IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 20 November 2023, accepted 29 November 2023, date of publication 7 December 2023, date of current version 13 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3340515

RESEARCH ARTICLE

Predictive Reliability Analysis of Power Distribution Systems Considering the Effects of Seasonal Factors on Outage Data Using Weibull Analysis Combined With Polynomial Regression

YUTTANA DECHGUMMARN[1,2], (Member, IEEE), PRADIT FUANGFOO[2], AND WARAYUT KAMPEERAWAT[1]

[1]Department of Electrical Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
[2]Provincial Electricity Authority (PEA), Bangkok 10900, Thailand

Corresponding author: Warayut Kampeerawat (warayut@kku.ac.th)

This work was supported by the Provincial Electricity Authority (PEA), Bangkok, Thailand.

ABSTRACT In this study, the reliability of power distribution systems is analyzed using a novel strategy of predictive reliability analysis based on the lifetime failure rate cycle in a bathtub curve shape and considering the standard Weibull distribution to determine the trend of the failure rate in each period using median rank regression for parameter estimation. The proposed strategy consists of three processes. The first process involves separating the external and internal factors that influence power outages in the power distribution system from the seasonal multimodal shape in the empirical distribution of the dataset using a bisection algorithm of residuals of polynomial regression. Second, clustering and characterization of each component in the power distribution system according to the condition of the total factor bathtub curve leads to the introduction of the use of shape parameters as the total factor deterioration index (TFDI) with linear regression trends of log scale shape parameters of the useful period. A simple approximation of the system's overall total factor bathtub curve using a sixty-year forecast is the final process presented that can be used in reliability planning to address lifecycle risks. The actual time-to-outage dataset between 2015 and 2020 of the Provincial Electricity Authority, Region 1, Northeastern Thailand, which covers the area of distribution line life in the three periods of the bathtub curve, was used as the test data. The numerical results obtained from the proposed process provide a comprehensive prediction of the reliability of the electrical distribution system for risk response planning. The results show the proportion and amount of internal deterioration versus external disturbances, helps to group components according to health and usability and prioritizes them according to risk. Furthermore, it clarifies the important moments of status transition. All of these factors make it possible to improve reliability in the right place at the right time. Every method that we have chosen to improve for use in analysis is simple, provides a clear visualization of every step, and can be used in practice.

INDEX TERMS Time to outage, failure rates, bathtub curve, standard Weibull distribution, predictive reliability, electric distribution system, polynomial regression, bisection algorithm.

I. INTRODUCTION

Every economic and social activity must continuously use electrical power. Even residential customers have this problem, reflected in the form of complaints at a higher level than other power quality problems. This has led the electric utility provider to establish numerical guarantee criteria for this issue that are acceptable to both sides. The economic and social impacts of a power outage are difficult to concretely assess. The topic of dealing with power outages is called

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Liu.

reliability analysis for power systems. The level of this problem is technically assessed in terms of frequency, duration, and the number of customers, while the value and amount of energy are used as alternatives. Collecting data on outage times can contribute to reliability evaluations in three types: assessment, prediction, and forecasting, which are useful for improved reliability planning.

Reliability analysis of power distribution systems is based on insights in the form of indices (SAIFI, SAIDI, etc.) obtained from the reliability assessment process. This is used along with descriptive statistical analysis of surrounding data to plan the restructuring of the electricity distribution network to increase the reliability of the system [1], [2]. These processes assume that the failure rate (λ) is a constant value or a time value only. The failure rate is an important characteristic of homologous components and is often regarded as time-varying in accordance with the Weibull probability distribution function model. All assumptions for the time-to-failure dataset are on the three trend lines that can be described by the Weibull probability distribution function, and the life cycle of the failure rate is described by the shape of the bathtub curve. This is useful for reliability prediction or condition-based reliability management as well as reliability in other engineering fields [3], [4], [5], [6], [7], [8].

Predictive reliability analysis for electrical distribution systems provides insights related to the lifespan of systems and components, as shown in Figure 1. This bathtub curve shows that total risk management is beyond the scope of maintenance work and extends to proactive planning from precommissioning agreements (quality testing and warranty) to periods when regular maintenance cannot prevent frequent failures. This is called the wear period. The goal of this risk management is to eliminate periods of high failure rates on both sides. It must be recognized and acted upon without a necessary cost assessment. This is the importance of the insights gained from predictive reliability analysis.

Many previous studies have presented arguments against this hypothesis, such as the inadmissibility of bathtub curve charts in the electronics industry and even in the field of electrical systems and components [9], [10]. Most studies note the presence of modal or multimodal in the empirical curve of the failure rate function or the probability density function (PDF) of the time to failure dataset, which can be obtained from various nonlinear methods such as the Kaplan-Meier estimator, histogram, kernel density estimation, and empirical cumulative distribution function [11], [12], [13], [14], [15], [16], [17], [18]. The results of analysis using the Weibull distribution on the time-to-failure dataset are monotonic functions. When analyzing the presumed lifetime failure data with the classical Weibull distribution model, the characteristics of the bathtub curves are not found. Therefore, improving the model to meet this need is a continuous area of research and development [4], [19], [20], [21].

In practice, the bathtub curve is an assumption of lifetime failure rates, and Weibull analysis is an important reliability tool used to predict trends in each period. The two-parameter Weibull distribution model (standard Weibull distribution) [20] is still the most commonly used in practice. Due to the simplicity of the estimation and interpretation methods, it is important to use only the shape parameter (β) and the scale parameter (α). The value of the shape parameter is an important condition for identifying the trend characteristics of three different failure rate values, consisting of a decreasing trend, a constant trend, and an increasing trend, which can be interpreted as whether a component or system is in the lifespan of the infant period, the useful period, or the wear-out period, respectively. This leads to a plan to respond with different management strategies for the three periods, such as burn in & warranty, corrective & preventive maintenance, and replacement & reallocation [19], [20], [21], as shown in Fig. 1. Conceptual details and recommendations for reliability predictions are provided in IEEE Std 1413.1-2002 [22].

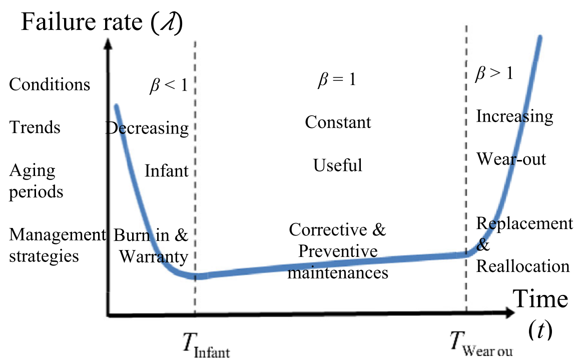


FIGURE 1. Predictive reliability analysis using the bathtub curve of the standard Weibull distribution.

Bathtub curves are used in analytical applications to manage assets over the lifetime of a component or system. In business, they also play an important role in the study of customer lifetime value and product life cycle. Bathtub curve shapes appear in power system resilience to measure the efficiency of power system recovery after severe or large-area damage with low probability but high impact. This chart, called the resilience curve or resilience trapezoid, shows that the bathtub curve is an operational chart that is not a data distribution model [23], [24], [25], [26], [27], [28].

There is a general understanding of the study topics: statistical survival analysis, engineering reliability, event history analysis in sociology, and duration analysis in economics are just a few different names. To analyze time-to-event datasets, expectations determine when and where we should worry. The hazard function (h), failure function (λ), and force of mortality (μ) have the same mathematical meaning. This function results in the shape of the bathtub curve. In contrast, the topics of statistical and psychrometric reliability refer to measurement reliability, which aims to assess measurement errors and suggest ways to improve testing to reduce errors.

The earliest survival analysis is found in John Gront's 1662 publication, ''Natural and political observations mentioned in the following index and made upon the bills of mortality.

It has been continuously developed and its effectiveness has been improved, and it has become an important tool that is popularly used in actuarial analysis, epidemiology, and demography [29], [30], [31].

A curve chart of the number of people who survived until the age of 86 out of 100 Christiaan Huygens people in 1669 shows important observations on the shape of the distribution of the data. This is clearly a parametric model that is one of the earliest survival curves [32]. Presenting survival charts has become a highly popular method of presenting insights using a simple estimation method. The Kaplan–Meier nonparametric model has been widely applied (since 1958) in many academic fields, particularly in health sciences. Kaplan-Meier curves, log-rank test, and Cox proportional hazards regression have become popular survival analysis techniques for analyzing time-to-event datasets [33].

The Weibull distribution was invented in 1937, and Weibull analysis is now the most popular parametric model method to analyze component and system lifetime failure data in reliability engineering. Starting from the time to failure (TTF) dataset, fitting the Weibull distribution model through parameter estimation leads to the derivation of the cumulative distribution function (CDF or $F(t)$) and the probability density function (PDF or $f(t)$), the reliability function ($R(t)$), and the failure rate function ($\lambda(t)$). A combination of the three periods gives rise to a bathtub curve. Although there are many parameter estimation methods, only three of the most popular are used in practice: Weibull probability plot (WPP), median rank regression (MRR), and maximum likelihood estimation (MLE) [19], [20], [34], [35], [36].

Revisiting the beginnings of studies on this topic not only honors those involved but also means uncovering the right understanding to begin data analysis. It also includes clarification of the nature of the analytical dataset and the optimum timing of observation and recording for a large number of components with similar characteristics. The original datasets that gave rise to the reliability function analysis occurred in a system that was virtually shielded from external influences, which ultimately led to the analytical concept of time series decomposition with outage time series presented in this article. The verification of the suitability of the datasets used in previous studies presents several interesting points. For example, Aarset's (1987) 50-device failure time dataset [37] was used to develop a bathtub-shaped lifetime failure rate model. It should be noted that the maximum time value from the data (85 h) was observed over the actual lifetime. In addition, the bathtub shape appears as an empirical distribution function of the failure probability density function (PDF) and not as a shape on the failure rate function. However, this type of analytical work was not part of this study.

The analysis of time-to-event datasets seems to be closely related to time-series analysis [4], [38]. Both datasets have a time domain; the range of the time-series dataset is an exact value, whereas the range of the time-to-event dataset does not exist. Thus, the task of analyzing the time-to-event dataset begins by mapping the data to the appropriate probabilities using the desired probability distribution model to form the time-series dataset. However, pairs with sequence values can also be converted into time-series datasets. Therefore, any data analysis technique applicable to time series can be applied to time-to-event datasets. A basic decomposition analysis [39] should be considered first. To measure the goodness of fit of the model (goodness-of-fit test), it is preferable to use the distance between the empirical cumulative distribution function (eCDF) [18], [40], [41], [42] and the cumulative distribution function of the selected distribution. such as the Kolmogorov-Smirnov test, Anderson-Darling test, root mean square error (RMSE), coefficient of determination (R^2) [43], [44], and eCDF derived from the estimation. The goodness-of-fit test is useful for comparing performance between different models and is used to analyze data with a single model type.

Traditionally, reliability prediction for power distribution systems is defined as finding the identity parameters of similar components, such as the failure rate, repair rate, and annual outage time (λ, r,U), to be used for reliability assessment with analytic methods. This is similar to circuit analysis with resistance, inductance and capacitance (R, L, C), and it utilizes a fixed-time or constant failure rate [3]. However, it is different from other areas of reliability engineering that focus on analyzing the lifetime failure rate of components or systems [22]. The predictive reliability analysis of power distribution system outage datasets in previous research was meant to predict the consequences after changes in the distribution system structure using a constant failure rate for analysis, which is called predictive reliability assessment [45], [46], [47], [48], [49], [50], [51], [52]. While the predictive reliability analysis recommended by IEEE Std 1413.1-2002 for electrical distribution systems has not been clearly found in previous research, there have been only a few attempts to use this principle in the analysis of components such as transformers and cables [38], [53], [54], [55], [56], [57], [58], [59], [60].

There are three major problems in applying Weibull analysis or reliability function analysis to power-outage data.

First, the reason that taking time to outage data to analyze through survival analysis or in the form of a reliability function is not very popular in practice is that most of the failure rates of components and systems are caused by external factors rather than internal wear and tears on the components and the system itself. The interpretation and utilization of traditional analytical results does not make much sense.

Second, for distribution systems, there are a large number of components and their actual lifetimes are similar, which often results in consistently close estimates of the shape parameter (β) of each component. It is difficult to find the separation point of the shape parameter (β) to group the components according to the bathtub curve condition.

Predictive reliability indices for power distribution components and systems should be introduced into practice to assess their ability to use continuously. For example, an index called the health index for reliability of a components [61], [62] that measures the change in different materials in that component and predicts its deterioration. similar to the deterioration index in medicine [63], [64] or the lifetime capability index in a production lots from industrial plants [65]. Through proper interpretation, these indices can be applied as predictive reliability indices.

Third, there is no clear method for combining the failure rate function curves of the three periods into a complete bath curve for predictive reliability management planning for power distribution systems.

In this study, the reliability of the power distribution system is analyzed using predictive reliability analysis on the concept of bathtub curves and median rank regression for parameter estimation of the standard Weibull distribution (two-parameter), which is used to find trends in failure rates to solve the above three problems. A three-process analytical strategy is presented that consists of the following: first, an analysis to separate the external and internal factors that influence outages in the power distribution system using nonlinear optimization with a bisection algorithm of residuals of polynomial regression; second, the analysis clusters and identifies the conditional features of the combined factor bath curves of each component in the power distribution system using linear regression of the natural log scale of the shape parameter (ln β), which leads to the introduction of the shape parameter (β) as a useful total factor deterioration index (TFDI). Finally, a simple approximate charting of the tubular curve of the system with a 60-year forecast is performed. In every analysis process, the results are interpreted to formulate risk management strategies. The proposed process is tested with a time-to-outage dataset between 2015 and 2020 of the Provincial Electricity Authority, Region 1, northeastern Thailand, which covers the area of distribution lines that complete three periods of the bathtub curve.

For the parameter estimation process of the proposed strategy, median-rank regression (MRR) is selected to estimate the parameters of the standard Weibull distribution because MRR is a more deterministic method than the Weibull probability plot. MRR can provide estimated results that are equivalent to maximum likelihood estimation (MLE) and still provide acceptable results when the amount of data is small. Moreover, MRR is easy to implement in practice. Polynomial regression is used to estimate the trend line of the outage time datasets to analyze separate outage datasets caused by internal and external influences. The main reasons for using polynomial regression are as follows. Polynomial regression is an optimal empirical alternative model that can be observed in the chart of the outage time function. The results obtained from polynomial regression show that the original curvature shape of the dataset is preserved and can be further adjusted to achieve a trend line representing the most noticeable seasonal excess. Polynomial regression has simple calculation steps and is suitable for implementation in practice.

The novelty of our proposed strategy can be summarized as follows:

1) A novel technique for predictive reliability analysis of a power distribution system is presented. The proposed strategy can provide the results of risk management and estimation of the bathtub curve for both the overall system and individual components.
2) A novel technique for separating internal and external factors from outage data is presented based on trend analysis with polynomial regression.
3) A novel total factor deterioration index (TFDI) is presented for simple and effective risk management.

The subsequent sections of this paper consist of the following: the theorems adapted for use in this paper are presented in section II, a step-by-step analysis of the three processes is presented in section III, practical examples with real data are presented in section IV, and the conclusion of the proposed predictive reliability analysis strategy of the power distribution system is presented in section V, which is the final part of this paper.

II. MODELS AND ALGORITHMS FOR PREDICTIVE RELIABILITY ANALYSIS

The predictive reliability analysis for power distribution systems presented in this paper uses bath curve charts of failure rates as the main application to manage lifecycle risk and find trends in each period. This analysis is performed using Weibull analysis, which uses median-rank regression to estimate parameters because it is simple and practical and works well when the amount of data is small [66], [67], [68]. In addition to extracting as much insight as possible, a polynomial regression model and bisection algorithm are used to exclude outages caused by external disturbances, where the linear regression model is used for conditional clustering of components on the bathtub curve. In the following, these models and algorithms are presented in principle, and the way they can be utilized to achieve the objectives is described as follows.

A. STANDARD WEIBULL DISTRIBUTION

The two-parameter Weibull distribution is called the standard Weibull distribution because it is the most popular in practice [20] and has only two main parameters: the shape parameter (β) and the scale parameter (α). The failure rate trend can be determined using the value of the shape parameter. Below is the mathematical model of the Weibull distribution function.

For $t \geq 0$, $\beta > 0$, and $\alpha > 0$, t is the failure time, β is the shape parameter, and α is the scale parameter.

The probability density function (PDF) is given by

$$f(t; \alpha, \beta) = \left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \qquad (1)$$

The cumulative distribution function (CDF) and the reliability function are given by

$$F(t;\alpha,\beta)=1-\exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \quad (2)$$

$$R(t;\alpha,\beta)=1-F(t)=\exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \quad (3)$$

The hazard or failure rate function is given by

$$h(t;\alpha,\beta)=\lambda(t;\alpha,\beta)=\frac{f(t)}{R(t)}=\left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1} \quad (4)$$

Utilization of parametric statistical models: The model is optimized to fit the data using several methods for optimizing parameter estimation. In this section, we review and discuss the three most popular methods used in practice: maximum likelihood estimation (MLE), Weibull probability plot (WPP), and median rank regression (MRR).

1) MAXIMUM LIKELIHOOD ESTIMATION

This approach is accepted in computational results when the amount of data is sufficiently large. The solution to this problem is an optimization model with a maximum of log likelihood function built from the probability density function (PDF), which is used as an objective function. $B>0$, $\alpha>0$ is used as a subject to a constraint where the partial derivative equal to zero together with a numerical computational method is used as an optimization method to determine the parameters. Although it does not start from an estimate of the empirical distribution, to check its fit with real data, this empirical distribution function is used as a reference value in the calculations and comparison. The details of the method can be found in references [20], [69], [70], and [71]. The final result from Cohen's analysis is another good alternative, as shown in the following equation.

For estimated β :

$$\left[\frac{\sum_{i=1}^{i=n}x_i^{\beta}\ln x_i}{\sum_{i=1}^{i=n}x_i^{\beta}}-\frac{1}{\beta}\right]=\frac{1}{n}\sum_{i=1}^{i=n}\ln x_i \quad (5)$$

and for estimated α:

$$\alpha=\sum_{i=1}^{i=n}\frac{x_i^{\beta}}{n} \quad (6)$$

2) WEIBULL PROBABILITY PLOT

Because the popularity of statistical model performance is often measured with reference values from the empirical cumulative distribution function (eCDF), it is easier to initiate the process of parameter estimation from this function estimation than to begin with the log-likelihood function. For the Weibull distribution, Benard's median rank approximation is commonly used to estimate this value, where $i=1, 2, 3, \ldots, n$ is the sequence number of the failure-time data in ascending order. eCDF or $F(t_i)$ can be determined using the following equation:

$$F(t_i)=\frac{i-0.3}{n+0.4} \quad (7)$$

Equation (2), when properly executed with the log-log transform, yields a linear relationship model, as shown in Equation (8), which can be used to generate a log-log-scale grid on paper to estimate the parameters of the Weibull function. In the Weibull probability plot paper (WPPP), the slope of two or more approximate coordinates (F, t) is the shape parameter (β) and the intersection point of this line, and the straight line $F(t)=63.21\%$ is the value of the scale parameter (α). The final result of the log-log transformation is as follows:

$$\ln\left(\ln\left(\frac{1}{1-F(t)}\right)\right)=\beta\ln t-\beta\ln\alpha \quad (8)$$

Compared to a linear relationship

$$y=\mathrm{m}x+\mathrm{c} \quad (9)$$

this will obtain

$$y=\ln\left(\ln\left(\frac{1}{1-F(t)}\right)\right) \quad (10)$$

$$x=\ln t,\quad \mathrm{m}=\beta \text{ and } \mathrm{c}=\beta\ln\alpha \quad (11)$$

This method of estimating parameters, although highly variable, is sufficiently useful when performing on-site maintenance tasks in which only WPP paper and pens are available for reliability prediction. Therefore, WPP remains popular today.

3) MEDIAN RANK REGRESSION

The results from (8)–(11) provide an alternative for the need to estimate β and α as accurately as possible using the linear regression least square error technique. This estimate can be obtained using the following equation:

$$\beta=\frac{n\sum_{i=1}^{i=n}x_iy_i-\sum_1^n x_i\sum_1^n y_i}{n\sum_{i=1}^{i=n}x_i^2-\left(\sum_{i=1}^{i=n}x_i\right)^2} \quad (12)$$

$$\alpha=\exp\left[\frac{\beta\sum_{i=1}^{i=n}x_i-\sum_{i=1}^{i=n}y_i}{n\beta}\right] \quad (13)$$

The combination of median rank approximation and linear regression analysis is known as median rank regression (MRR).

Parameter estimation for the Weibull distribution using the above and other methods is compiled and presented in [20], [69], [70], and [71]. The final result from the Weibull analysis has only one trend among the three periods of the bathtub curve. This is called a monotonic result and is shown in Fig. 2.

B. POLYNOMIAL REGRESSION

A special form of multiple linear regression or multivariate linear regression where the variables $\{x_1, x_2, x_3, \ldots, x_n\}$ are replaced by $\{x, x^2, x^3 \ldots, x^n\}$ for $i=1, 2, 3, \ldots, n$ becomes a form of a polynomial regression function. The

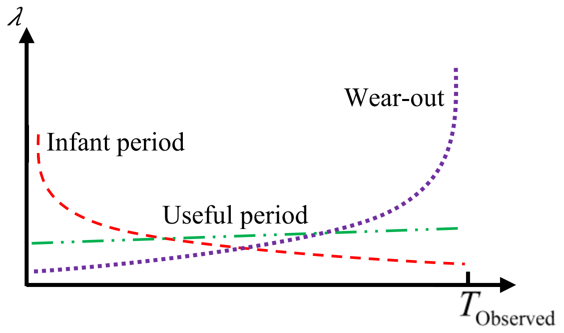


FIGURE 2. Monotonic pattern of the analytical results obtained by Weibull analysis.

model can be adapted to a wide range of nonlinear datasets, and the parameters $\{w_0, w_1, w_2, w_3, \ldots, w_n\}$ can be estimated using least-squares analysis. The model differences can be determined using (14)–(16): [72]

Simple linear regression function:

$$y = w_0 + w_1 x \tag{14}$$

Multiple linear regression function:

$$y = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 + \ldots + w_n x_n \tag{15}$$

Polynomial regression function:

$$y = w_0 + w_1 x + w_2 x^2 + w_3 x^3 + \ldots + w_n x^n \tag{16}$$

The model can be generalized to a matrix for variables **X**, response vector **Y**, error vector ε, and vector of parameters **w** as follows:

$$\mathbf{Y} = \mathbf{wX} + \varepsilon \tag{17}$$

With ordinary least squares estimation, the parameters can be estimated as follows:

$$\mathbf{w} = (\mathbf{X^T X})^{-1} \mathbf{X^T Y} \tag{18}$$

and the model performance is measured with the coefficient of determination R^2 defined as

$$\mathrm{R}^2 = 1 - \frac{\sum_1^n \left(y_i - \widehat{y}_i\right)^2}{\sum_1^n (y_i - \bar{y}_i)^2} \tag{19}$$

where y_i is the observed value, $\widehat{y}_i$ is the fitted value, and $\bar{y}_i$ is the arithmetic mean of the dependent variable Y for the i^{th} case. In this study, polynomial regression of the outage time function is used in conjunction with the bisection algorithm to separate the outage data influenced by intrinsic and extrinsic factors. The choice of the number of expressions for variable x is mainly determined by the most pronounced seasonal component together with the consideration of the value of R^2.

C. BISECTION ALGORITHM

To find the roots of continuous nonlinear equations in the form $F(z) = 0$, the bisection method is one of the most popular basic methods. A simple rule of thumb for determining the range of variables that has a solution is that the bounding function of both variables is $[-, +]$. Subsequently, this interval is divided into two parts and tested with the same principle. This is repeated until a solution is obtained that makes the value of the function equal to 0 or as close to 0 as acceptable. This process of dividing a variable range into two and then testing an objective function to select a range based on optimum conditions, known as the bisection algorithm, is widely adopted in solving optimization problems [73], [74], [75], [76]. In this study, this method was adapted to filter out outage data that are likely to be influenced by exogenous factors in a seasonal pattern using an objective function as a shape parameter and standard Weibull analysis (SWA) with MRR as a constraint. Variable (z) is the residual boundary of polynomial regression whose range is between the highest residual boundary [−m,m], which can be expressed as a mathematical model as follows:

Objective function:

$$\min \beta(z) \tag{20}$$

Subject to the constraints:

$$r(i) = T(i) - P(i) \tag{21}$$

$$z \in [-m, m] \tag{22}$$

$$F(t_l) = \frac{l - 0.3}{n + 0.4} \tag{23}$$

$$y = \ln\left(\ln\left(\frac{1}{1 - F(t_l)}\right)\right) \tag{24}$$

$$x = \ln t_l \tag{25}$$

$$\beta = \frac{n \sum_1^n x_l y_l - \sum_1^n x_l \sum_1^n y_l}{n \sum_1^n x_l^2 - \left(\sum_1^n x_l\right)^2} \tag{26}$$

where:

$T(i)$ is the outage time function,
$P(i)$ is the polynomial regression function of $T(i)$,
$r(i)$ is the residual of the polynomial regression function,
m is the maximum absolute value of $r(i)$,
$z \in [-m, m]$ is the residual boundary range used to extract time-to-outage data influenced by seasonal external factors, and
t_l is the time-to-outage dataset in which exogenously influenced data are already trimmed in seasonality from the z-boundary condition, where $l \in \{1,2, 3\ldots, n\}$ is the sequence number of the outage time data, which is arranged from least to greatest.

The bisection algorithm solution algorithm of this optimization model begins after the regression process to obtain the optimal $P(i)$, and the implementation of this algorithm follows the following sequence:

Step 1: Evaluate $r(i)$ with (21) and obtain the maximum upper bound in $[0, m]$, which is used for the initial upper bound of z_0(0^{th} iteration).

Step 2: Divide the range by values between 0 and m with an appropriately predicted value or divide the range by half by setting it to obtain two upper bounds [0, a] and (a, m]. Find the middle value of each range as $b = (0 + a)/2$ and $c = (a + m)/2$, and then obtain the boundaries $[-b, b]$ and $[-c, c]$.

Step 3: Truncate the data from step 2 with the residual boundary values from $[-b, b]$ and $[-c,c]$, resulting in two attenuated datasets (t_{bl} and t_{al}). The obtained dataset is used to estimate the shape parameter (β) using (23)–(26). With smaller β values, this range provides a suitable solution. Then, it becomes the new upper bound range in step 1 ([0, a] or (a, m] replaces [0, m]), and the middle value becomes the new range divider used in step 2 (b or c replaces a).

Step 4: Repeat steps 1-3 until the lowest and unchanged β values are achieved. If all iterations are k cycles, the appropriate range is variable to. z_k

This algorithm will be demonstrated again with actual outage data in Section IV.

III. THE PROPOSED STRATEGY OF PREDICTIVE RELIABILITY EVALUATION

In this paper, the processes of predictive reliability evaluation for power distribution systems are presented based on the life cycle failure rate usage in bathtub curves and a two-parameter or standard Weibull distribution model. It is used to extract trends for three periods (infant, useful, and wear-out) using a MRR parameter estimation method to gain insights that can be used to plan support for each period. The analysis strategy is divided into three main processes to analyze the power outage dataset in the power distribution system. A life cycle risk management plan can be developed that consists of a separate analysis of outage events caused by internal and external factors using a bisection algorithm of residuals in polynomial regression, clustering components analysis of conditional health levels on individual component bath curves using linear regression for a useful period and estimating a simple bath curve of the overall power distribution system using a forecast that is out of range for 60 years. An overview of the strategies presented in this article is shown in Fig. 3, and the details of the three main processes are as follows.

A. SEPARATING OUTAGE DATA INFLUENCED BY INTERNAL AND EXTERNAL FACTORS

This process is proposed to analyze and assess the impact of internal and external factors on outages. The mixture of these influences gives a multimodal nature to the empirical distribution of time-to-failure (histogram) data using a nonlinear optimization with a bisection algorithm of residuals in polynomial regression. The results of the analysis can distinguish outage events that are influenced by internal degradation (trend) or external factors and disturbance characteristics.

The analysis process is illustrated in Fig. 4. This process is initiated by introducing the entire time-to-failure dataset into a median rank regression estimated Weibull analysis and comparing the resulting PDF model to the empirical distribution characteristics derived from the histogram to consider the consensus and accept the results obtained from the Weibull analysis. Simultaneously, a parallel process is performed to extract only the intrinsically influenced power-outage dataset

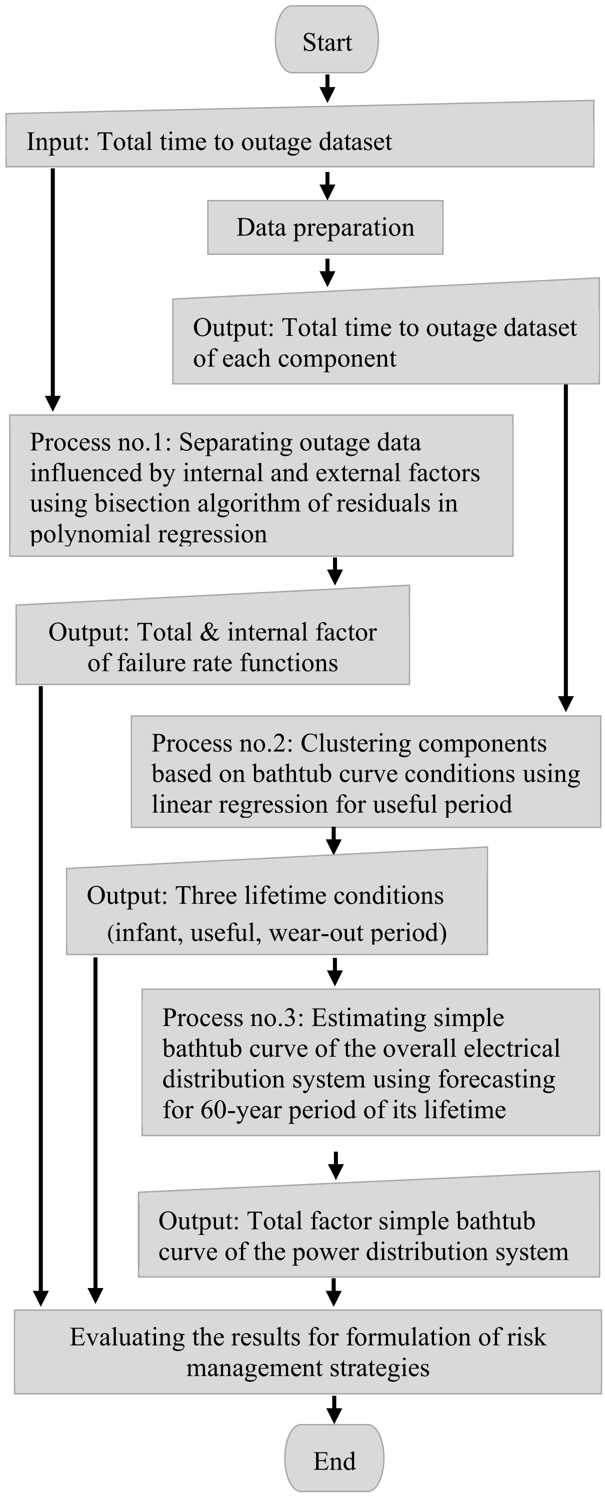


FIGURE 3. Flowchart of the proposed strategy of predictive reliability evaluations.

using the bisection algorithm of residuals of polynomial regression, and the resulting dataset (t_I) is included in the Weibull analysis. Finally, the results are compared, and measures to prevent power outages are found to be in line with the results obtained from the analysis.

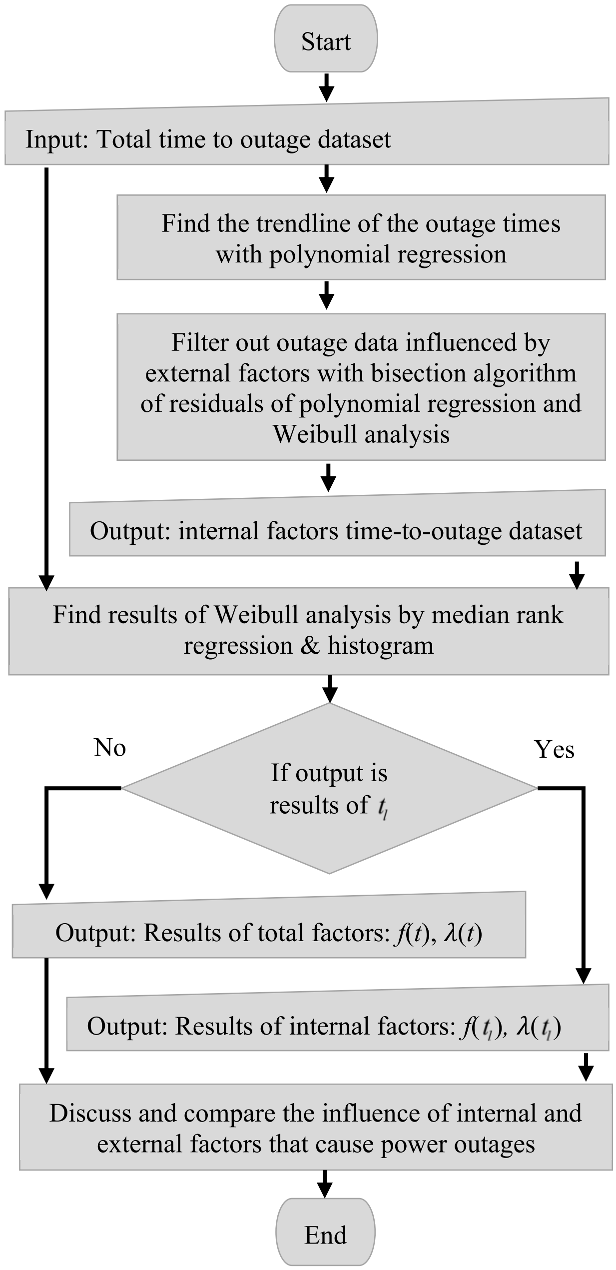


FIGURE 4. The process for filtering outage data is influenced by internal and external factors.

B. CLUSTERING OF COMPONENTS BASED ON BATHTUB CURVE CONDITIONS

This process presents a segmented analysis and characterizes the conditional features on the total-factor bath curves of each component in the power distribution system by determining the boundary β for the useful period using linear regression estimates of $\ln\beta$ to estimate the life period of the components and recommend the use of the shape parameter as a total factor deterioration index (TFDI).

The analysis process is illustrated in Fig. 5. This analytical process begins by importing the outage dataset of each component to determine the shape parameter of each component with Weibull analysis using MRR parametric estimation. Subsequently, a range of shape parameters (β) is performed to group the components within the useful period for all the factors that influence the outage. By linear regression analysis of $\ln(\beta)$, the components can be grouped into three groups (infant, useful, and wear-out periods). Finally, a contingency plan is developed according to the conditions of each group and prioritized by risk using the shape parameter as the total factor deterioration index (TFDI).

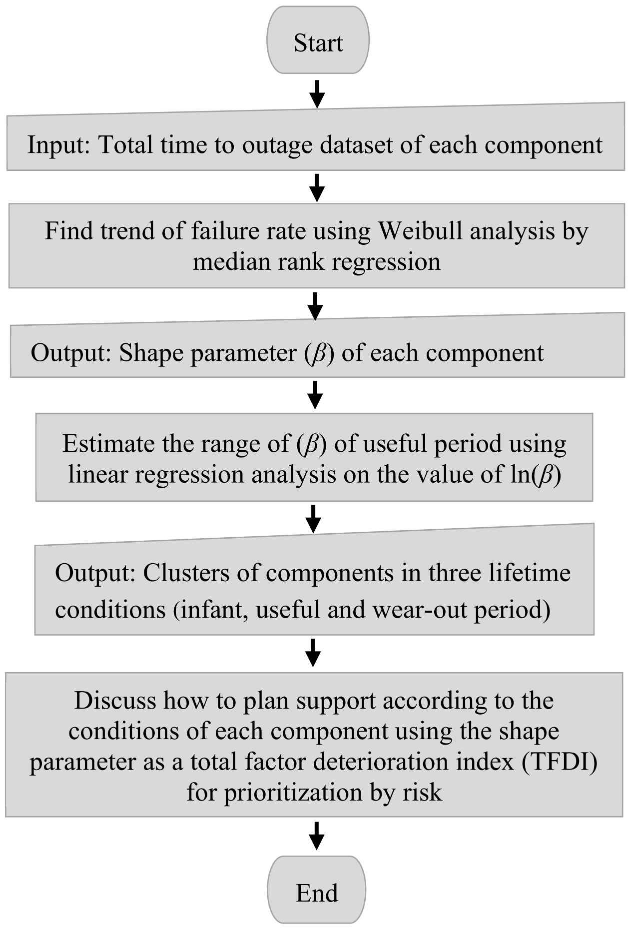


FIGURE 5. The process for evaluating the operating conditions of each component based on the bathtub curve.

C. ESTIMATING SIMPLE BATHTUB CURVE FOR REPRESENTATION OF THE OVERALL ELECTRICAL DISTRIBUTION SYSTEM

In this process, an analysis is presented to plot the bathtub curve of the system as a total factor by taking the outage datasets grouped in Section B. It is then introduced into the Weibull analysis by MRR to determine the shape parameters

(β) of each period and scale parameters (α) from all outage data. Then, the failure rates are estimated out of range as an estimated time close to the actual lifespan (60 years) of each period. The end of the infant period (T_{Infan}) and the beginning of the useful period are estimated at the same failure rate point, and the transition point from the useful period to the wear-out period ($T_{\text{Wear-out}}$) is estimated when the cumulative probability distribution function (CDF) of the useful period is equal to 1. These two transition points allow a simple bath curve approximation to be used as an overall representation of the power distribution system to plan reliability management based on life conditions. The analysis process is illustrated in Fig. 6.

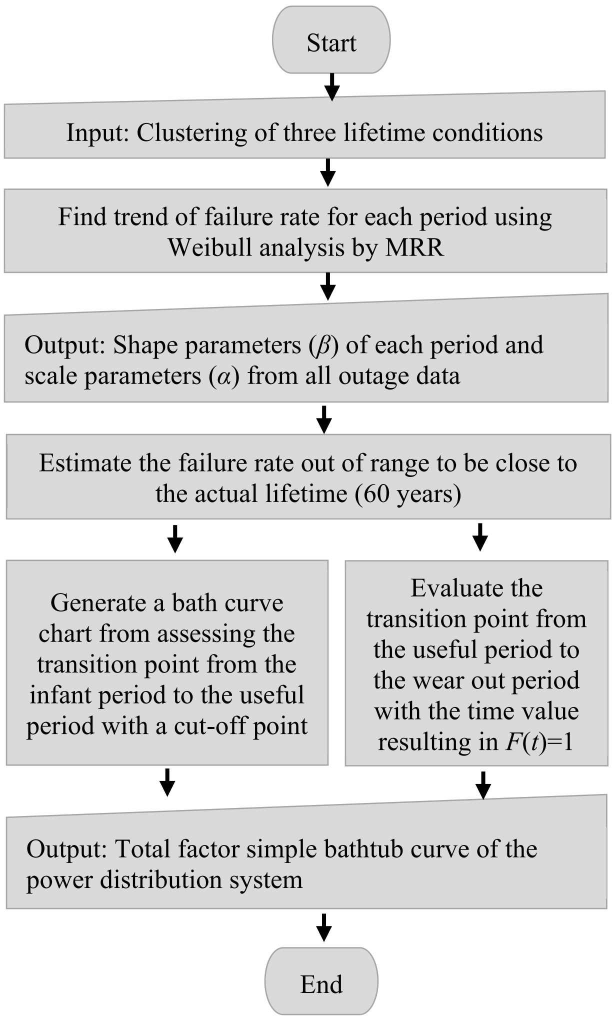


FIGURE 6. **The process for estimating a simple bath curve.**

While the outage datasets are used to analyze the proposed strategy, two conditions are considered. First, the data collection period must include at least two outage events and must not be too long for the data to be within one period of the bathtub curve. Second, the data collection area must cover all three periods with a conditionally qualified component on the bathtub curve.

Regarding some alternative models, Weibull family models are the main choice for analyzing time-to-event datasets to obtain bathtub curve conditional failure rate functions due to their ability to adapt to a wide range of data distributions, including normal distributions and exponential distributions. For models used to perform trend analysis on outage time datasets, various models can be used, such as parametric models (linear regression, polynomial regression, multiple regression, moving-average model, etc.) and nonparametric models (kernel regression, spline smoothing, locally weighted scatterplot smoothing, etc.). However, using them together to separate outage datasets based on internal and external factors has not attracted comparable research. We consider the choice and improvement of the model for use in this study based on the achievability of our proposed strategy and the simplicity of the numerical calculations that can be used in practice.

The essential validation processes to assess the performance, accuracy, and suitability of the proposed strategy are performed and discussed as follows.

1) The Weibull analysis results are validated by comparing their shape compatibility with the histograms of the empirical distribution.
2) MRR is used for evaluating the desired parameters in the proposed strategy. MRR has been used and validated by several previous studies of parameter estimation.
3) Polynomial regression is used for modeling the outage time series because it is flexible and adaptable to a variety of curve shapes and can clearly show the proportion of seasonal components while preserving the shape of the basic trend curve. In addition, the suitability of the model can be considered based on the clarity of the seasonal component and the R^2 value.
4) The performance of the newly improved bisection algorithm's solution is confirmed by convergence of the solution to a single optimal point.
5) The bathtub curve estimation method used in the proposed process is derived from the books recommended in reference no. [9]. The accuracy of the bathtub curve can be accurately considered from the information indicating the actual service life of each component. However, the lack of actual service information presents a difficulty for the validation of the bathtub curve in practice. The bathtub curve may be validated by considering the characteristics and relevant parameters.

IV. CASE STUDIES AND RESULTS

In this study, a predictive reliability analysis of power distribution systems is presented to plan the lifetime reliability of components and systems. This demonstrates the implementation of the proposed strategy with a time-to-outage dataset in the electricity distribution system of the Provincial Electricity Authority in Area 1, northeastern Thailand, between 2015 and 2020. It consists of 855 distribution lines, each of which has a number of outage data varying from 1 to 157, totaling 15,380 power outage records. The data collection

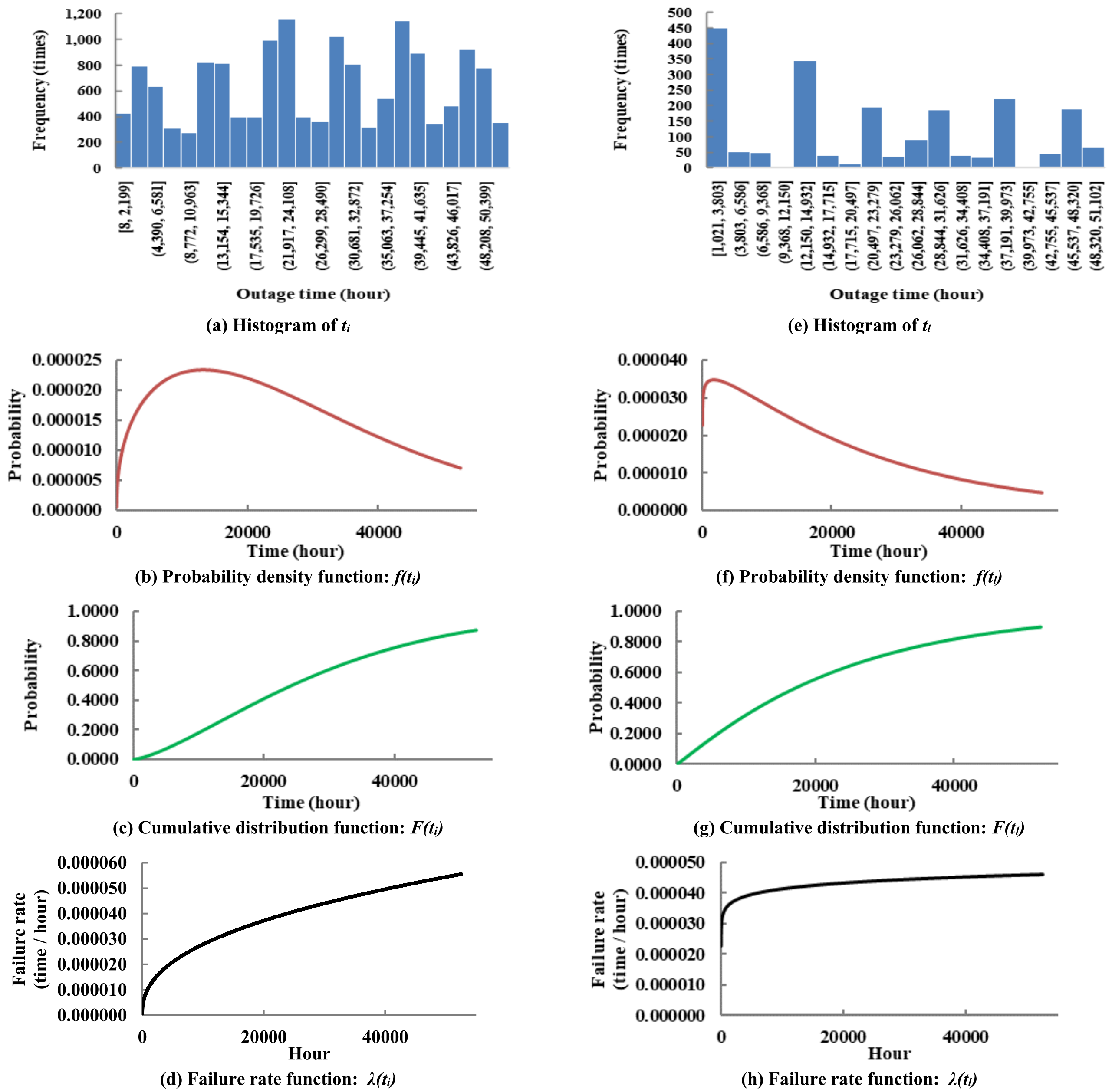


FIGURE 7. The Weibull analysis results of the total factor dataset: (a)-(d) and internal factor dataset: (e)-(h).

area has distribution lines covering all life stages from inception to deterioration, and some distribution lines can be identified as being more than 55 years old. Collecting data for six years makes it possible to determine that the data obtained for each distribution line are contained in only one range on the bathtub curve chart. The three proposed strategies for predictive reliability analysis of power distribution systems are described in detail below. This dataset is available for downloads, as shown in the Appendix.

A. PROCESS 1: SEPARATING OUTAGE DATA INFLUENCED BY INTERNAL AND EXTERNAL FACTORS

This analytical strategy uses all time-to-outage data to characterize the failure rate function relative to the duration of the bathtub curve by comparing the results of the empirical probability density distribution models from the histogram chart with the pattern obtained from the Weibull analysis.

The empirical distribution is multimodal due to the influence of exogenous factors on the seasonal pattern, which is different from the result obtained from the Weibull analysis (resulting $\beta = 1.414867$ and $\alpha = 31{,}509.48$), as shown in Fig. 7 (a)–(d). When Weibull analysis is performed on the seasonally excluded data (trend), the results were $\beta = 1.065085$ and $\alpha = 24{,}299.63$, and the distributions are consistent with the empirical distributions shown in Fig. 7 (e)–(f). A comparison of the failure rate function between the total factor and internal factor is shown in Fig. 7 (d) and (h), which shows the influence of external

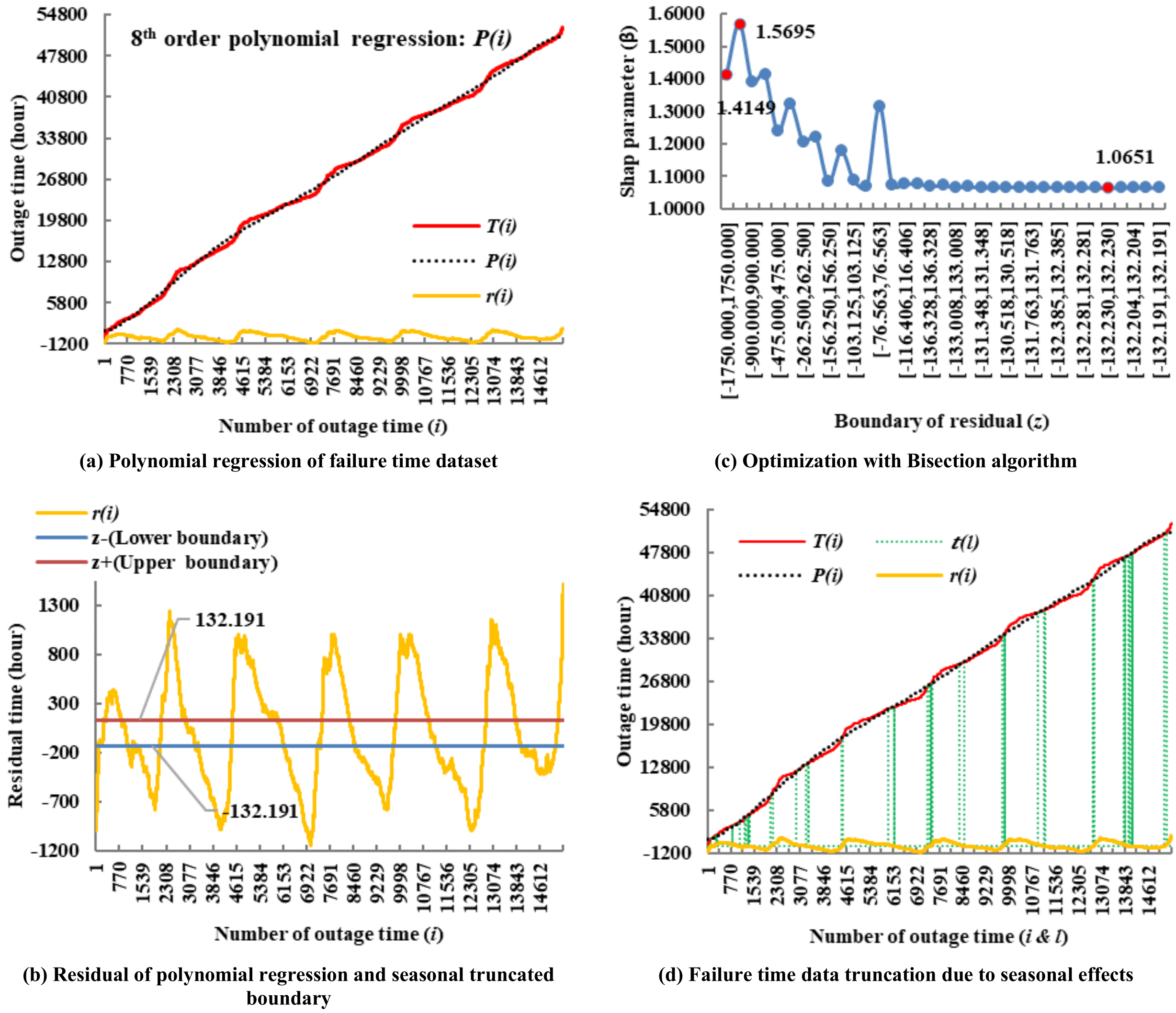


FIGURE 8. The chart shows each step of the process of extracting outages caused by seasonal external influences.

factors that can change the Weibull analysis results from the wear-out period to a useful period.

The trend and seasonal components are then separated using the bisection algorithm of residual polynomial regression. Polynomial regression is an optimal empirical alternative model that can be observed in the chart of the outage time function, as shown in Fig. 8 (a). The 8$^{\text{th}}$-order polynomial function is chosen because it clearly shows seasonal components as in (27) and the chart in Fig. 8 (a), which shows the curve of the outage time function, $T(i)$, polynomial function, $P(i)$, and residual function, $r(i)$. The results of linear regression (first-order polynomial regression) polynomial functions of orders less than the 8th show a clearly incomplete proportion of seasonal components and have worse R^2 values, while the results of polynomial functions of orders higher than 8 show that an excessively high fit to the dataset causes the proportion of seasonal components to be reduced.

$$\begin{aligned} P(i) =& (-1.50\times10^{-27})i^8+(8.10\times10^{-23})i^7+(-1.66\times10^{-18})i^6 \\ &+(1.54\times10^{-14})i^5+(-4.65\times10^{-11})i^4+(-1.94\times10^{-7})i^3 \\ &+(1.46\times10^{-3})i^2+1.45i+999.99 \\ \mathrm{R}^2 =& 0.99853 \end{aligned} \tag{27}$$

The bisection algorithm has a reasonable residual boundary of $z = [-132.191, 132.191]$, as shown in Fig. 8 (b) and (c). The variation in residual time $r(i)$ seems to be irregular in Fig. 8 (b). The strange fluctuation in the residual time is caused by the random behavior of power outages in the electrical distribution system, and the periodic behavior is influenced by seasonal effects. In Fig. 8 (d), the residual outage dataset is presented in terms of t_l.

Putting all time-to-outage datasets into the Weibull analysis allows the prediction of the overall power distribution system in the wear-out period. However, contradicting the empirical distributions from the histogram prevents the accuracy of the predictions from being confirmed. The repeated periodic multimodal appearance suggests the presence of seasonal external disturbances. By truncating the data to only data from internal influences, the analytical results are consistent with the empirical distribution. The overall power distribution system in the useful period is shown in Fig. 9.

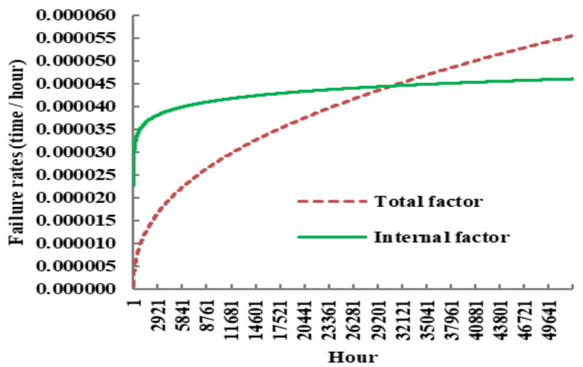


FIGURE 9. Comparing the failure rates of the t_i and t_l datasets.

From the evaluation results, if seasonal external factors cause 86.63% of power outages, 13.37% of the outages are caused by internal factors. Therefore, an appropriate response plan should focus on preventing external causes and ensuring that the power distribution system is durable to seasonal external factors. The factors for outage events in power distribution systems consist of internal and external factors. Internal factors may include age-related deterioration, manufacturing defects, and installation errors. External factors are disturbances from the environment, such as weather, trees, living things, and other objects. Considering internal and external factors in risk analysis, internal factor risk analysis is a reliability prediction task, while external risk analysis is often referred to as outage cause analysis in reliability assessment tasks.

There are two observations for the analysis of general time-to-failure datasets. First, the presence of a modal or multimodal pattern on the empirical distribution may be caused by interference from external factors, and it causes an error in the deterioration analysis of the system or components by the Weibull analysis method. Second, the useful period clustering with time-to-outage datasets considering both internal and external factors in practice for electric power distribution systems can be given the value of the shape parameter (β) as an interval instead of 1, as in the traditional model.

B. PROCESS 2: CLUSTERING OF COMPONENTS BASED ON BATHTUB CURVE CONDITIONS

The analysis presented in this strategy uses Weibull analysis on discrete datasets or power distribution line sections to predict the combined extrinsic and intrinsic degradation from 855-line section components. The β values of each component are obtained closely and continuously over a wide range. Therefore, it is necessary to perform clustering to separate the component groups into three periods based on bath curve characteristics using a natural log scale of shape parameter ($\ln\beta$) linear regression model to determine the β range of the useful period to separate the infant and wear-out periods. The results of β estimation with the MMR of each component and the estimation of β for the useful period are shown in (28)–(30) as follows.

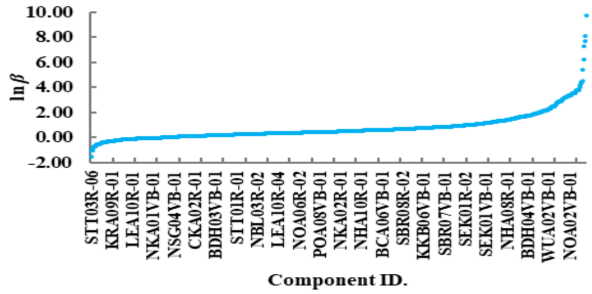
Linear regression of $\ln\beta$:

$$\ln\beta = 0.001862C_{\text{no.}} - 0.221318 \tag{28}$$

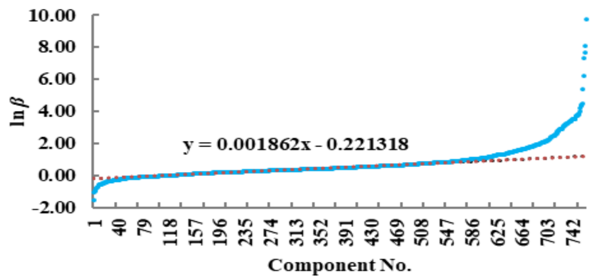
Range of useful period:

$$-0.161 \le \ln\beta \le 0.919 \tag{29}$$

$$0.851 \le \beta \le 2.508 \tag{30}$$



(a) The shape parameter chart in natural logarithmic scale (ln β) in ascending order and identified by component ID.



(b) The linear regression in the near-zero -slope range of the shape parameter in the natural logarithm scale (ln β), in ascending order and identified by component No..

FIGURE 10. Results of shape parameter estimation of each component and estimation of shape parameter ranges in the useful period.

The results of this analysis are shown in Fig. 10 (a)–(b), which clearly show the changes in the slopes of the curves for the three distinct intervals according to the condition of the bath curve. This near-zero slope difference range is the component within the useful period. By converting the variable from Component ID to Component No. ($C_{\text{no.}}$), a linear regression analysis is performed. The range of shape parameters for grouping components within this range can be estimated as in (30). When the shape parameters of each component are arranged in descending order, they are found to be consistent under different service life conditions. They are therefore suitable for use as well as other reliability indices (SAIFI, SAIDI, etc.), which may also be referred to as the total factor deterioration index (TFDI).

The results of the ratio parameter estimation for each component in the form of a histogram showing the normal distribution pattern are presented in Fig. 11.

TABLE 1. List of components in the wear-out period arranged in priority with TFDI.

No.	Component ID	β-TFDI	Factors (%)		No.	Component ID	β-TFDI	Factors (%)	
			External	Internal				External	Internal
1	TLI01BVB-01	16678.097119	100.0%	0.0%	35	BUA08R-02	22.087102	100.0%	0.0%
2	SCM1BVB-01	3252.407886	100.0%	0.0%	36	NOA09R-02	21.650836	100.0%	0.0%
3	UDC09R-01	2106.792149	100.0%	0.0%	37	BCA02R-04	20.826695	81.8%	18.2%
4	NQA03R-01	1491.787483	100.0%	0.0%	38	TLI05VB-01	19.438838	100.0%	0.0%
5	UDU03BR-01	510.926010	100.0%	0.0%	39	SOB05R-01	18.317139	100.0%	0.0%
6	KKD10VB-01	217.460322	100.0%	0.0%	40	SSK03VB-01	17.737466	60.0%	40.0%
7	SBR09R-03	89.033834	100.0%	0.0%	41	SOB09VB-01	17.715127	100.0%	0.0%
8	NQA03R-02	84.196948	100.0%	0.0%	42	POA01R-02	17.695252	100.0%	0.0%
9	STT01BVB-01	76.801542	100.0%	0.0%	43	SOB07VB-01	17.551433	88.9%	11.1%
10	WUA06VB-01	67.095512	66.7%	33.3%					
11	LEA06R-03	53.588380	100.0%	0.0%					
12	POA03R-01	47.777441	50.0%	50.0%					
13	KKD06VB-01	43.897010	100.0%	0.0%	173	BDH02R-01	2.778346	100.0%	0.0%
14	NQA04R-01	43.165308	100.0%	0.0%	174	PON05R-04	2.773059	80.0%	20.0%
15	PON05R-06	42.443170	100.0%	0.0%	175	SBR06R-03	2.765351	83.3%	16.7%
16	THA01TR-01	41.948867	66.7%	33.3%	176	NNA01R-04	2.756735	88.9%	11.1%
17	BPU06VB-01	40.464808	100.0%	0.0%	177	PFA03R-02	2.752450	91.4%	8.6%
18	2232000	34.742525	100.0%	0.0%	178	WWA04R-01	2.741465	75.0%	25.0%
19	WUA04R-01	34.387111	100.0%	0.0%	179	BUA09R-01	2.739948	91.7%	8.3%
20	KKD08VB-01	34.068352	85.7%	14.3%	180	KKC06R-04	2.735024	100.0%	0.0%
21	NQA04R-02	32.833598	88.5%	11.5%	181	KKC06VB-01	2.714588	100.0%	0.0%
22	LEA09R-02	32.802038	60.0%	40.0%	182	NBL03VB-01	2.690278	78.9%	21.1%
23	KKD09VB-01	32.522848	66.7%	33.3%	183	THA01R-01	2.672105	94.4%	5.6%
24	UDD06VB-01	32.093031	100.0%	0.0%	184	NNA08VB-01	2.636963	91.7%	8.3%
25	NQA07R-03	29.721631	100.0%	0.0%	185	SOA03R-04	2.627139	81.8%	18.2%
26	SOB09R-01	29.550351	100.0%	0.0%	186	SCM01R-01	2.619782	86.7%	13.3%
27	NOA02VB-01	28.985753	100.0%	0.0%	187	SEK01R-02	2.615338	87.7%	12.3%
28	WUA01R-03	27.584634	90.9%	9.1%	188	UDC03R-01	2.606501	93.3%	6.7%
29	SOB02VB-01	27.000262	80.0%	20.0%	189	SEK03R-01	2.586689	98.1%	1.9%
30	PFA03R-05	26.203458	50.0%	50.0%	190	CKA02R-03	2.563633	90.9%	9.1%
31	UDC09R-02	26.113843	75.0%	25.0%	191	UDC06R-01	2.552577	100.0%	0.0%
32	SEK06R-03	25.192051	100.0%	0.0%	192	NOA03VB-01	2.542005	92.9%	7.1%
33	WUA10R-02	23.941465	91.7%	8.3%	193	2222000	2.521287	92.9%	7.1%
34	NOA01R-02	23.535433	75.0%	25.0%	194	SCM02R-01	2.514791	90.9%	9.1%

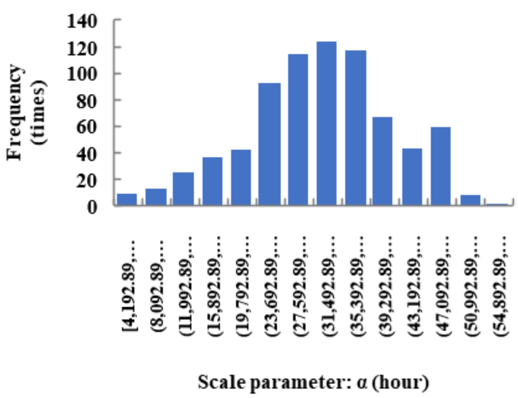


FIGURE 11. The scale parameter of each component obtained from Weibull analysis shows a normal distribution pattern.

The results of the analysis are as follows. Of the 855 components, 92 components had only one outage in a 6-year period. These components could not be analyzed by Weibull analysis, but it could be determined that these devices were in a useful period. A total of 59 components were in the infant period (6.90%), 602 were in the useful period (70.41%), and 194 were in the wear-out period (22.69%). Most of the components were within the range of normal use, with corrective and preventive protection. The proportions differed considerably between the 194 components that were already degraded and 59 that were just beginning to be used. Therefore, it can be suggested that the amount of new construction work is too small and insufficient to support the replacement of deteriorated parts. There should be more work on new construction projects to replace the deteriorated components listed in Table 1, which can be ordered by importance from the value of β that can be used in the form of the total factor deterioration index (TFDI). The proportion of external factors that have a high influence on the deterioration of components is shown in Table 1. Therefore, the

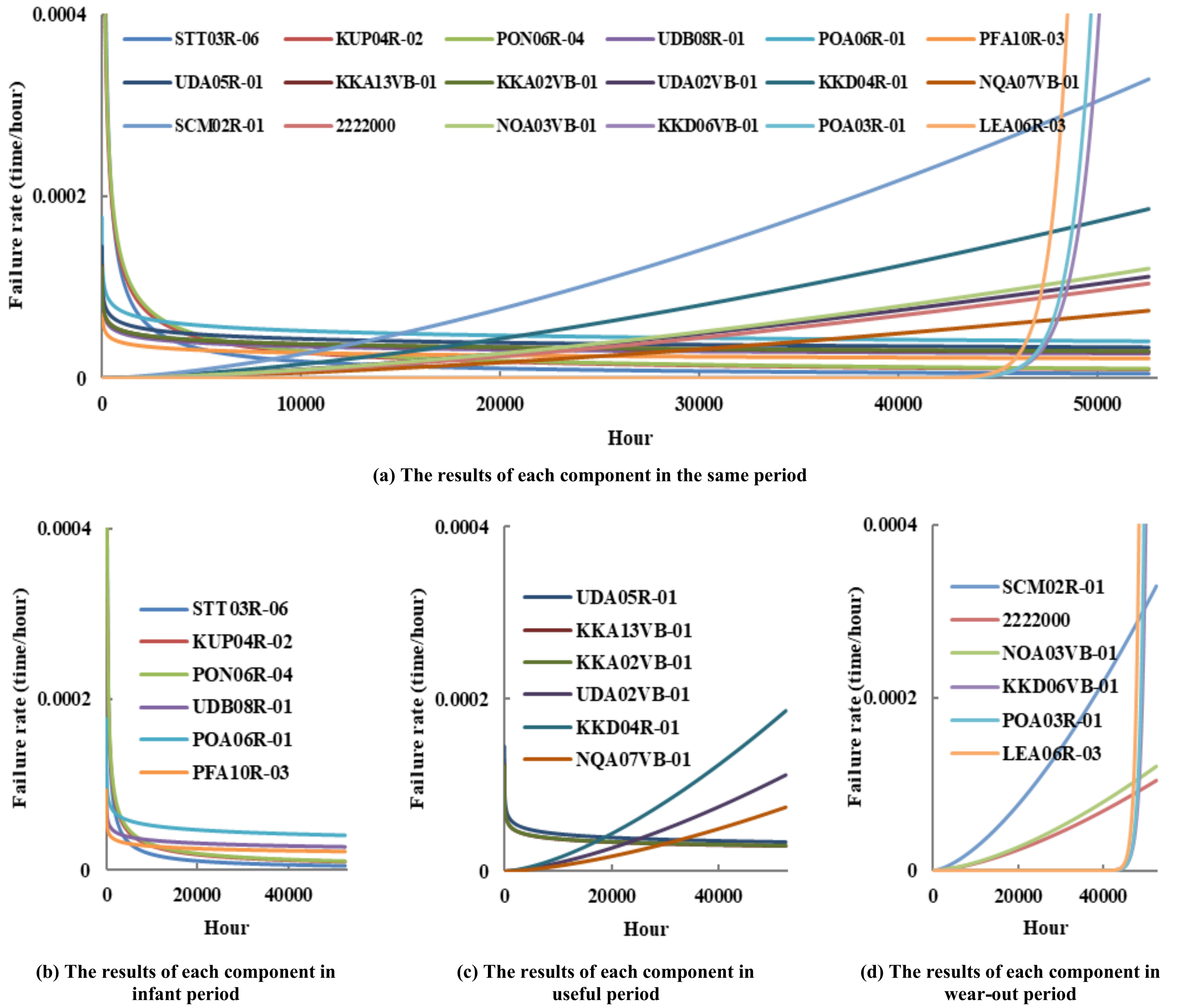


FIGURE 12. Weibull analysis results of each component based on the total factor bathtub curve.

allocation of new components must be additionally considered to be able to withstand disturbances from the external environment.

The chart of component failure rates in the first three and last three parts of each period in which the constituent shapes that can be combined as a characteristic of the bathtub curve are shown in Fig. 12. The superimposed shape of the bathtub curve is shown in the same period as in Fig. 12 (a) and can be separated into a discontinuous bathtub curve, as shown in Fig. 12 (b).

C. PROCESS 3: ESTIMATING SIMPLE BATHTUB CURVE FOR REPRESENTATION OF THE OVERALL ELECTRICAL DISTRIBUTION SYSTEM.

The simple total-factor bathtub curve estimation of this overall power distribution system is the last process of the predictive analyses presented in this paper. The failure rate is estimated from the range up to 60 years with the Weibull distribution model using the aggregated data as grouped in Section B to find the shape parameter (β) of each period. The infant period was 0.87304, the useful period was 1.4282441, and the wear-out period was 3.150178, while the scale parameter (α) was used together for every period of 31,509.48 hours, which was the value obtained from all data.

The results of the estimation of failure rates were out of range for up to 60 years for all three periods. Failure equal to 0.000029 times per hour (0.254 times per year) at approximately 17520 hours (2 years) is termed the end of the infant period (T_{Infan}), and the transition junction can be selected from the useful period to the wear-out period at the first point of the cumulative probability distribution function (CDF) equal to 1, which is approximately the point equal to 0.000109 times per hour (0.954 times per year) at approximately 271,560 hours (31 years). This is called the

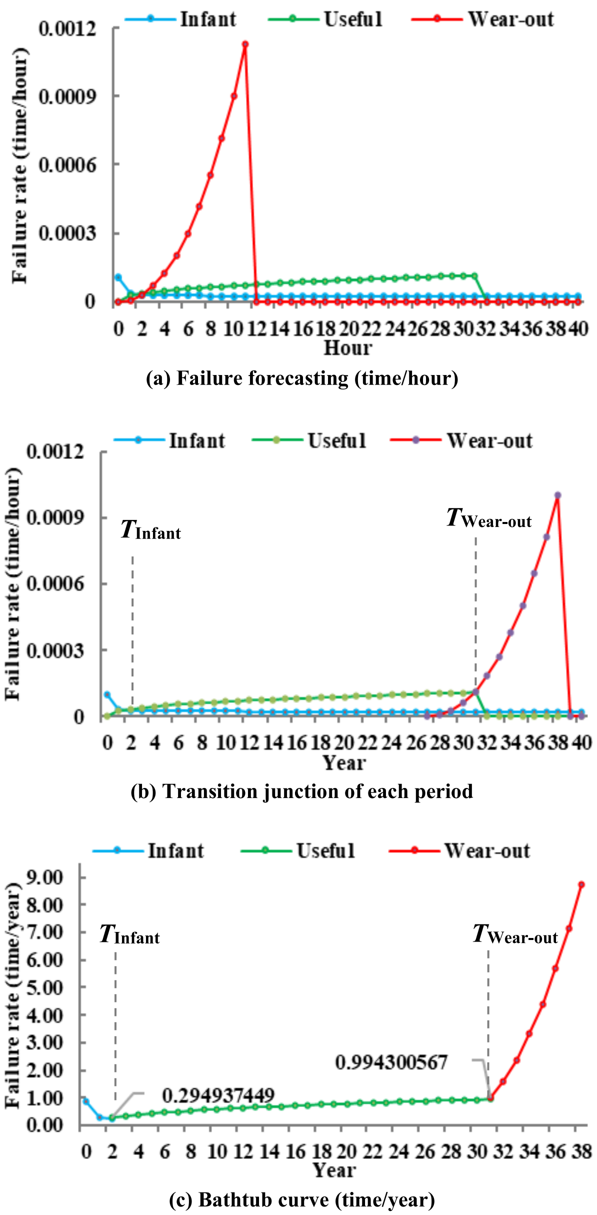


FIGURE 13. The simple total factor bathtub curve estimation of this overall power distribution system.

beginning of the wear-out period ($T_{\text{Wear-out}}$), as shown in Fig. 13 (a)–(c).

In Fig. 13 (c), the bathtub curve failure rate values are adjusted to be expressed in units of times per year and show that there is a useful period starting at 2 years and ending at 31 years. A summary of this insight can be advantageous because the appropriateness of the warranty period after construction must not be less than two years, and the period of monitoring the use and costs incurred during this period should be in agreement since commissioning. The service life of components that should be considered for the reallocation of the construction project to replace the original project should be considered starting at 31 years of service life.

Another important issue is setting operational targets to manage the reliability of the power distribution system, which should be based on the utilization of this bathtub curve. The effect of annual changes in the system and component failure rates is an important factor in planning and predicting future outcomes.

All analyses and calculations in this study were performed using Microsoft Excel. The dataset, demonstration of calculation methods and analysis results are provided in the files listed in the Appendix.

V. CONCLUSION

This study attempted to explore and create strategies and methods that apply the principles of predictive reliability analysis recommended by IEEE Std 1413.1-2002 to power distribution system outage datasets, which has not been clearly shown in previous research.

The test dataset from the Provincial Electricity Authority Region 1 in northeastern Thailand is a common feature of general electricity distribution time-to-outage datasets and can be representative of other time-to-event datasets with multimodal empirical distributions, in which the proposed strategies and processes can be used for predictive analysis even though the results are different from the test datasets.

The following is a summary of the test results of the proposed strategy.

First, it was revealed that the failure rate caused by internal factors of the system itself is in a useful period, and seasonal external factors cause 86.63% of power outages. This case study also indicates that the multimodal in the empirical distribution is caused by interference from external factors.

Second, in this study, the linear regression of ln β in the low-slope range was chosen to define the boundary of the useful period with β between 0.851 and 2.51, and a ratio that may lead to future problems was found. The number of components in the wear-out period was 3.6 times higher than in the infant period, so it may be interpreted that the construction for the replacement work cannot keep up with the deterioration. Therefore, it was proposed to use the shape parameter (β) as another reliability index in the form of the total factor deterioration index (TFDI) to prioritize risk.

Finally, the end of the infant period (T_{Infan}) and the beginning of the wear-out period ($T_{\text{Wear-out}}$) are 2 and 31 years, respectively, which can lead to component and system support planning, including burns in the test, warranty period, replacement, and reallocation. In addition, the bathtub curves can be used to set operational targets more sensibly than to make predictions based solely on past assessment results.

The proposed strategy is designed for a reliability analysis using the outage datasets in a distribution system with the multimodal or seasonal shape of an empirical distribution. If the proposed strategy is applied with the other time-to-failure datasets, the multimodal distribution of datasets should be confirmed before performing analysis.

This is a summary of the main points that this study presents.

1) The finding of multimodal patterns in the empirical distribution of power outage data is due to the high influence of external factors. The proposed method can isolate those factors and can find the true trend of the failure rate function. This leads to the decision to plan work to deal with external factors that come with the season.
2) Component clustering to accord the bathtub curve condition with shape parameters (β) for electrical distribution system components is difficult because their values vary greatly due to the influence of external factors. The log scale analysis method is the preferred choice. This study is the first to present the use of the factor deterioration index (TFDI) with shape parameters (β) as a predictive reliability index for a power distribution system.
3) The bathtub curve representing the power distribution system is estimated by forecasting the failure rate function of the three periods over the actual lifetime. It can be used to manage lifecycle risk with condition-based reliability management strategies.
4) The results of the Weibull analysis were confirmed in accordance with the histogram and the results of the regression analysis were tested with R^2, confirming the correctness of the bisection algorithm's optimality with good convergence to a single solution. A simple bathtub curve construction method is proposed for use with the condition that high tolerances of the results can be achieved.

After the development of the proposed strategy, the predictive reliability analysis may be enhanced by machine learning methods with more complicated datasets. Studying other factors, such as trees, animals, or storms, which influence the results of reliability assessment is an interesting issue and is useful for system planning in practice. Furthermore, developing an advance forecasting model for reliability analysis in long-term planning may achieve more efficient modern system operation in smart grid environments.

APPENDIX

The experimental dataset can be downloaded at https://shorturl.asia/b7tH6.

REFERENCES

[1] Y. Dechgummarn, P. Fuangfoo, and W. Kampeerawat, "Reliability assessment and improvement of electrical distribution systems by using multinomial Monte Carlo simulations and a component risk priority index," *IEEE Access*, vol. 10, pp. 111923–111935, 2022, doi: 10.1109/ACCESS.2022.3215956.
[2] A. Chowdhury and K. Don, *Power Distribution System Reliability: Practical Methods and Applications*. Hoboken, NJ, USA: Wiley, 2009.
[3] R. E. Brown, *Electric Power Distribution Reliability*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2017.
[4] M. Dong and A. B. Nassif, "Combining modified Weibull distribution models for power system reliability forecast," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1610–1619, Mar. 2019, doi: 10.1109/TPWRS.2018.2877743.
[5] M. Xie and C. D. Lai, "Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function," *Rel. Eng. Syst. Saf.*, vol. 52, no. 1, pp. 87–93, Apr. 1996, doi: 10.1016/0951-8320(95)00149-2.
[6] M. S. Alvarez-Alvarado, D. L. Donaldson, A. A. Recalde, H. H. Noriega, Z. A. Khan, W. Velasquez, and C. D. Rodríguez-Gallegos, "Power system reliability and maintenance evolution: A critical review and future perspectives," *IEEE Access*, vol. 10, pp. 51922–51950, 2022, doi: 10.1109/ACCESS.2022.3172697.
[7] Y. Fu, T. Yuan, and X. Zhu, "Optimum periodic component reallocation and system replacement maintenance," *IEEE Trans. Rel.*, vol. 68, no. 2, pp. 753–763, Jun. 2019, doi: 10.1109/TR.2018.2874187.
[8] A. Gaonkar, R. B. Patil, S. Kyeong, D. Das, and M. G. Pecht, "An assessment of validity of the bathtub model hazard rate trends in electronics," *IEEE Access*, vol. 9, pp. 10282–10290, 2021, doi: 10.1109/ACCESS.2021.3050474.
[9] G. Klutke, P. C. Kiessler, and M. A. Wortman, "A critical look at the bathtub curve," *IEEE Trans. Rel.*, vol. 52, no. 1, pp. 125–129, Mar. 2003, doi: 10.1109/TR.2002.804492.
[10] A. Ragab, M.-S. Ouali, S. Yacout, and H. Osman, "Remaining useful life prediction using prognostic methodology based on logical analysis of data and Kaplan–Meier estimation," *J. Intell. Manuf.*, vol. 27, no. 5, pp. 943–958, Oct. 2016, doi: 10.1007/s10845-014-0926-3.
[11] G. A. Bohoris, "Comparison of the cumulative-hazard and Kaplan–Meier estimators of the survivor function," *IEEE Trans. Rel.*, vol. 43, no. 2, pp. 230–232, Jun. 1994, doi: 10.1109/24.294997.
[12] J. Hu, Q. Sun, and Z.-S. Ye, "Condition-based maintenance planning for systems subject to dependent soft and hard failures," *IEEE Trans. Rel.*, vol. 70, no. 4, pp. 1468–1480, Dec. 2021, doi: 10.1109/TR.2020.2981136.
[13] O. Barnett and A. Cohen, "The histogram and boxplot for the display of lifetime data," *J. Comput. Graph. Statist.*, vol. 9, no. 4, pp. 759–778, Dec. 2000, doi: 10.1080/10618600.2000.10474912.
[14] P. Patil and D. Bagkavos, "Histogram for hazard rate estimation," *Sankhya B*, vol. 74, no. 2, pp. 286–301, Nov. 2012, doi: 10.1007/s13571-012-0036-1.
[15] N. Langrené and X. Warin, "Fast multivariate empirical cumulative distribution function with connection to kernel density estimation," *Comput. Statist. Data Anal.*, vol. 162, Oct. 2021, Art. no. 107267, doi: 10.1016/j.csda.2021.107267.
[16] R. J. Jackson and T. F. Cox, "Kernel hazard estimation for visualisation of the effect of a continuous covariate on time-to-event endpoints," *Pharmaceutical Statist.*, vol. 21, no. 3, pp. 514–524, May 2022, doi: 10.1002/pst.2183.
[17] B. Sukresno, A. Hartoko, B. Sulistyo, and Subiyanto, "Empirical cumulative distribution function (ECDF) analysis of Thunnus.Sp using Argo float sub-surface multilayer temperature data in Indian ocean south of Java," *Proc. Environ. Sci.*, vol. 23, pp. 358–367, Jan. 2015, doi: 10.1016/j.proenv.2015.01.052.
[18] Z. Li, Y. Zhao, X. Hu, N. Botta, C. Ionescu, and G. Chen, "ECOD: Unsupervised outlier detection using empirical cumulative distribution functions," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12181–12193, Mar. 2022, doi: 10.1109/TKDE.2022.3159580.
[19] T. J. Ikonen, F. Corona, and I. Harjunkoski, "Likelihood maximization of lifetime distributions with bathtub-shaped failure rate," *IEEE Trans. Rel.*, vol. 72, no. 2, pp. 759–773, Jun. 2023, doi: 10.1109/TR.2022.3190542.
[20] D. N. P. Murthy, M. Xie, and R. Jiang, *Weibull Models*. Hoboken, NJ, USA: Wiley, 2004.
[21] S. Ali, T. Zafar, I. Shah, and L. Wang, "Cumulative conforming control chart assuming discrete Weibull distribution," *IEEE Access*, vol. 8, pp. 10123–10133, 2020, doi: 10.1109/ACCESS.2020.2964602.
[22] *IEEE Guide for Selecting and Using Reliability Predictions Based on IEEE 1413*, Standard 1413.1-2002, 2003 doi: 10.1109/IEEESTD.2003.94232.
[23] M. I. Mazhar, M. Salman, and I. Howard, "Assessing the reliability of system modules used in multiple life cycles," in *Proc. 4th World Congr. Eng. Asset Manag. (WCEAM)*, Ateny, Grecja, Sep. 2009, pp. 567–573, doi: 10.1007/978-0-85729-320-6_65.
[24] W. J. Roesch, "Using a new bathtub curve to correlate quality and reliability," *Microelectron. Rel.*, vol. 52, no. 12, pp. 2864–2869, Dec. 2012, doi: 10.1016/j.microrel.2012.08.022.
[25] B.-D. Mohammed, U. Kumar, and M. D. N. Prabhakar, "Life cycle of engineered objects," in *Introduction to Maintenance Engineering*. Hoboken, NJ, USA: Wiley, 2016, pp. 107–122, doi: 10.1002/9781118926581.ch5.
[26] Q. Gong, Y. Xiong, Z. Jiang, J. Yang, and C. Chen, "Timing decision for active remanufacturing based on 3E analysis of product life cycle," *Sustainability*, vol. 14, no. 14, p. 8749, Jul. 2022, doi: 10.3390/su14148749.

[27] N. Bhusal, M. Abdelmalak, M. Kamruzzaman, and M. Benidris, "Power system resilience: Current practices, challenges, and future directions," *IEEE Access*, vol. 8, pp. 18064–18086, 2020, doi: 10.1109/ACCESS.2020.2968586.

[28] A. N. Tari, M. S. Sepasian, and M. T. Kenari, "Resilience assessment and improvement of distribution networks against extreme weather events," *Int. J. Electr. Power Energy Syst.*, vol. 125, Feb. 2021, Art. no. 106414, doi: 10.1016/j.ijepes.2020.106414.

[29] J. Graunt, "Natural and political observations mentioned in a following index, and made upon the bills of mortality," in *Mathematical Demography: Selected Papers*, D. P. Smith and N. Keyfitz, Eds. Berlin, Germany: Springer, 1977, pp. 11–20, doi: 10.1007/978-3-642-81046-6_2.

[30] C. Suchindran and K. Namboodiri, *Life Table Techniques and Their Applications*. Cambridge, MA, USA: Academic, 1987.

[31] M. Corazza, C. Perna, C. Pizzi, and M. Sibillo, Eds., *Mathematical and Statistical Methods for Actuarial Sciences and Finance*. Cham, Switzerland: Springer, 2022, doi: 10.1007/978-3-030-99638-3.

[32] H. Wainer and P. F. Velleman, "Statistical graphics: Mapping the pathways of science," *Annu. Rev. Psychol.*, vol. 52, no. 1, pp. 305–335, Feb. 2001, doi: 10.1146/annurev.psych.52.1.305.

[33] P. Schober and T. R. Vetter, "Kaplan–Meier curves, log-rank tests, and cox regression for time-to-event data," *Anesthesia Analgesia*, vol. 132, no. 4, pp. 969–970, Apr. 2021, doi: 10.1213/ane.0000000000005358.

[34] W. Weibull, "A statistical distribution function of wide applicability," *J. Appl. Mech.*, vol. 18, no. 3, pp. 293–297, Apr. 2021, doi: 10.1115/1.4010337.

[35] E. Machado de Assis, G. A. C. Lima, A. Prestes, F. Marinho, and L. A. N. Costa, "*q*-Weibull applied to Brazilian hydropower equipment," *IEEE Trans. Rel.*, vol. 68, no. 1, pp. 122–132, Mar. 2019, doi: 10.1109/TR.2018.2864550.

[36] J. C. Fothergill, "Estimating the cumulative probability of failure data points to be plotted on Weibull and other probability paper," *IEEE Trans. Electr. Insul.*, vol. 25, no. 3, pp. 489–492, Jun. 1990, doi: 10.1109/14.55721.

[37] B. Abba, H. Wang, and H. S. Bakouch, "A reliability and survival model for one and two failure modes system with applications to complete and censored datasets," *Rel. Eng. Syst. Saf.*, vol. 223, Jul. 2022, Art. no. 108460, doi: 10.1016/j.ress.2022.108460.

[38] J. Wang and H. Yin, "Failure rate prediction model of substation equipment based on Weibull distribution and time series analysis," *IEEE Access*, vol. 7, pp. 85298–85309, 2019, doi: 10.1109/ACCESS.2019.2926159.

[39] V. Zarnowitz and A. Ozyildirim, "Time series decomposition and measurement of business cycles, trends and growth cycles," *J. Monetary Econ.*, vol. 53, no. 7, pp. 1717–1739, Oct. 2006, doi: 10.1016/j.jmoneco.2005.03.015.

[40] S. Picoli, R. S. Mendes, and L. C. Malacarne, "*q*-exponential, Weibull, and *q*-Weibull distributions: An empirical analysis," *Phys. Stat. Mech. Appl.*, vol. 324, no. 3, pp. 678–688, Jun. 2003, doi: 10.1016/S0378-4371(03)00071-2.

[41] J. Yu, "Empirical characteristic function estimation and its applications," *Econ. Rev.*, vol. 23, no. 2, pp. 93–123, Dec. 2004, doi: 10.1081/etc-120039605.

[42] R. A. Raj, S. Murugesan, S. Ramanujam, and A. A. Stonier, "Empirical model application to analyze reliability and hazards in pongamia oil using breakdown voltage characteristics," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 29, no. 5, pp. 1948–1957, Oct. 2022, doi: 10.1109/TDEI.2022.3194490.

[43] A. de Souza, J. F. Júnior, M. C. Abreu, F. Aristone, W. A. Fernandes, R. S. C. Nunes, G. H. Cavazzana, C. M. D. Santos, C. J. Reis, and U. Dumka, "Variation of ozone in the Pantanal environment based on probability distributions," *Ozone, Sci. Eng.*, vol. 45, no. 2, pp. 130–146, Mar. 2023, doi: 10.1080/01919512.2022.2041392.

[44] T. Shimokawa and M. Liao, "Goodness-of-fit tests for type-I extreme-value and 2-parameter Weibull distributions," *IEEE Trans. Rel.*, vol. 48, no. 1, pp. 79–86, Mar. 1999, doi: 10.1109/24.765931.

[45] P. Manohar and C. R. Atla, "Development of predictive reliability model of solar photovoltaic system using stochastic diffusion process for distribution system," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 12, no. 1, pp. 279–289, Mar. 2022, doi: 10.1109/JETCAS.2022.3148147.

[46] C. Williams, C. McCarthy, and C. Cook, "Predicting reliability improvements," *IEEE Power Energy Mag.*, vol. 6, no. 2, pp. 53–60, Mar. 2008, doi: 10.1109/MPE.2007.915185.

[47] N. Balijepalli, S. S. Venkata, and R. D. Christie, "Predicting distribution system performance against regulatory reliability standards," *IEEE Trans. Power Del.*, vol. 19, no. 1, pp. 350–356, Jan. 2004, doi: 10.1109/TPWRD.2003.820192.

[48] C. A. McCarthy and H. Liu, "Predictive analysis ranks reliability improvements," *IEEE Comput. Appl. Power*, vol. 12, no. 4, pp. 35–40, Oct. 1999, doi: 10.1109/67.795136.

[49] I. Adaji, J. C. Su, and S. Gautam, "Heuristic approach to ring main unit placement within a self-healing distribution network," in *Proc. IEEE PES/IAS PowerAfrica*, Kigali, Rwanda, Aug. 2022, pp. 1–5, doi: 10.1109/PowerAfrica53997.2022.9905271.

[50] J. R. Aguero, J. Spare, E. Phillips, C. O'Meally, J. Wang, and R. E. Brown, "Distribution system reliability improvement using predictive models," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2009, pp. 1–7, doi: 10.1109/pes.2009.5275476.

[51] R. Billinton and C. Wu, "Predictive reliability assessment of distribution systems including extreme adverse weather," in *Proc. Can. Conf. Electr. Comput. Eng. Conf.*, Toronto, Ont., Canada, Mar. 2001, pp. 719–724, doi: 10.1109/CCECE.2001.933530.

[52] D. P. Ross, L. A. A. Freeman, and R. E. Brown, "Overcoming data problems in predictive reliability distribution modeling," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo. Developing New Perspect.*, Atlanta, GA, USA, Nov. 2001, pp. 742–748, doi: 10.1109/tdc.2001.971330.

[53] E. O. Amuta, S. T. Wara, A. F. Agbetuyi, and B. A. Sawyerr, "Weibull distribution-based analysis for reliability assessment of an isolated power micro-grid system," *Mater. Today, Proc.*, vol. 65, pp. 2215–2220, Jan. 2022, doi: 10.1016/j.matpr.2022.06.244.

[54] M. F. El-Naggar, A. S. Abdelhamid, M. A. Elshahed, and M. E. M. Bekhet, "Dynamic reliability and availability allocation of wind turbine subassemblies through importance measures," *IEEE Access*, vol. 10, pp. 99445–99459, 2022, doi: 10.1109/ACCESS.2022.3203423.

[55] W. Huang, C. Shao, M. Dong, B. Hu, W. Zhang, Y. Sun, K. Xie, and W. Li, "Modeling the aging-dependent reliability of transformers considering the individualized aging threshold and lifetime," *IEEE Trans. Power Del.*, vol. 37, no. 6, pp. 4631–4645, Dec. 2022, doi: 10.1109/TPWRD.2022.3152745.

[56] T.-W. Kim, Y. Chang, D.-W. Kim, and M.-K. Kim, "Preventive maintenance and forced outages in power plants in Korea," *Energies*, vol. 13, no. 14, p. 3571, Jul. 2020, doi: 10.3390/en13143571.

[57] M. Buhari, V. Levi, and S. K. E. Awadallah, "Modelling of ageing distribution cable for replacement planning," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3996–4004, Sep. 2016, doi: 10.1109/TPWRS.2015.2499269.

[58] J. Suntaranurak, T. Suwanasri, and C. Suwanasri, "Lifetime estimation of switching devices using Weibull distribution analysis," in *Proc. 17th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol. (ECTI-CON)*, Phuket, Thailand, Jun. 2020, pp. 413–416, doi: 10.1109/ECTI-CON49241.2020.9158112.

[59] F. Yi, C. Jia, W. Zhang, X. Yang, and P. Zhang, "Simulation of HVDC transmission system failure rate bathtub curve based on Weibull distribution," in *Proc. IEEE 2nd Int. Future Energy Electron. Conf. (IFEEC)*, Nov. 2015, pp. 1–5, doi: 10.1109/IFEEC.2015.7361599.

[60] Z. Tang, W. Zhou, J. Yu, and C. Zhou, "Comparison of Weibull distribution and crow-AMSAA model used in cable failure analysis," in *Proc. Spring Congr. Eng. Technol.*, May 2012, pp. 1–4, doi: 10.1109/SCET.2012.6341895.

[61] Y. Liu, J. Xv, H. Yuan, J. Lv, and Z. Ma, "Health assessment and prediction of overhead line based on health index," *IEEE Trans. Ind. Electron.*, vol. 66, no. 7, pp. 5546–5557, Jul. 2019, doi: 10.1109/TIE.2018.2868028.

[62] M. Dong, W. Li, and A. B. Nassif, "Long-term health index prediction for power asset classes based on sequence learning," *IEEE Trans. Power Del.*, vol. 37, no. 1, pp. 197–207, Feb. 2022, doi: 10.1109/TPWRD.2021.3055622.

[63] K. Singh, T. S. Valley, S. Tang, B. Y. Li, F. Kamran, M. W. Sjoding, J. Wiens, E. Otles, J. P. Donnelly, M. Y. Wei, J. P. McBride, J. Cao, C. Penoza, J. Z. Ayanian, and B. K. Nallamothu, "Evaluating a widely implemented proprietary deterioration index model among hospitalized patients with COVID-19," *Ann. Amer. Thoracic Soc.*, vol. 18, no. 7, pp. 1129–1137, Jul. 2021, doi: 10.1513/annalsats.202006-698oc.

[64] R. Wu, A. Smith, T. Brown, J. P. Hunt, P. Greiffenstein, S. Taghavi, D. Tatum, O. Jackson-Weaver, and J. Duchesne, "Deterioration index in critically injured patients: A feasibility analysis," *J. Surgical Res.*, vol. 281, pp. 45–51, Jan. 2023.

[65] C.-W. Wu, M.-H. Shu, and T.-C. Wang, ''An adaptive lot-traceability sampling plan for Weibull distributed lifetime with warranty return rate consideration and a smart information system,'' *Ann. Oper. Res.*, vol. 331, no. 1, pp. 1–15, Jun. 2023, doi: 10.1007/s10479-023-05438-8.

[66] D. Olteanu and L. Freeman, ''The evaluation of median-rank regression and maximum likelihood estimation techniques for a two-parameter Weibull distribution,'' *Qual. Eng.*, vol. 22, no. 4, pp. 256–272, Aug. 2010, doi: 10.1080/08982112.2010.505219.

[67] U. Genschel and W. Q. Meeker, ''A comparison of maximum likelihood and median-rank regression for Weibull estimation,'' *Qual. Eng.*, vol. 22, no. 4, pp. 236–255, Aug. 2010, doi: 10.1080/08982112.2010.503447.

[68] Q. Zhou, X. Huang, B. Wei, and L. Ye, ''Impulse life evaluation method of MOV based on Weibull distribution,'' *IEEE Access*, vol. 9, pp. 34818–34828, 2021, doi: 10.1109/ACCESS.2021.3062454.

[69] A. C. Cohen, ''Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples,'' *Technometrics*, vol. 7, no. 4, pp. 579–588, Nov. 1965, doi: 10.1080/00401706.1965.10490300.

[70] C.-D. Lai, *Generalized Weibull Distributions*. Berlin, Germany: Springer, 2014, doi: 10.1007/978-3-642-39106-4_2.

[71] H. Rinne, *The Weibull Distribution: A Handbook*. New York, NY, USA: Chapman & Hall, 2008, doi: 10.1201/9781420087444.

[72] E. Ostertagová, ''Modelling using polynomial regression,'' *Proc. Eng.*, vol. 48, pp. 500–506, Jan. 2012, doi: 10.1016/j.proeng.2012.09.545.

[73] P. Velleman, *Practical Analysis in One Variable*. Berlin, Germany: Springer, 2002. Accessed: Oct. 27, 2023. [Online]. Available: https://maa.org/press/maa-reviews/practical-analysis-in-one-variable

[74] A. N. Elghandour, A. M. Salah, Y. A. Elmasry, and A. A. Karawia, ''An image encryption algorithm based on bisection method and one-dimensional piecewise chaotic map,'' *IEEE Access*, vol. 9, pp. 43411–43421, 2021, doi: 10.1109/ACCESS.2021.3065810.

[75] H. Xing, Y. Mou, M. Fu, and Z. Lin, ''Distributed bisection method for economic power dispatch in smart grid,'' *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3024–3035, Nov. 2015, doi: 10.1109/TPWRS.2014.2376935.

[76] E. Barnett and C. Gosselin, ''A bisection algorithm for time-optimal trajectory planning along fully specified paths,'' *IEEE Trans. Robot.*, vol. 37, no. 1, pp. 131–145, Feb. 2021, doi: 10.1109/TRO.2020.3010632.

YUTTANA DECHGUMMARN (Member, IEEE) received the B.Eng. degree in electrical engineering from Khon Kaen University, Khon Kaen, Thailand, in 2004, and the M.Eng. degree in electrical engineering from Kasetsart University, Bangkok, Thailand, in 2012. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Khon Kaen University. Since 2008, he has been an Electrical Engineer with the Provincial Electricity Authority (PEA), Thailand. His main research interests include power system operation and planning, power system reliability, power system optimization, data science, and machine learning.

PRADIT FUANGFOO received the B.Eng. degree (Hons.) in electrical engineering from Kasetsart University, Bangkok, Thailand, in 1993, the M.Eng. degree in electrical engineering from Chulalongkorn University, Bangkok, in 1996, and the Ph.D. degree from The University of Texas at Arlington, Arlington, TX, USA, in 2006. He is currently an Assistant Governor with the Provincial Electricity Authority (PEA), Thailand. He has authored or coauthored many journal articles and conference papers focused on power engineering. His main research interests include power system analysis, power system reliability, power distribution planning, power system protection, power quality, and smart grids.

WARAYUT KAMPEERAWAT received the B.Eng. (Hons.) and M.Eng. degrees in electrical engineering from Khon Kaen University, Thailand, in 2005 and 2007, respectively, and the Ph.D. degree in electrical engineering and information systems from The University of Tokyo, Japan, in 2019. Since 2019, he has been an Assistant Professor with the Department of Electrical Engineering, Khon Kaen University. His research interests include power system operation and planning, power system reliability, electric vehicles, energy management systems, and renewable energy resources.