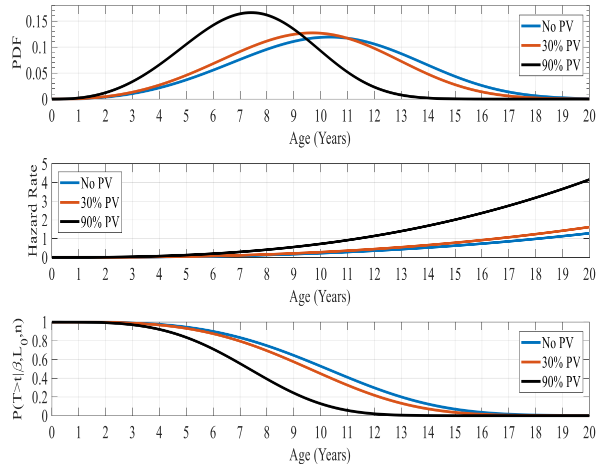

Figure 3 PDF, Hazard Rate and Survivor Function for VR-1

The Bayesian method enables us to make an inference on the second set of OLTCs based on the data in the first set since the failure/retirement time of the second set is unknown at the present time instant. The inference is valid because the OLTCs in both the sets experience similar failure mechanisms. Figures 3 and 4 show the failure density function, hazard rate and the survival function of the two voltage regulators of IEEE 34 bus system under different scenarios.

The regulators are assumed to have survived until the present time instant and hence are not censored. From Figure 3 and Figure 4, it is clear that while a low PV penetration may not significantly reduce the lifetimes of voltage regulators, sustained high PV penetration has a significant impact on the device's lifetime. This is further evinced by the hazard rates of both the regulators, which show a marked increase towards the end of device lifetime with high PV penetration.

It is interesting to note how the failure density changes after new data are acquired. Figure 5 shows the impact of censoring number on the failure density function. The transition of failure density from the poor initial estimate (blue curve) to the final estimate

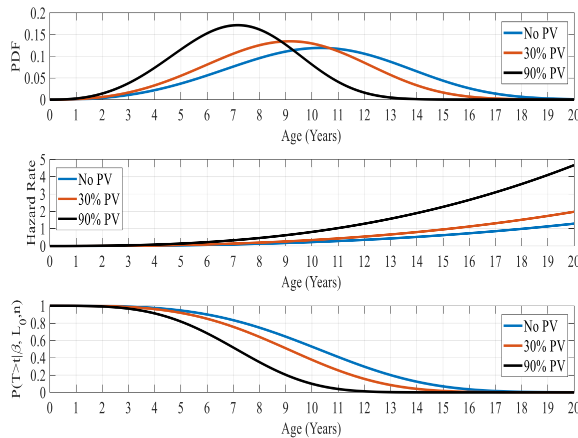

Figure 4 PDF, Hazard Rate and Survivor Functions for VR-2

(green curve) is remarkably fast. The mean time to failure in the initial estimate with $r = 1$ is 114200 years, considering no PV penetration. With $r = 20$,

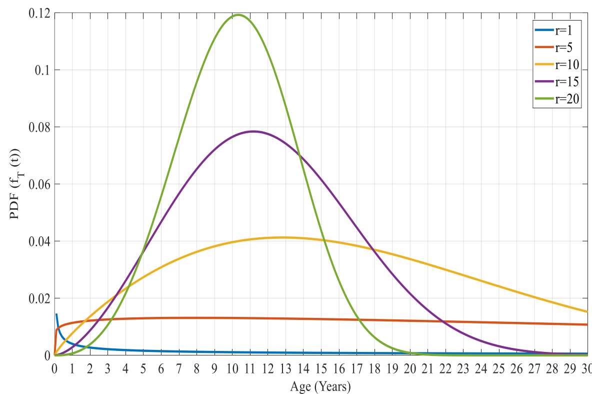

Figure 5 Impact of Censoring on Failure Time Prediction

the mean time to failure is 10.255 years. This suggests the method is very robust. Since the inference on the surviving assets is drawn from the set of failed/retired assets, the proposed method will perform better with a large amount of censored failure data.

Table 2 list the point estimates of the model parameters as a function of the censoring number. As more failure data are acquired, the failure prediction improves. The most likely $(\hat{\beta}, \hat{\theta}, \hat{\alpha})$ with $r = 20$ is (3.52,9.16,0.1108) for voltage regulator-1in the No PV scenario. Table 3 and Table 4 list some statistical properties of the failure density function shown in Figure 3 and Figure 4. The impact of heavy PV penetration on device lifetime is quite apparent. This can be realized by observing the mean time to failure (MTTF) of the two voltage regulators under low PV and high PV penetration and compare that with the No PV scenario. With a high PV penetration, the MTTF of voltage regulator 1 is 7.33 years and for voltage regulator 2, the MTTF with high PV penetration is 7.09 years.

Table 2 Impact of censoring on Point Estimates of Model Parameters

Parameter	r=1	r=5	r=10	r=15	r=20
$\hat{\beta}$	0.43	1.11	1.83	2.65	3.5217
$\widehat{L_0}$	7655	30.7	13.09	10.06	9.1617
$\hat{n}$	0.88	0.34	0.21	0.14	0.1108

Table 3 Statistical Properties of Aging Distribution of VR-1

Property	No PV	30%PV	90%PV
Mean Time to Failure (MTTF)	10.25	9.58	7.33
Median	10.238	9.59	7.34
Mode	10.33	9.68	7.41

Table 4 Statistical Properties of Aging Distribution of VR-2

Property	No PV	30%PV	90%PV
Mean Time to Failure (MTTF)	10.25	9.06	7.09
Median	10.238	9.07	7.10
Mode	10.33	9.15	7.17

9. Conclusions

Accurate prediction of service times of power equipment is a critical issue in asset management. The development of statistical models that can incorporate the most salient features thought to strongly determine the device lifetime is pivotal to reliability and the economic operation of the power system. Such models are especially sought in situations where a portion of the installed population fails or is retired

progressively, assuming similar failure mechanism for the entire population. A good example of this is on-load tap changers and switched capacitors on distribution feeders with significant solar or wind generation.

In this work, we propose the use of IPL-Weibull distribution to accurately model the impact of operational stress on the tap-changers given a high penetration of solar generation. Also, a Bayesian approach to estimate the model parameters is presented. The method uses point estimates of the model parameters, which are obtained via Bayesian updating of the acquired failure data. The model is flexible and can incorporate subjective information, if available, in the form of the prior distribution of parameters.

The future work in this direction will involve the quantification of uncertainty in parameter estimation. This can be done by obtaining the posterior predictive distribution of the times-to-failure of devices not censored by the present time instant. For the IPL Weibull model, numerical integration techniques are not sufficient to derive the posterior predictive distribution. However, simulation procedures like Metropolis Hastings algorithm or Gibbs sampler can be used to simulate to posterior predictive distribution of the IPL Weibull failure model.

10. Acknowledgment

This publication was made possible by NPRP grant # 9-204-2-103 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

11. References

[1] R. E. Brown, *Electric power distribution reliability*. 2017.

[2] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems*. 1996.

[3] E. R. Ziegel and B. Dodson, “Weibull Analysis,” *Technometrics*, 1995, doi: 10.2307/1269925.

[4] J. Yan, “Bayesian Survival Analysis,” *J. Am. Stat. Assoc.*, 2004, doi: 10.1198/jasa.2004.s359.

[5] M. Begovic and P. Djuric, “On predicting the times to failure of power equipment,” in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2010, doi: 10.1109/HICSS.2010.290.

[6] Q. Chen and D. M. Egan, “A bayesian method for transformer life estimation using Perks' hazard function,” *IEEE Trans. Power Syst.*, 2006, doi: 10.1109/TPWRS.2006.881129.

[7] Wasserman, “All of Statistics : A Concise Course in Statistical Inference Brief Contents,” *Simulation*, 2004, doi: 10.1007/978-0-387-21736-9.

[8] A. van Schijndel, “Power Transformer Reliability Monitoring,” Eindhoven University of Technology, 2010.

[9] X. Zhang, E. Gockenbach, V. Wasserberg, and H. Borsi, “Estimation of the lifetime of the electrical components in distribution networks,” *IEEE Trans. Power Deliv.*, 2007, doi: 10.1109/TPWRD.2006.876661.

[10] M. M. Begovic, A. Peerzada, W. Rohouma, and R. Balog, “On Accelerated Aging of Mechanical Assets in Distribution Systems with Renewable Generation,” in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020, doi: 10.24251/hicss.2020.362.

[11] H. Rinne, *The Weibull Distribution*. 2008.

[12] S. K. Sinha and I. Guttman, “Bayesian analysis of life—testing problems involving the weibull distribution,” *Commun. Stat. - Theory Methods*, 1988, doi: 10.1080/03610928808829627.