Research Article

A Poisson-Fault Model for Testing Power Transformers in Service

Dengfu Zhao,1,2 Zheng Zhao,2,3 Qihong Duan,1 and Gongnan Xie4

1 School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
2 School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
3 State Power Economic Research Institute, Beijing 10052, China
4 School of Mechanical Engineering, Northwestern Polytechnical University, Xian, Shaanxi 710072, China

Correspondence should be addressed to Qihong Duan; k.h.tuan@stu.xjtu.edu.cn

Received 5 December 2013; Accepted 28 April 2014; Published 14 May 2014

Academic Editor: Yang Tang

Copyright © 2014 Dengfu Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a method for assessing the instant failure rate of a power transformer under different working conditions. The method can be applied to a dataset of a power transformer under periodic inspections and maintenance. We use a Poisson-fault model to describe failures of a power transformer. When investigating a Bayes estimate of the instant failure rate under the model, we find that complexities of a classical method and a Monte Carlo simulation are unacceptable. Through establishing a new filtered estimate of Poisson process observations, we propose a quick algorithm of the Bayes estimate of the instant failure rate. The proposed algorithm is tested by simulation datasets of a power transformer. For these datasets, the proposed estimators of parameters of the model have better performance than other estimators. The simulation results reveal the suggested algorithms are quickest among three candidates.

1. Introduction

The failure of power transformers is one of most frequent causes of long interruptions of power systems. Many researchers have studied evaluation schemes of a piece of equipment of power systems; see [1–3] and reference therein, and the IEEE C57-110-98 standard [4] is developed to determine the capability of a power transformer. However, due to the long lifetime of a power transformer as well as the complexity of an experiment, few statistical experiments about the lifetime or the failure rate of power transformers have been performed. A singular experiment, in which eight large transformers are tested, is reported in [5]. On the other hand, there are many power transformers in service, and we can obtain data of periodic inspections and maintenance for them. The fact inspires us to find from the data the relation between the failure rate and the working condition of a power transformer.

While investigating a statistical method for this problem, we face three difficulties. The first difficulty is as follows: data is collected according to a standard test procedure [6, 7], but the procedure does not take into account our problem. What we know is the working condition of a power transformer at an inspection as well as whether or not a transformer power fails. As time between successive inspections is far less than a typical lifetime of a transformer power, we can assume that the number of failures that occur between two successive inspections has a Poisson distribution; see [8, 9] for more details about a Poisson model of similar cases. The rate of the Poisson distribution, which is proportional to the instant failure rate, is associated with the current working condition of a power transformer. According to an idea of Bayesian experimental design [10], we assume that, for a power transformer, the instant failure rate with a given working condition has a prior normal distribution. Then the Bayes estimate is the best point estimate to the instant failure rate with the quadratic loss function.

The second difficulty is that the mean and the variance matrix of the prior distribution might not be easily determined. Some empirical models, including the inverse power
law model [11] and the Arrhenius model recommended by the IEC loading guide [12], may motivate us. These models study the influence of working temperature and operating environment on the lifetime of a power transformer. Substituting in each model by an appropriate parameter, we find that the model is an ideal case of an autoregression model. Considering observed errors as well as possible bedrock errors of these empirical models, we assume that the instant failure rate of a power transformer with respect to working conditions is an autoregressive process. Furthermore, by applying a classical method of time series and a technique of martingales, we can establish estimates to parameters for the prior distribution. Comparing with [6], we establish an acceptable method for investigating the instant failure rate of a power transformer.

The last difficulty appears while computing the Bayes estimate to the instant failure at different working conditions. Let \( N \) denote the total number of working conditions and let \( L \) denote the number of failures. The Bayes estimate involves a \( \text{L} \)th moment of \( N \)-dimensional normal random variable. The time complexity of a classical algorithm to the moment is \( O(N^2) \). Although a Monte Carlo simulation is not sensitive to the dimension \( N \), its performance is not good when \( L \) is slightly large.

Recently, there is some advance in filtered estimates for Poisson process observations; see [13, 14] and references therein for example. By modulating the preceding estimates for our case, we can establish an evolutionary equation for the Bayes estimate. A closed-form solution to the evolutionary equation is obtained and a corresponding quick algorithm is then established. Similarly to [8, 13], we establish a closed-form solution to the problem. Moreover, the corresponding algorithm is exact. As a comparison, [13] used an approximation algorithm. Finally, the simulation result reveals that the algorithm is quicker than a Monte Carlo simulation, which is suggested by [15] for a similar problem.

The rest of the paper is organized as follows. We construct the model in Section 2 and discuss its relation with available empirical models. Section 3 presents the setting of the prior distribution. In Section 4, we investigate the Bayes estimate of the instant failure rate. After establishing a classical algorithm and a Monte Carlo simulation of the Bayes estimate and exploring its performance, we present a quick algorithm according to a new theory in Section 4. Simulation results are presented in Section 5. Section 6 concludes the paper.

2. Constructing a Poisson-Fault Model

In this section, after a literature review about some empirical models, we propose a Poisson-fault model of a power transformer.

2.1. Empirical Models. The Arrhenius model is accepted widely. The IEC loading guide [12] uses the model and many authors investigated it; see [6, 7, 16] for more details. According to the Arrhenius reaction rate equation, the average lifetime \( t \) of a piece of equipment at different temperature in the Arrhenius model is given by

\[
t = A \exp (-kT).
\]  

Here \( T \) is the temperature, \( A \) is a nonthermal constant, and \( k \) is the multiplication of the activation energy and the Boltzmann constant.

The power law model [11] can describe dielectric failures of a capacitor under increased voltages and the lifetime of a multicomponent system. The average lifetime \( t \) in that model is given by

\[
t = \gamma V^{-n}.
\]  

Here \( V \) represents a voltage level or a level of a severe operating environment. With the nonlinear transformation \( U = \log(V) \) of the level, we have that \( t = \gamma \exp(-\eta U) \). Hence, we may regard the power law model as another edition of the Arrhenius model (1).

Typically, \( T \) in (1) may be substituted by other parameters. In most applications, it is assumed that the increased level of the working condition affects the average of the lifetime and the corresponding instant failure rate.

2.2. Poisson-Fault Model. In general, it is assumed that the number of failures has a normal distribution in a given time interval. The assumption is correct if test duration is long. As we know, when using a short test duration, the number of failures has a Poisson distribution; see [8, 9] for more details.

We consider a right-continuous and complete filtration \( \mathcal{F} = \{\mathcal{F}_n : n = 0, 1, \ldots\} \) on a complete probability space. Two nonnegative stochastic processes \( X = \{X_n : n = 1, 2, \ldots\} \) and \( M = \{M_n : n = 1, 2, \ldots\} \) are adapted to the filtration \( \mathcal{F} \). We refer to [17] for an introduction of filtrations and stochastic processes adapted to filtration. We call \( M \) a Poisson-fault model with the intensity \( X \) if these two processes satisfy the following conditions.

(C1) For \( n = 1, 2, \ldots, E\epsilon_n = 0, D\epsilon_n \) is a positive constant \( \sigma^2 \) and \( \epsilon_{n+1} \) is independent with respect to \( \mathcal{F}_n \). For given constants \( a, b \) such that \( |a| < 1, b > 0 \),

\[
X_{n+1} = aX_n + b + \epsilon_{n+1}.
\]  

(C2) The distribution of \( M_{n+1} \) conditional on a \( \sigma \)-field \( \mathcal{F}_n \) is a Poisson distribution with a parameter \( X_{n+1} \). That is, for \( k = 0, 1, \ldots, \)

\[
P(M_{n+1} = k \mid \mathcal{F}_n \vee \sigma(X_{n+1})) = \frac{X_{n+1}^k}{k!} \exp (-X_{n+1}).
\]  

Here, \( \mathcal{F}_n \) describes the full information at level 1, \ldots, \( n \). And the combination \( \mathcal{F}_n \vee \sigma(X_{n+1}) \) describes the information
of both $F_n$ and $\sigma^2$. In the above model for a power transformer, $n$ represents the level of working conditions. $M_n$ represents the number of failures between two successive inspections at level $n$.

2.3. Relation between Models. To motivate the model, suppose that $F_n$ is the instant failure rate of a power transformer under a working condition with level $n$. When there are $p$ power transformers at a working condition and the length of time interval between inspections is $d$, we have that the average number $X_n$ of failures is proportional to the failure rate $F_n$. Concretely,

$$X_n = dpF_n.$$  

(5)

Moreover, $M_n$, the number of failures at the working condition $n$, has a Poisson distribution with parameter $X_n$. Hence, we conclude that a Poisson-fault model describes an instant failure rate under different working conditions, and an instant failure rate is an autoregressive process.

Now we point out the relationship between the Arrhenius model (1) and our model. We consider a special case in which $b = 0$ and $e_n = 0$ for $n = 2, 3, \ldots$. In this case, we have

$$X_n = a^{n-1}X_1.$$  

(6)

We may write the corresponding “temperature” $T = C - n\Delta$, where $C$ is a constant and $\Delta$ is the change of two successive working conditions. Moreover, the failure rate $F_n$ is the reciprocal of the average lifetime for a piece of equipment for a short time interval $d$. Hence, we have that the average lifetime $t$ at the level $n$ is given by $t = A \exp(-kT)$, where $A = a^{(1-C)/\Delta} (dpX_1)^{-1}$ and

$$k = -\frac{\log(a)}{\Delta}.$$  

(7)

By comparing this equation with (1), we may conclude that the autoregressive process $X$ incorporates and extends features of the Arrhenius model (1).

3. Prior Distribution and Its Parameter

For the latent autoregressive process $X$, $(X_1, \ldots, X_N)$ has a prior normal distribution. In this section, we propose a method to estimate parameters of the prior distribution.

3.1. Background of Time Series. Here we review some results of the classical theory of time series; see [18] for a detailed introduction to time series.

We have that for $n \geq m$, $\text{Cov}(X_m, X_{m+1}) = a\text{Cov}(X_m, X_n)$ and $DX_{n+1} = a^2 DX_n + \sigma^2$. Provided that $X_n$ is a stable stochastic process, we have $EX_n = b/(1-a)$ and $DX_n = \sigma^2/(1-a^2)$. That is, the parameter of the prior distribution is determined by providing $a$, $b$, and $\sigma^2$. Now we turn to estimate these key parameters.

With the conditions $|a| < 1$ and $b \geq 0$, the intensity process $X_n$ is an asymptotic stable stochastic process. Hence, three sequences $EX_n$, $EX^2_n$, and $EX_{n+1}X_n$ are convergent. Therefore, these sequences are bounded.

Assume $(x_1, \ldots, x_N)$ is a dataset of observations for the intensity process $X_n$, $n = 1, \ldots, N$. It is well known that we can solve linear equations

$$A \left( \begin{array}{c} \tilde{a} \\ \tilde{b} \end{array} \right) = \beta, \quad A = \left( \begin{array}{cc} \sum_{n=1}^{N-1} x_n^2 & \sum_{n=1}^{N-1} x_n \\ \sum_{n=1}^{N-1} x_n & N-1 \end{array} \right),$$

(8)

for estimators $\tilde{a}$ and $\tilde{b}$. As $N$ tends to $\infty$, $\tilde{a}$ and $\tilde{b}$ converge to $a$ and $b$ with a probability 1, respectively. Moreover, as $Ee_n = 0$, we can obtain an estimator $\tilde{\sigma}^2$ by the following formula:

$$\tilde{\sigma}^2 = \frac{1}{N-1} \sum_{n=1}^{N-1} (x_{n+1} - \tilde{a}x_n - \tilde{b})^2.$$  

(9)

3.2. Parameter Estimation to the Prior Distribution. Assume that we face a history record of failures. We can then obtain $M_n$, $n = 1, \ldots, N$ by counting failures. However, we cannot obtain exactly $X_n$ due to noises. In this case, we must establish new estimators for $a$ and $b$ only based on observations of $M_1, \ldots, M_N$. In a simple heuristic, we may replace $X_i$ by $M_i$ in (8) to (9) and obtain estimators $\tilde{a}^*$ and $\tilde{b}^*$ for $a$ and $b$, respectively. Based on simulation results in Section 5, we will conclude that $\tilde{a}^*$ and $\tilde{b}^*$ are bad estimators. In the following, we construct a good approximation of $X_n$ based on $M_1, \ldots, M_N$ and obtain better estimators for $a$ and $b$.

**Theorem 1.** $M = \{M_n : n = 1, 2, \ldots\}$ is a Poisson-faults model with an intensity $X = \{X_n : n = 1, 2, \ldots\}$. As $N$ tends to $\infty$, with a probability 1, one has that

$$\frac{1}{N-1} \sum_{n=1}^{N-1} (M_n - X_n) \longrightarrow 0,$$

$$\frac{1}{N-1} \sum_{n=1}^{N-1} (M_{n+1}^2 - M_n X_n) \longrightarrow 0,$$

$$\frac{1}{N-1} \sum_{n=1}^{N-1} (M_n^2 - M_n - X_n^2) \longrightarrow 0.$$  

(10)

See Appendix A for the proof of Theorem 1.
Then, it follows from Theorem 1 and (8) that we can solve linear equations
\[
\bar{\mathbf{A}} \left( \bar{\mathbf{a}} \right) = \bar{\mathbf{b}}, \quad \bar{\mathbf{A}} = \begin{pmatrix}
\sum_{n=1}^{N-1} M_n^2 - M_n, & \sum_{n=1}^{N-1} M_n \\
\sum_{n=1}^{N-1} M_n, & N - 1
\end{pmatrix}.
\]
(11)

for estimators \( \bar{a} \) and \( \bar{b} \). Moreover, it follows from Theorem 1 and the convergence of \( \bar{a} \) and \( \bar{b} \) that \( \bar{a} \) and \( \bar{b} \) converge to \( a \) and \( b \) with probability 1, respectively. Therefore, as \( N \) tends to \( \infty \), \( \bar{a} \) and \( \bar{b} \) converge to \( a \) and \( b \) with a probability 1, respectively. Similarly, from (9), we can construct an estimator
\[
\bar{\sigma}^2 = \bar{b}^2 + \frac{1}{N-1} \sum_{n=1}^{N-1} (M_n^2 - M_n) + (M_{n+1}^2 - M_{n+1})
\]
\[
- 2\bar{a}M_nM_{n+1} - 2\bar{b}M_{n+1} + 2\bar{a}\bar{b}M_n
\]
(12)

for the parameter \( \sigma^2 \).

4. Bayes Estimate of the Instant Failure Rate

In this section, from observations of \( M_1, \ldots, M_N \), we try to estimate the instant failure rate \( X_n \). According to the general theory of Poisson compound decision problem [9], we establish a formula to the estimate. After investigating the time complexity of the formula, we turn to another quick algorithm.

4.1. Formula Involving Moments

**Theorem 2.** \( M = \{M_1, \ldots, M_N\} \) is a Poisson-fault model with the intensity \( X \). \( \{m_1, \ldots, m_N\} \) is a dataset of observations of \( M \). Assume that the prior normal distribution of \( X_1, \ldots, X_N \) has a mean \( \bar{\mu} \) and a covariance matrix \( B \). Let \( U = (U_1, \ldots, U_N) \) be a normal random variable with mean \( \mu = \bar{\mu} - \mathbf{B} \bar{e} \) and covariance \( B \), where \( \mathbf{e} \) denotes \( N \)-dimensional vector whose every entry is 1. One has that, for \( n = 1, \ldots, N \),
\[
E[X_n | M_1 = m_1, \ldots, M_N = m_N] = \frac{E(U_n \prod_{i=1}^{n-1} U_i^{m_i})}{E\prod_{i=1}^{n-1} U_i^{m_i}}.
\]
(13)

See Appendix B for the proof of Theorem 2.

From Theorem 2, we can obtain the Bayes estimate \( \mathbb{E}[X_n | M_1, \ldots, M_N] \) by computing \( L + 1 \)th and \( L \)th moments of a normal random variable \( U \), where \( L = m_1 + \cdots + m_N \) is the total number of failures. The moment generating function \( \phi(t) \) of \( U \) provides a tool to calculate the moments; see [20] for details of a moment generating function. For \( t = (t_1, \ldots, t_N) \), the moment generating function \( \phi(t) \) of \( U \) is defined by
\[
\phi(t) = E \exp \left( \sum_{i=1}^{N} t_i U_i \right) = \exp \left( \sum_{i=1}^{N} t_i \mu_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} t_i t_j \phi_{ij} \right),
\]
(14)

where \( (\mu_1, \ldots, \mu_N) = \mu \) and \( (\phi_{ij}) = \mathbf{B} \). In general, \( n \)th derivative of \( \phi(t) \) evaluated at \( t = 0 \) equals a corresponding \( n \)th moment of \( U \). Concretely, for nonnegative integers \( n_1, \ldots, n_N \),
\[
E \left( \prod_{i=1}^{N} U_i^{n_i} \right) = \frac{\partial^{(n_1 + \cdots + n_N)}}{\partial t_1^{n_1} \cdots \partial t_N^{n_N}} \phi(t) \bigg|_{t=(0,\ldots,0)}.
\]
(15)

On the other hand, differentiation of \( \phi(t) \) with respect to \( t_k \) yields
\[
\frac{\partial}{\partial t_k} \phi(t) = \left( \mu_k + \sum_{j=1}^{N} t_j \phi_{kj} \right) \phi(t).
\]
(16)

Hence, for nonnegative integers \( s_1, \ldots, s_N \),
\[
\frac{\partial^{(s_1 + \cdots + s_N)}}{\partial t_1^{s_1} \cdots \partial t_N^{s_N}} \phi(t) = \left( \mu_k + \sum_{j=1}^{N} t_j \phi_{kj} \right) \frac{\partial^{(s_1 + \cdots + s_N-1)}}{\partial t_1^{s_1-1} \cdots \partial t_k^{s_k-1} \cdots \partial t_N^{s_N}} \phi(t) + \sum_{j=1}^{N} t_j \phi_{kj} \frac{\partial^{(s_1 + \cdots + s_N-1)}}{\partial t_1^{s_1-1} \cdots \partial t_k^{s_k-1} \cdots \partial t_N^{s_N}} \phi(t).
\]
(17)

Moreover, it follows from (15) that
\[
E \left( U_k \prod_{i=1}^{N} U_i^{s_i} \right) = \mu_k E \left( \prod_{i=1}^{N} U_i^{s_i} \right) + \sum_{j=1}^{N} s_j \phi_{kj} E \left( U_j^{-1} \prod_{i \neq j}^{N} U_i^{s_i} \right),
\]
(18)

As we can obtain the 1st and 2nd moments of \( U \) from its mean \( \mu \) and covariance matrix \( B \), the preceding (18) provides us with a formula to calculate any higher-order moment of \( U \). However, the formula involves \( N \) order tensor. The time complexity of a corresponding algorithm for \( E(\prod_{i=1}^{N} U_i^{m_i}) \) is \( O(N^2) \), where \( L = m_1 + \cdots + m_N \).

As an alternate way to calculate \( E(\prod_{i=1}^{N} U_i^{m_i}) \), we had studied a Monte Carlo method [15, 21]. In general, to achieve the specified accuracy of \( E(\prod_{i=1}^{N} U_i^{m_i}) \), the runtime of a Monte Carlo algorithm is proportional to the variance of the random variable \( \prod_{i=1}^{N} U_i^{m_i} \). Our simulation shows that the performance of the Monte Carlo algorithm is not good when \( L \) is slightly large.
4.2. Methodology of a Quick Algorithm. So we turn to another technique to calculate the Bayes estimate. The methodology is similar to [13, 14].

Assume that \( f(x_n) \) and \( g(e_n) \), \( n = 2, 3, \ldots \), are probability density functions (pdf) of random variables \( X_1 \) and \( e_n, n = 2, 3, \ldots \), respectively. As \( e_n \) is independent with respect to \( \mathcal{F}_{n-1} \), we have that the multivariate pdf of \( (X_1, e_2, \ldots, e_N) \) is

\[
h_N(x_1, e_2, \ldots, e_N) = f(x_1) \prod_{n=2}^{N} g(e_n).
\]

Suppose that \( X_n = x_n, n = 1, \ldots, N \), is a dataset of the intensity process. It follows from the definition of the Poisson-fault model that the conditional probability function of \( M_1, \ldots, M_N \) is

\[
P(M_1 = m_1, \ldots, M_N = m_N | X_1 = x_1, \ldots, X_N = x_N) = \prod_{n=1}^{N} \frac{x_n^{m_n}}{m_n^n!} \exp(-x_n).
\]

Therefore,

\[
P(M_1 = m_1, \ldots, M_N = m_N) = \int_{R^N} \prod_{n=1}^{N} \left( \frac{x_n^{m_n}}{m_n^n!} \exp(-x_n) \right) h_N d\tau_N.
\]

Here \( h_N = h_N(x_1, e_2, \ldots, e_N) \), \( d\tau_N = dx_1 de_2 \cdots de_N \), and \( x_n = ax_{n-1} + b + e_n \) for \( n = 2, \ldots, N \). From the above discussion, we can derive the following expectation of \( X_j \) conditional on a dataset of \( M_1, \ldots, M_N \)

\[
E[X_j | M_1 = m_1, \ldots, M_N = m_N] = \int_{R^N} x_j \prod_{n=1}^{N} \left( \frac{x_n^{m_n}}{m_n^n!} \exp(-x_n) \right) h_N d\tau_N.
\]

To compute auxiliary integrals in (21) and (22), we study the following auxiliary function \( I_n(k_1, \ldots, k_n) \) for \( n = 1, \ldots, N \):

\[
I_n = \int_{R^N} \prod_{i=1}^{n} x_i^{k_i} \exp \left( -\sum_{i=1}^{n-1} x_i - 1 - a^{N-n-1} x_n \right) h_{n-1} d\tau_{n-1}.
\]

As \( x_n = ax_{n-1} + b + e_n \), writing \( c_n = (1 - b^{N-n-1})(1 - b) \), we have that

\[
I_n = \exp(-c_n b) \times \int_{R} \exp(-c_n a x_{n-1}) \times \prod_{i=1}^{n-1} x_i^{k_i} \exp \left( -\sum_{i=1}^{n-1} x_i \right) h_{n-1} \times \left( \int \, g(e_n) \, (ax_{n-1} + b + e_n)^{k_n} \exp(-c_n e_n) \, de_n \right) \times d\tau_{n-1}.
\]

Hence, an evolutionary equation of \( I_n \) follows from the binomial formula

\[
I_n(k_1, \ldots, k_n) = e^{-c_n b} \sum_{i=0}^{k_n} \binom{k_n}{i} \exp \left( -\sum_{i=1}^{n-1} x_i - 1 - a^{N-n-1} x_n \right) h_{n-1} (k_1, \ldots, k_{n-1} + i).
\]

That is, based on expectations of several simple functions of \( x_1 \) or \( e_n, n = 2, 3, \ldots \), we can obtain iteratively \( I_n(k_1, \ldots, k_n) \) from (25). Moreover, it follows from (21) and (22) that

\[
E[X_j | M_1 = m_1, \ldots, M_N = m_N] = \frac{I_N(m_1, \ldots, m_j, 1, \ldots, m_N)}{I_N(m_1, \ldots, m_N)}.
\]

4.3. Algorithm. Comparing (26) with (18), we note that there is still a tensor \( I_N(m_1, \ldots, m_N) \) in (26). We try to simplify \( I_N(m_1, \ldots, m_N) \).

We know that \( e_n \) has a normal distribution \( N(0, \sigma^2) \) and we assume that \( X_1 \) has a prior normal distribution \( N(x_1, \sigma^1) \). Then we can derive a straightforward procedure to compute \( E[X_j | M_1 = m_1, \ldots, M_N = m_N] \).

Assume that a random variable \( Y \) has a normal distribution \( N(\mu, \sigma^2) \). Following the method of Section 4.1, for any nonnegative integer \( k \), we have

\[
E[(b + Y)^k e^{-c Y}] = \exp \left( \frac{c^2 \sigma^2}{2} - c \right) \sum_{i=0}^{k} C(k, i) (b + \mu Y - c \sigma^2)^{k-i} I_i.
\]

Here \( I_{2i+1} = 0 \) and \( I_2 = \prod_{i=1}^{j} (2j - 1) \) for \( i = 0, 1, \ldots \).

In the following, the right side of (27) is denoted by \( E(k, c, b, \mu, \sigma^2) \). Based on the method of Section 4.2, we can derive the following algorithm to compute \( I(k_1, \ldots, k_N) \).

To calculate \( I_N(m_1, m_2, \ldots, m_N) \), it is obvious that the time complexity of the preceding algorithm is \( O(NL^2) \), where \( L = m_1 + \cdots + m_N \). Moreover, we can obtain \( E[X_j | M_1 = m_1, \ldots, M_N = m_N] \) from (26).

5. Simulation Results

We will study the performance of a power transformer from the viewpoint of a Poisson-fault model. It is reckless to apply our method to an empirical test: it is difficult to explore all features of a power transformer based on empirical datasets due to the loss of the intensity \( x_n \), and hence we cannot evaluate the performance of the method. Through simulations, we may obtain observations of both \( X_n \) and \( M_n \).

In the section, we focus on simulation experiments about failures of a power transformer.

5.1. Dataset. The dataset should be as similar as possible to real world dataset. In the following, we start from a model used in [12] for transformers to deduce reasonable \( a, b, e_n \),
and $X_1$ for (3). In the model used in [12], the relative average lifetime

$$t = A_2^{-2(T-98)/6} = 82371.3A_2 \exp (-0.1155T).$$

(28)

Here $T$ is the hottest spot temperature of a power transformer in degrees Celsius. $A_2$ is the average lifetime of a transformer at $T = 98$. A typical $A_2$ is two decades. Comparing (28) to (6), we have that $k = 0.1155$. It follows from (7) that

$$a = e^{-\Delta k}.$$  

(29)

Here the benchmark interval $\Delta$ is a change in temperature between two successive working conditions. In the following simulations, we choose $\Delta = 3$ and $a = 0.707$. Besides, the level $n$ is from 1 to 9, corresponding to the temperature $T = 134, \ldots, 110$, respectively. That is, the total number of levels is $N = 9$.

Now we turn to the choices of $b$ and $\epsilon$. $b$ represents the bedrock errors which does not appear in the Arrhenius model. As the Arrhenius model is a widely accepted model, $b$ is a small number. In the following simulations, we use different settings 0.01, 0.02, 0.05, and 0.1 for $b$. $\epsilon_n$ represents a difference of the behavior between a real transformer and the Arrhenius model. It is also a small random variable. In following simulations, we assume that $\epsilon_n$ has a normal distribution $N(0, 0.01)$.

$X_1$ is the average of the number of failures at level 1, which corresponds to the most severe working condition. We assume that $T = 134$ at level 1. Then the average lifetime $t = 0.0156A_2$ from (28). For a typical $A_2 = 20$, it follows from (5) that

$$X_1 = \frac{dp}{0.312}.$$  

(30)

$dp$ may have a wide range for choosing. In the following simulations, we set $dp = 8$. Besides, the model allows a small noise for $X_1$. Hence, we assume that $X_1$ has a normal distribution $N(25.64, 0.01)$.

The simulation of $M_n$ based on $X_n$ is according to the classical method to simulation of a Poisson process; see [13] for example.

5.2. Simulation Results. Based on datasets provided by the method introduced in Section 5.1, we perform three experiments to test performances of the model and the method introduced in the above two sections.

In the first experiment, we compare estimators $\hat{a}$ and $\hat{b}$ with $a^*$ and $b^*$. For different value of $b$, we repeat simulations 400 times to obtain 400 datasets. We present the means and standard errors of $\hat{a}$ and $\hat{b}$ of these datasets in Table 1. We also present the means and standard errors of $a^*$ and $b^*$ in Table 2. Noting that the true value of $a$ is 0.707 from (29), Tables 1 and 2 show that the standard errors of both $\hat{a}$ and $a^*$ are small, but the mean of $\hat{a}$ is closer to the true value than that of $a^*$. Moreover, from these tables, we may conclude that the performances of both the mean and the standard errors of $\hat{b}$ are better than that of $b^*$.

In the second experiment, we set $b = 0.01$ and run 400 simulations. For every dataset, based on the estimators $\hat{a}$ and $\hat{b}$ for the dataset, we deduce by our algorithm the expectations $E[X_n | M_1, \ldots, M_9]$ for $n = 1, \ldots, 9$. In Figure 1, we compare these expectations with $X_n$, $n = 1, \ldots, 9$. As there are 400 datasets of $X_n$, the benchmark curve is the mean of $X_n$. We may conclude from Figure 1 that the mean of expectations $E[X_n | M_1, \ldots, M_9]$ is close to the true value and the standard errors are small.

The third experiment concerns the time complexity of the method. We randomly selected 12 groups of nonnegative $m_1, \ldots, m_9$ such that $L = m_1 + \ldots + m_9$ is 3, 4, \ldots or 14. We use three methods to calculate $E[X_5 | M_1, \ldots, M_9 = m_9]$; the algorithm in Section 4.3, formula (18), and a Monte Carlo simulation suggested in Section 4.1. The Monte Carlo simulation stops when runtime exceeds half an hour or the standard error is less than 0.1. The outputs presented in Figures 2 and 3 show the corresponding runtime.

For small $L$, the result of Bayes estimate are similar for three methods. The performance of the Monte Carlo simulation is not good for $L \geq 11$ because its runtime exceeds its limit. On the other hand, the runtime of formula (18) is unacceptable for $L = 9$. In fact, we do not test (18) for $L \geq 10$ due to its long runtime at $L = 9$.
6. Conclusions

In this paper, we proposed a Poisson-fault model to describe the failure of a power transformer at different working conditions. The model is an extension to the Arrhenius model. We modified the traditional time series techniques and proposed a framework to calculate parameters of the model. After exploring three methods with different time complexity, we established a quick algorithm to estimate the failure rate of a power transformer. The algorithm is tested on simulation datasets for a power transformer. By simulation studies, we have shown that the estimators we proposed have better performance than the estimators based on the classical method of time series. Moreover, the runtime of our method is the best among three methods. The estimator of the intensity is near the true value and has a small standard error.

From the viewpoint of an engineer, we use a slightly ideal assumption that the change of two successive working conditions is constant. If the change is not a constant, we do not know whether the runtime of the proposed method is still acceptable to applications. On the other hand, from the viewpoint of statistician, it is an unnatural assumption that a parameter of the model has a prior normal distribution, because the posterior distribution is not a normal distribution. To extend the proposed method to a case with general change and to establish a natural prior distribution for a parameter are future research directions. Ongoing work is investigating.

Appendices

A. Proof of Theorem 1

Lemma A.1. Assume that a stochastic process \( \{Y_n : n = 1, 2, \ldots \} \) is adapted to filtration \( \mathcal{F}_n = \{\mathcal{F}_n : n = 1, 2, \ldots \} \), \( E[Y_n | \mathcal{F}_{n-1}] = 0 \), and \( \sup_n EY_n^2 < \infty \). With probability 1, one has that \( (1/N) \sum_{n=1}^N Y_n \rightarrow 0 \).

Proof. For convenience, we use the symbol a.e. \( P \) for an event occurring with probability 1. We consider a stochastic process \( \{Z_n : n = 0, 2, \ldots \} \) such that \( Z_0 = 0 \) and \( Z_n = Z_{n-1} + (1/n)Y_n \).
for \( n \geq 1 \). It is obvious that \( \{Z_n\} \) is adapted to the filtration \( \mathcal{F} \) and
\[
E[Z_n \mid \mathcal{F}_{n-1}] = Z_{n-1} + \frac{1}{n} E[Y_n \mid \mathcal{F}_{n-1}] = Z_{n-1}, \quad \text{a.e. } P.
\]
(A.1)
That is, \( \{Z_n\} \) is a \( \mathcal{F} \)-martingale.
According to the property of a martingale,
\[
E[Z_N^2] = \sum_{n=1}^{N} E(Z_n - Z_{n-1})^2 + E[Z_0^2] = \sum_{n=1}^{N} \frac{EY_n^2}{n^2} \leq \left( \sup_n EY_n^2 \right) \sum_{n=1}^{N} \frac{1}{n^2} < \infty.
\]
(A.2)
Then, it follows from Doob’s convergence theorem (see, e.g., [17, 22]) that there is a random variable \( Z_{\infty} \) such that \( P(|Z_{\infty}| < \infty) = 1 \) and \( Z_N \to Z_{\infty}, \) a.e. \( P \) as \( N \) tends to \( \infty. \) Moreover, according to Kronecker’s lemma [23] it follows from \( Z_N = \sum_{n=1}^{N} (Y_n/n) \) that as \( N \) tends to \( \infty, \)
\[
\frac{1}{N} \sum_{n=1}^{N} Y_n = \frac{1}{N} \sum_{n=1}^{N} \frac{Y_n}{n} \to 0, \quad \text{a.e. } P.
\]
(A.3)
Based on Lemma A.1, we can present the proof of Theorem 1 by constructing different stochastic processes and filtrations.

Case 1 ((1/(N − 1)) \( \sum_{n=1}^{N-1} (M_n - X_n) \to 0, \) a.e. \( P \).) We study filtration \( \{\mathcal{F}_n \mid \sigma[X_{n+1}] : n = 0, 1, \ldots\} \) and a stochastic process \( Y_n = M_n - X_n. \) It follows from the definition of a Poisson-faults model that \( E[M_n \mid \mathcal{F}_{n-1} \cup \sigma[X_n]] = X_n \) and \( E[M_n^2 \mid \mathcal{F}_{n-1} \cup \sigma[X_n]] = X_n^2 + X_n. \) Hence, \( E[Y_n \mid \mathcal{F}_{n-1} \cup \sigma[X_n]] = 0 \) and
\[
\sup_n EY_n^2 = \sup_n E \left[ Y_n^2 \mid \mathcal{F}_{n-1} \cup \sigma[X_n] \right] = \sup_n \left( E \left[ M_n^2 \mid \mathcal{F}_{n-1} \cup \sigma[X_n] \right] - 2X_n E[M_n \mid \mathcal{F}_{n-1} \cup \sigma[X_n]] + X_n^2 \right) = \sup_n EX_n < \infty.
\]
(A.4)
Therefore, the result follows from Lemma A.1.

Case 2 ((1/(N − 1)) \( \sum_{n=1}^{N-1} (M_n^2 - M_n - X_n^2) \to 0, \) a.e. \( P \).) Considering the filtration \( \mathcal{F} \) and a stochastic process \( Y_n = M_n^2 - M_n - X_n, \) we can obtain the result by mimicking Case 1.

Case 3 ((1/(N − 1)) \( \sum_{n=1}^{N-1} (M_{n+1}M_n - X_{n+1}X_n) \to 0, \) a.e. \( P \).) Firstly, we study filtration \( \{\mathcal{F}_n = \mathcal{F}_{2n} : n = 0, 1, \ldots\} \) and a stochastic process \( Y_n = M_{2n+1}M_n - X_{2n+1}X_n. \) We have
\[
E[Y_n \mid \mathcal{F}_{2n+1} \cup \sigma[X_{2n+1}]] = X_{2n+1} (M_{2n} - X_{2n}).
\]
(A.5)
Substituting \( X_{2n+1} = aX_{2n} + b + \epsilon_{2n+1} \) into the above equation and noting that \( \epsilon_{2n+1} \) is independent with respect to \( \mathcal{F}_{2n}, \) we have that
\[
E[Y_n \mid \mathcal{F}_{2n}] = E \left[ E \left[ Y_n \mid \mathcal{F}_{2n+1} \cup \sigma[X_{2n+1}] \right] \mid \mathcal{F}_{2n} \right] = (aX_{2n} + b) (M_{2n} - X_{2n}).
\]
(A.6)
Hence, as \( \mathcal{F}_{n-1} = \mathcal{F}_{2(n-1)} \subset \mathcal{F}_{2n-1} \cup \sigma[X_{2n}] \subset \mathcal{F}_{2n} \) and \( E[M_{2n} - X_{2n} \mid \mathcal{F}_{2n-1} \cup \sigma[X_{2n}]] = 0, \) we have that
\[
E\left[ Y_n \mid \mathcal{F}_{n-1} \right] = E \left[ (aX_{2n} + b) \right] \times E \left[ M_{2n} - X_{2n} \mid \mathcal{F}_{2n-1} \cup \sigma[X_{2n}] \right] \mid \mathcal{F}_{n-1} = 0.
\]
(A.7)
Moreover, similarly to Case 1, we can obtain \( \sup_n Y_n^2 < \infty. \) Then, it follows from Lemma A.1 that
\[
\frac{1}{N-1} \sum_{n=0}^{N-1} (M_{2n+1}M_{2n} - X_{2n+1}X_{2n}) \to 0, \quad \text{a.e. } P.
\]
(A.8)
Similarly, for filtration \( \{\mathcal{F}_n = \mathcal{F}_{2n} : n = 1, 2, \ldots\} \) and a stochastic process \( Y_n = M_{2n}M_{2n-1} - X_{2n}X_{2n-1}, \) we can obtain that
\[
\frac{1}{N-1} \sum_{n=1}^{N-1} (M_{2n}M_{2n-1} - X_{2n}X_{2n-1}) \to 0, \quad \text{a.e. } P.
\]
(A.9)
The result follows from (A.8) and (A.9).

B. Proof of Theorem 2
For a sample \( x = (x_1, \ldots, x_N) \) of \( X = (X_1, \ldots, X_N), \) we have that
\[
P \left( X_1 \in (x_1, dx_1), \ldots, X_N \in (x_N, dx_N) \right) = f(x) \prod_{i=1}^{N} dx_i,
\]
where \( f(x) \) is the prior pdf of \( X. \) That is,
\[
f(x) = (2\pi)^{-N/2}|B|^{-1/2} \exp \left( -\frac{1}{2} (x - \mu)^T B^{-1} (x - \mu) \right).
\]
(B.2)
According to our model, the conditional probability of \( M_1 = m_1, \ldots, M_N = m_n, \) given that \( X = x, \) is
\[
P \left( m \mid x \right) = \prod_{i=1}^{N} \left( \frac{x_{m_i}}{m_i!} \exp(-x_i) \right).
\]
(B.3)
Hence the posterior pdf of \( X \) is
\[
g(x \mid m) = C f(x) P \left( m \mid x \right),
\]
(B.4)
where the constant \( C \) satisfies
\[
1 = \int_{\mathbb{R}^N} g(x \mid m) dx_1 \cdots dx_N.
\]
(B.5)
Comparing the preceding equation with the pdf \( f_U \) of \( U = (U_1, \ldots, U_N) \), we have that

\[
g(x \mid m) = C' f_U(x) \prod_{i=1}^{n} x_i^{m_i}.
\] (B.6)

Then, it follows from (B.5) that

\[
1 = \int_{R^N} C' f_U(x) \prod_{i=1}^{n} x_i^{m_i} dx_1 \cdots dx_N = C' E \prod_{i=1}^{n} U_i^{m_i}.
\] (B.7)

That is,

\[
C' = \left( E \prod_{i=1}^{n} U_i^{m_i} \right)^{-1}.
\] (B.8)

Moreover, it follows from (B.6) and (B.6) that

\[
E[X_i \mid M_1 = m_1, \ldots, M_N = m_N] = \int_{R^N} g(x \mid m) dx_1 \cdots dx_N
\] (B.9)

\[
= C' \left( E U_i \prod_{i=1}^{n} U_i^{m_i} \right).
\]

The results follow from the preceding and (B.8).

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgment**

The authors would like to acknowledge the funding of the National Natural Science Foundation of China (Grant no. 70971109).

**References**


