GU J Sci, Part C,9(1): 108-121 (2021)

Gazi Üniversitesi

Fen Bilimleri Dergisi

PART C: TASARIM VE TEKNOLOJİ

http://dergipark.gov.tr/gujsc

Selecting and Analyzing Appropriate Probability Distributions for Reliability of Electrical Transmission Lines

Yunus KARA[1,*], Mehmet Rahmi CANAL[2], İbrahim SEFA[3], Fatih Emre BORAN[4]

[1] *Nevşehir Hacı Bektaş Veli Üniversitesi, Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, 50300, NEVŞEHİR*

[2] *İnönü Üniversitesi, Mühendislik Fakültesi, Biyomedikal Mühendisliği Bölümü, 44280, MALATYA*

[3] *Gazi Üniversitesi, Teknoloji Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 06560, ANKARA*

[4] *Gazi Üniversitesi, Teknoloji Fakültesi, Enerji Sistemleri Mühendisliği Bölümü, 06560, ANKARA*

Article Info

Research article
Received: 30/01/2021
Revision: 22/02/2021
Accepted: 22/02/2021

Makale Bilgisi

Araştırma makalesi
Başvuru: 30/01/2021
Düzeltme: 22/02/2021
Kabul: 22/02/2021

Keywords

Goodness of fit
Anderson-Darling Test
Log-normal Distribution
Reliability Analysis
MTBF
RCM

Anahtar Kelimeler

Uyum iyiliği
Anderson-Darling Test
Log-normal Dağılım
Güvenirlik Analizi
MTBF
RCM

Abstract

Reliability is the ability of a system and/or system components to function within a specified time under specified conditions. One of the most important criteria of electricity transmission systems is to be able to keep the energy on the system continuously within the limits and to be interrupted for the least possible time. The reliability study offers important advantages such as determining the appropriate operating range of the system and making the necessary intervention. In this study, it is aimed to show that the transmission lines can be interrupted for the least amount of time within the specified limits when the maintenance and the operation activities of electricity transmission lines are carried out with Reliability Centered Maintenance (RCM) in the foreground. In this direction, reliability analysis was conducted using transmission line fault data. Also, the Anderson-Darling goodness of fit test was performed to determine from which statistical distribution the transmission line fault data came from. The reliability of the transmission line is evaluated with the Log-normal distribution, which is determined by the goodness of fit. As a result of the reliability assessment, the reliability of the transmission line was found to be low. The advantages that transmission lines can provide when maintenance and operating activities are conducted based on reliability are presented.

Elektrik İletim Hatlarının Güvenirliği için Uygun Olasılık Dağılım Seçimi ve Analizi

Öz

Güvenirlik, bir sistemin veya sistem bileşenlerinin belirlenen süre içerisinde, belirtilen koşullarda işlevini yerine getirme yeteneğidir. Elektrik iletim sistemlerinin en önemli kriterlerinden birisi sistem üzerindeki enerjiyi limitler dâhilinde sürekli tutabilmek ve mümkün olan en az süre kesintiye gitmektir. Güvenirlik çalışması, sistemin uygun çalışma aralığının tespit edilip gerekli müdahalenin yapılması gibi önemli avantajlar sunmaktadır. Bu çalışmada, elektrik iletim hatlarının bakım ve işletim faaliyetlerinin Güvenirlik Merkezli Bakım (Reliability Centered Maintance -RCM) ön planda tutularak yapıldığında, belirtilen sınırlar dahilinde iletim hatlarının en az sürede kesintiye uğrayabileceğinin gösterilmesi amaçlanmıştır. Bu doğrultuda, iletim hattı arıza verileri kullanılarak güvenilirlik analizi yapılmıştır. Ayrıca, iletim hattı arıza verilerinin hangi istatistiksel dağılımdan geldiğini belirlemek için Anderson-Darling uyum iyiliği testi yapılmıştır. Uyum iyiliği ile belirlenmiş olan Log-normal dağılım ile iletim hattının güvenirlik değerlendirilmesi yapılmıştır. Güvenirlik değerlendirilmesi sonucunda iletim hattının güvenirliğinin düşük olduğu görülmüştür. İletim hatlarının bakım ve işletim faaliyetlerinin güvenirlik merkezli yapıldığında sağlayabileceği faydalar sunulmuştur.

*Corresponding author, e-mail: yunuskara@nevsehir.edu.tr
DOI: 10.29109/gujsc.868923

1. INTRODUCTION

Long-lasting operation of the system without failure provides a safer environment in terms of efficiency and cost, as well as the employee health and safety. Electricity transmission lines are asked to remain continuously energized. It is undesirable to cut off the energy on the line. It is essential to keep the lines energized and to complete the interruption as soon as possible. In order to meet these conditions, maintenance work is increased and provided in the current situation. But, increasing maintenance work will increase the cost and number of personnel. At the same time, it will prolong the staff's exposure to risk in the work area. In the literature, the process called RAMS (Reliability, Availability, Maintainability and Safety) has gained importance in order to minimize the afore-mentioned problems. RAMS, which is the basic principle of preventive maintenance, is effective in reducing the maintenance time, the number of failures and risk level, and extending the operating time of the system. Due to the preventive advantage of RAMS activities, the health and the safety of employees coincide with the proactive prevention approach. The most important step of RAMS studies is Reliability Centered Maintenance (RCM). RCM is a technique used to develop the cost-effective maintenance plans [1]. Reliability-centered maintenance provides the opportunity to determine the appropriate operating range of the system and to make the necessary intervention.

IEEE (The Institute of Electrical and Electronics Engineers) defines reliability as "the ability of a system or component to perform the specified functions for a certain period of time" [2]. Reliability is an interdisciplinary work subject used in almost all engineering fields that try to bring safety, time, cost and quality to the appropriate values as much as possible. Typical criteria for reliability are error rate/frequency, average downtime, and average time between failures. Mathematically, reliability can be expressed as the probability that the time to failure of the component or system is greater than or equal to a certain time (t), denoted by $R(t) = P(T \geq t)$, where R (t) is reliability, P is probability, T is any time, and t is the determined time [3].

Reliability depends on the quality of the data used in the studies. System reliability models are applied to the probability theory with the aim of adapting error data to statistical distribution models and this process is created by data analysis. The purpose of data analysis is to obtain reliability and hazard functions, and two methods are used for this. The first one of these methods is the nonparametric method. In this context, a nonparametric method is a method that takes into account experience and observations based on the field experience rather than applying a mathematical method. On the other hand, the second one is parametric method. It is usually the preferred method and is characterized (simulated) by an appropriate distribution of the fault data [3]. There are many distributions used in parametric evaluation of the reliability. Among these distributions, the most common ones are normal, log-normal, Weibull, gamma, and exponential distributions [4, 5].

According to IEEE P1366 standard electrical distribution reliability indices, the mostly used indices are system average interruption frequency index (SAIFI) and system average interruption duration index (SAIDI) [6]. SAIFI is commonly used as a reliability indicator and is calculated using the error rate (λ) criterion. Error rate (failure rate) criterion plays an important role in reliability and survival studies. With the failure rate model, a mathematical model of the distribution can be created for a specific life span [7]. Mean Time to Between Failure (MTBF) is the inverse of the error rate, and the MTBF criterion is one of the most widely used criteria in reliability calculations [8]. Reliability assessments aim to calculate the frequency and the expected duration of the interruption. Reliability calculation with the use of distribution functions can provide many advantages, such as job planning, scheduling, warehouse inventory, the cost estimation of the interruption frequency and the expected duration of the outage.

In order to determine the reliability and the lifespan of the systems, past error data of the systems are required. The reliability/security study depends on the quality of the data and is basically conducted with two approaches. The first method of these approaches, the empirical approach (non-parametric method), derives the reliability and hazard functions directly from fault data [3]. It is usually carried out with data from experienced people in the field. As an example, the reliability ratio of the N component system can be calculated by proportioning the number of non-deteriorating equipment to the total number of equipment at the end of the t period. The other method is the parametric approach. As the name suggests, the parametric approach is characterized by parameters, in which fault data are defined with a probability distribution and distribution parameters are estimated [3]. Goodness of fit tests need to be applied to examine whether the data fit a particular distribution [9, 10]. Reliability analysis and life expectancy estimation were applied to different areas in the literature, using parametric probability distributions. Gorjian Jolfaei et al. evaluated the reliability of the power generation system in order to minimize the operating and maintenance costs in waste-water treatment plants with the two-parameter Weibull model [11]. Kumar and Krishnan estimated the reliability evaluation and the expected first failure time with two-parameter Weibull Distribution using the MTBF criteria of twenty-four diesel compressors of the same brand and model [12]. Li et al. determined the lower and upper limits of reliability of electrical elements with three-parameter generalized inverse Weibull [13]. Volkanovski et al. developed a reliability assessment aimed at identifying weak points of the system in order not to interrupt electrical power systems [14]. Using the real data of the last five years used in the reliability assessment of the electrical power system, Tur assessed the reliability with the indices specified in the IEEE P1366 standard, and determined the most suitable feeder by commissioning the switching points determined to make the power system more reliable with four different configurations [15]. Deng et al. made their reliability assessment with Weibull distribution in order to improve the safe working coefficient of mobile equipment, reduce maintenance costs, and extend equipment life [16]. Roshan et al. conducted a study to find the most suitable model among the four distributions used in reliability evaluations, which are log-normal, Weibull, gamma, and exponential distribution [17]. Gupta et al. used the log-normal distribution to evaluate the reliability of variables showing positively skewed [18]. Raqab et al. examined the distinctions of Weibull, log-normal, and log-logistic distributions used for the lifespan estimation of the positively skewed data [19]. Dey and Kundu analyzed the discrimination of the positively skewed data, which is important in reliability assessment [20].

In this study, unlike other studies, in order to evaluate the reliability of the system as a whole, reliability analysis of Konya4-named feeder located in Yeşilhisar Transformer Center at the Kayseri 11th Regional Directorate under General Directorate of Turkish Electricity Transmission Corporation (TETC) was performed. First of all, fault record data emerging between 2014-2019 were collected. The goodness of fit test was used to determine which parametric distribution was appropriate for the collected fault record data. Then, the reliability value of Konya4 feeder was calculated by using the probability distribution. According to the calculated reliability value, it is concluded that when the maintenance and the operation activities of the electricity transmission lines are based on the reliability-centered maintenance studies, intervention and operation continuity within the specified limits can be achieved. The parametric probability distributions used in reliability evaluation are explained in Section 2. The goodness of fit test and reliability analysis application example for determining the proper distributions used in the reliability analysis of Konya4 feeder are included in Section 3. Results and evaluations are given in Section 4.

2. DISTRIBUTIONS USED FOR RELIABILITY ASSESSMENT

Different probability distributions are used for reliability assessment. Distributions are divided into two as discrete probability and continuous probability according to whether the variable is discrete and continuous. Fiondella and Xing state that the mean time between errors is evaluated by continuous probability distributions [21]. There are many continuous distributions in the literature for reliability assessment. Verma et al. classify the distributions of statistics that are widely used in reliability according to their application areas. This classification is shown in Table 1 [3].

Table 1 *Statistical distributions used for reliability assessment [3]*

Distribution Type	***Application areas in reliability assessments***
Poisson Distribution	*Determining error occurrence rates per hour or per element*
Binomial Distribution	*In the reliability assessment of KooN systems*
Exponential Distribution	*Determining the lifetime distribution of non-repairable complex systems and the lifetime of multi-equipment systems*

Weibull Distribution	*Predicting downtime of equipment that is often subject to wear and fatigue*
Log-normal Distribution	*Modeling the repair time and modeling the lifetimes of metals and transistors*
Normal Distribution	*Analysing lifetime distribution and load resistance of components under load.*
Gama Distribution	*Modeling system failure timing and modeling time between maintenance*

2.1. Exponential Distribution

The exponential distribution is the most widely used distribution in reliability and risk assessment. Exponential distribution is used to model the useful life of the system [3]. Pham defined the probability density function (PDF) f(t) as a mathematical function that defines the probability of each element of a discrete set or the result range or possible values of a continuous variable. The probability density function for the exponential distribution is defined in Equation 1 [8]:

$$f(t) = \lambda e^{\lambda t} \qquad 0 < t < \infty, \qquad f(t) = 0 \qquad t < 0 \qquad 1$$

I λ is the hazard (failure) rate, MTBF = $1/\lambda$.

Pham defined the cumulative density function (CDF) F(t) as a function that gives the probability of a random T variable to take values less than or equal to some t values. The cumulative density function for the exponential distribution is defined in Equation 2 [8].

$$F(t) = \int_0^t f(t)dt = 1 - e^{-\lambda t} \qquad 2$$

The reliability function R(t) is a time-varying function and it is complementary to the cumulative function. In cases where the failure time is modeled, the cumulative density function represents the failure probability and the reliability function survival probability. Pham defined the reliability function for the exponential distribution in Equation 3 [8].

$$R(t) = 1 - F(t) = e^{-\lambda t} \qquad 3$$

The hazard function *h(t)* is the ratio of the probability density function to the reliability. Pham defined the hazard function for the exponential distribution in Equation 4 [8].

$$h(t) = \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda \qquad 4$$

The mean value *E(t)* is the mean value of a distribution. Verma et al. defined the mean value for the exponential distribution in Equation 5 [8].

$$E(t) = \int_0^\infty t f(t) = \frac{1}{\lambda} \qquad 5$$

Var (t) Variance is a statistical expression that determines the average distance of a variable set from the mean value in that set. Verma et al. defined the variance for the exponential distribution in Equation 6 [3].

$$Var(t) = E(T^2) - (mean)^2 = \left(\int_0^\infty t^2 f(t)dt\right) - \left(\frac{1}{\lambda}\right)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \qquad 6$$

2.2. Normal Distribution

Normal distribution is the most important and widely used distribution in statistics and probability fields. It is known as the Gaussian distribution, and is used to represent attrition information by which fatigue and aging will be modeled [3].

Pham defined the probability density function (PDF) *f(t)* for the normal distribution in Equation 7 [8].

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}, \qquad -\infty \leq t \leq +\infty \tag{7}$$

In the equation, μ is the population mean, and σ is the population standard deviation, which is the square root of the variance.

Pham defined the cumulative density function (CDF) *F(t)* for the normal distribution in Equation 8 [8].

$$F(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \tag{8}$$

The reliability function *R(t)* for Pham normal distribution is defined in Equation 9 [8].

$$R(t) = \int_{t}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \tag{9}$$

Here, Z transform is done for integral analysis. The z transformation is defined in Equation10.

$$z = \frac{t-\mu}{\sigma} \tag{10}$$

Re-expression of the cumulative density function by substituting the Z transform is defined in Equation 11.

$$\emptyset(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{11}$$

Pham defined the hazard function h (t) for the normal distribution in Equation12 [8].

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1-\emptyset(z)} \tag{12}$$

2.3. Log-normal Distribution

Log-normal distribution is the normal distribution with the natural logarithm of the variable that takes a continuous positive value. Log-normal distribution is used to model failure cycles of metals, lifetimes and repair times of transistors, and bearings [3].

Pham defined the probability density function (PDF) *f(t)* for the log-normal distribution in Equation13 [8].

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma t} e^{-\frac{1}{2}\left(\frac{lnt-\mu}{\sigma}\right)^2}, \qquad 0 < t \tag{13}$$

Pham defined the cumulative density function (CDF) *F(t)* for the log-normal distribution in Equation 14 [8].

$$F(t) = \Phi\left[\frac{lnt-\mu}{\sigma}\right] \tag{14}$$

In this equation Φ is the standard normal distribution cumulative density function.

The reliability function *R(t)* for Pham log-normal distribution is defined in Equation15 [8].

$$R(t) = 1 - \Phi\left[\frac{lnt-\mu}{\sigma}\right] \tag{15}$$

Pham defined the hazard function *h(t)* for the normal distribution in Equation 16 [8].

$$h(t) = \frac{f(t)}{R(t)} = \frac{\frac{1}{\sqrt{2\pi}\sigma t} e^{-\frac{1}{2}\left(\frac{lnt-\mu}{\sigma}\right)^2}}{1-\Phi\left[\frac{lnt-\mu}{\sigma}\right]} \tag{16}$$

Verma et al. defined the mean value for the log-normal distribution in Equation 17 [3].

$$E(t) = e^{\mu+\frac{\sigma^2}{2}} \tag{17}$$

Verma et al. defined the variance for the log-normal distribution in Equation 18 [3].

$$Var(t) = e^{(2\mu+\sigma^2)}\left(e^{\sigma^2} - 1\right) \tag{18}$$

2.4. Weibull Distribution

The Weibull distribution has a wide application in reliability calculation due to its flexibility in modeling different distribution types. This distribution method can be used to model time for failure of lamps, relays, capacitors, germanium transistors, ball bearings, automobile tires, and some motors. In addition to being the most useful distribution function in reliability analysis, it is used in the classification of fault types, troubleshooting, preventive maintenance, and inspection activity programming [3].

Pham defined the probability density function (PDF) f(t) for the Weibull distribution in Equation 19 [8] as

$$f(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \;, \qquad 0 < t \tag{19}$$

where β is the shape parameter, and η is the scale parameter.

Pham defined the cumulative density function (CDF) *F(t)* for the Weibull distribution in Equation 20 [8].

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \tag{20}$$

Pham defined the reliability function *R(t)* for the Weibull distribution in Equation 21 [8].

$$R(t) = e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \tag{21}$$

Pham defined the hazard function *h(t)* for the Weibull distribution in Equation 22 [8].

$$h(t) = \frac{\beta t^{\beta-1}}{\alpha^{\beta}} \tag{22}$$

Verma et al. defined the mean value for the Weibull distribution in *E(t)* Equation 23 [3].

$$E(t) = \alpha\Gamma\left(1 + \frac{1}{\beta}\right) \tag{23}$$

In this equation, *Γ(x)* is the gamma function and it is defined in Equation 24.

$$\Gamma(x) = \int_0^{\infty} y^{x-1} e^{-y} dy \tag{24}$$

Verma et al. defined the variance Var (t) for the Weibull distribution in Equation 25 [3].

$$Var(t) = \sigma^2 = \alpha^2\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right] \tag{25}$$

3. RELIABILITY ANALYSIS EXAMPLE OF KONYA 4 FEEDER

The reliability analysis of the 380kV high voltage line, the total line length of 223km and the bidirectional current carrying capacity of the feeder, named Konya4 at the Kayseri 11th Regional Directorate Yeşilhisar Transformer Center under General Directorate of Turkish Electricity Transmission Corporation, was made. Restrictions such as land structure and general weather conditions that may cause failure in Konya4 feeder were ignored. Thus, the steps given below were followed for the reliability analysis.

Step 1. Obtaining fault record data

The mean times between failures (MTBF) data were extracted from the last six years of fault records of the Konya4 feeder, and reliability assessment was aimed to be made with these data. The data do not have a specific frequency. In addition, the data are considered as data for malfunctions that cause the system to stop and that require manual or automatic intervention. It is aimed to evaluate the reliability of Konya4 feeder by choosing the most suitable one among the most used exponential, normal, log-normal and Weibull distributions in the literature. Thus, the data of the last six years of Konya4 feeder are given in Figure 1 below.

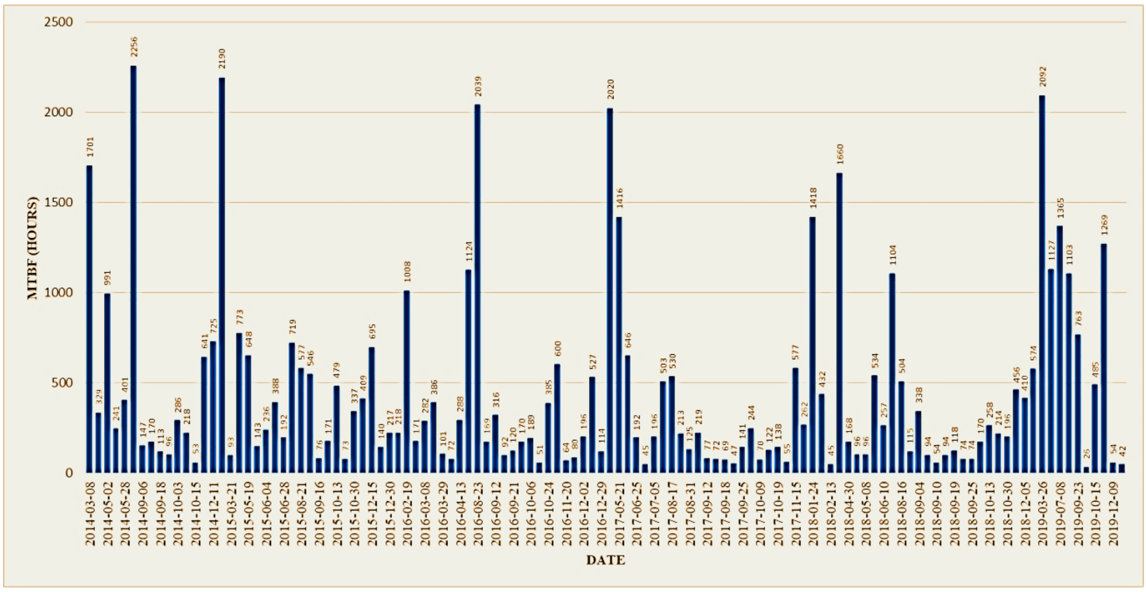

Figure 1. *Fault MTBF times of Konya4 Feeder between 2014-2019 (hours)*

Step 2. Determining the probability distribution with the Goodness of Fit Test

Goodness of fit test is a statistical test for choosing the appropriate distribution in the parametric approach. It is used to determine whether the data are suitable for a certain probability distribution. In other words, it tests whether the data comes from a specific or partially specified probability distribution [22]. The null hypothesis (H0) and the alternative hypothesis (H1) are defined to determine if random data fit a distribution.

H0: The data does not differ from the probability distribution; it is the specified distribution.

H1: There is a difference between the probability distribution of the data; it is not the distribution specified [3].

The classical approach to obtaining the null hypothesis fit test that a data has a certain probability distribution is to divide the possible values of random variables into a limited number of regions. The number of sample values entering each region is then determined and compared with the theoretical expected numbers under the specified probability distribution, and the null hypothesis is rejected when they are significantly different [22]. When performing the statistical evaluation of the data in good fit, the difference between the data and the distribution model is evaluated and this difference is desired to be small. If it is less than the value indicated at the end of the evaluation, H0 is accepted, or H1 is accepted [3, 23].

Goodness of fit can be expressed as the first step in data evaluation. There are different goodness of fit tests in the literature. Kolmogorov-Simirnov, Anderson-Darling, Cramer-von Missses, Pearson Chi-square, Shapiro-Wilk, Shapiro-Francia, Weisberg-Bingham, D'Agustino, Fillibun and Jarque-Bera can be given as examples of the most used of these tests [24]. The data used in the study are about the time between failures of the transmission line. In the goodness of fit studies, these durations are considered as hours, sorted according to chronological order and the Anderson-Darling test, which is the goodness of fit test, is classified according to periods during the application phase.

Step 2.1. Anderson-Darling Goodness of Fit Test

As a goodness of fit test, the Anderson Darling test calculates the Anderson-Darling statistics between the cumulative distribution function (CDF) and the empirical probability density function (EDF). For example, the Anderson Darling test is used to determine whether it belongs to a particular distribution [25]. Anderson-Darling goodness of fit test is widely used and has good strength properties [26]. Anderson-Darling goodness of fit test has the advantage of a more sensitive test by making use of other distributions (Normal, Lognormal, Exponential, Weibull, Logistic, etc.) in calculating the p value. Although calculating the p value for each given statistical distribution is a disadvantage, software such as Minitab is turned it into an advantage.

Let the data x_1≤ x_2≤⋯≤x_n be as follows in ascending order (Anderson-Darling equation for n data is shown in Equation 26 and Equation 27):

$$A_n^2 = n\int_{-\infty}^{+\infty}[F_n(x) - F(x)]^2/[F(x)\{1 - F(x)\}]dF(x) \qquad 26$$

Here, Fn (x) and F (x; 0) are EDF and CDF, respectively.

$$A_n^2 = -n - (1/n)\sum_{i=1}^{n}(2i - 1)\ln(u_i) + \ln(1 - u_{n-i+1}) \qquad 27$$

Here, $u_i = F(x_i); i = 1,2,\dots,n$

The calculated test statistic is then compared with a critical value according to the importance of the values. Generally, if the test statistic is less than the critical value, the null hypothesis (H0) is accepted. If not, the alternative hypothesis (H1) is accepted. Critical value depends on test severity and sample size [3, 25, 27].

The p probability value is used to interpret the goodness of fit test results. As a result of the goodness of fit test of the data set, the p value is compared with the degree of significance (α) to determine whether the data differ from the statistical distributions. The significance level (α) is a value selected in the goodness of fit test evaluation process and is compared with the p value at the end of the test. In the Anderson-Darling test, the significance level, α = 0.05, is generally used. In addition, the degree of importance (α) can be equal to values of 0.2, 0.1, 0.05, 0.02 and 0.01 [3, 25, 28]. These values are the percentage of the error made in rejecting the null hypothesis. For example, when the alpha value is selected as 0.1, it means rejecting the null hypothesis with a tolerance of 10% of the evaluation. The degree of importance (α) value is chosen and the p value is a value that can be calculated. The p value can be calculated statistically by different methods, but the comments are the same. If the p value is less than alpha, the null hypothesis is rejected; if the p value is greater than or equal to alpha, the null hypothesis is accepted [29]. In this study, the importance level was determined as α = 0.05.

Minitab 19 statistics program was used to determine which distribution best fits the MTBF data of Konya4 feeder between 2014-2019, which is given in Figure 1. Normal, exponential, log-normal and Weibull distributions were tested according to Anderson-Darling goodness of fit. Thus, the test results obtained are shown in Table 2.

Table 2. *Anderson-Darling compatibility results of distributions of Konya4 Feeder*

Distribution	***AD***	***p***
Normal	*10.744*	*<0.005*
Exponential	*2.253*	*0.005*
Weibull	*2.006*	*<0.01*
Log-normal	*0.629*	*0.099*

** In the Anderson-Darling test, α was chosen as 0.05.*

As can be seen in Table 2, the p value is less than 0.005 in the normal distribution, it is equal to 0.005 in the exponential distribution, it is less than 0.01 in the Weibull distribution, and it is equal to 0.099 in the log-normal distribution. According to these results, since p is 0.099> α = 0.05, the null hypothesis (H0: Konya4 feeder data come from the Log-normal probability distribution family) is accepted, and the H1 hypothesis is rejected. (H1: There is a difference between the data of Konya4 feeder with normal, exponential, and Weibull probability distributions). Thus, the compatibility test curves made with Anderson-Darling are shown in Figure 2.

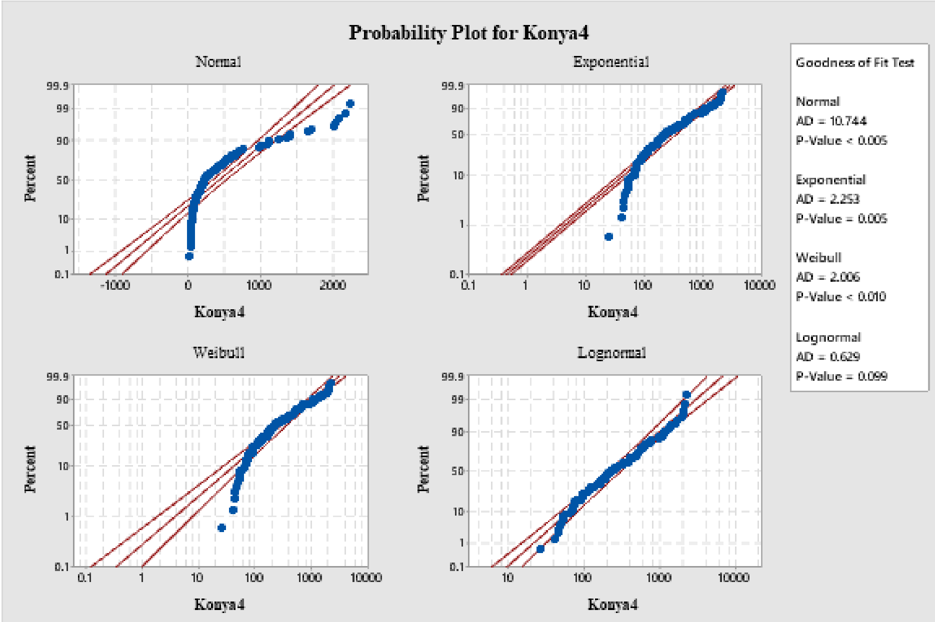

Figure 2. *Anderson-Darling compatibility curves of Konya4 Feeder distributions*

Figure 2 shows that, according to Anderson-Darling compatibility curves, the most harmonious distribution among the values remaining in the 95% confidence interval (CI) is the log-normal distribution.

Normality test and the normality test of Konya4 feeder data were performed in order to reveal whether the data conformed to the normal distribution. Skewness and kurtosis values in the normality test should be in the range of ± 1 according to Hair et al. According to Tabachnic and Fidel, the skewness and kurtosis values in the normality test results should be in the range of ± 1.5 [30, 31]. When the normality test is applied, according to Anderson-Darling, to the MTBF data of Konya4 feeder, it is seen that the skewness is 1.98, that the kurtosis is 3.53, and that the curve is positively skewed according to Figure 3.

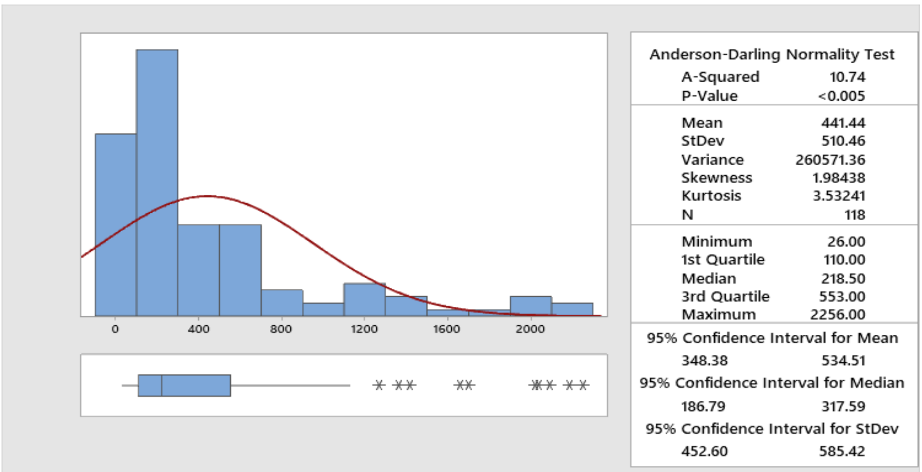

Figure 3. *Normality curve of Konya4 feeder*

Generally, the data that do not show normality can be normalized by transforming. According to Tabachnic and Fidel, data showing positive skewness is normalized by taking its natural logarithm (ln) [31]. When the natural logarithm of the Konya4 feeder is taken and subjected to the normality test again, it is seen that the skewness: 0.2 and the kurtosis: -0.76 values and the distribution become close to normal as can be seen in Figure 4. Thus, it is verified to perform reliability assessment with log-normal distribution showing the best performance in goodness of fit. It coincides with the studies in the literature that use log-normal distribution in reliability evaluation for the data showing positively skewed.

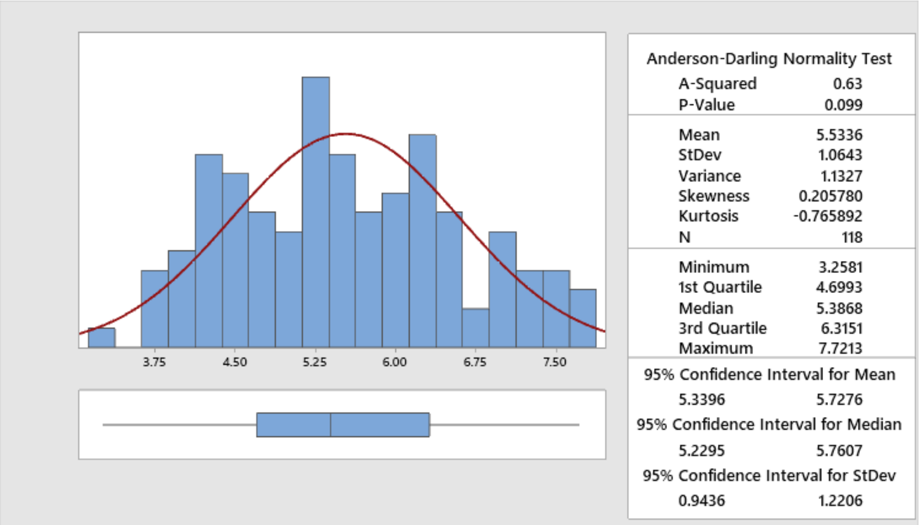

Figure 4. *Normality test curve with natural logarithm of Konya4*

Step 3. Calculation of the reliability value of Konya4 feeder

In the previous section, it was determined that the most appropriate distribution to be used for the reliability studies of the MTBF data of Konya4 feeder is the log-normal distribution. These equations used in the reliability assessment of the log-normal distribution in Section 2.3 are defined in 14, 15, 16, 17, and 18. Log-normal distribution is a parametric distribution, and firstly, location parameter (μ) and scale parameter (σ) are calculated, which are the parameters of the log-normal distribution. Microsoft Excel program is used in calculations.

The location parameter, μ, is defined in Equation 28.

$$\mu = \frac{\sum_{i=1}^{n} lnt_i}{n} \tag{28}$$

Scale parameter σ is defined in Equation 29.

$$\sigma = \sqrt{\frac{n\sum_{i=1}^{n}(lnt_i)^2 - \left(\sum_{i=1}^{n} lnt_i\right)^2}{n(n-1)}} \tag{29}$$

In the calculations made using Equation 28 and Equation 29, location parameter and scale parameter were found as follows, respectively:

μ= 5.533591979 and σ= 1.064268901

The probability density function according to the log-normal distribution (PDF) is found as follows when evaluated according to Equation 30:

$$f(t) = \frac{1}{\sigma t\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{lnt-\mu}{\sigma}\right)^2} = 0.002456939 \tag{30}$$

When the cumulative density function (CDF) according to log-normal distribution is evaluated according to Equation 31, it was found as follows:

$$F(t) = \Phi\left[\frac{lnt-\mu}{\sigma}\right] = 0.236737331 \tag{31}$$

The reliability function *R(t)* according to the log-normal distribution was found as follows when evaluated according to Equation 32:

$$R(t) = 1 - \Phi\left[\frac{lnt-\mu}{\sigma}\right] = 1 - 0.236737331 = 0.763262669 \tag{32}$$

When the hazard function *h(t)* according to the log-normal distribution is evaluated according to Equation 33, it is found as follows:

$$h(t) = \frac{f(t)}{R(t)} = \frac{0.002456939}{0.763262669} = 0.00322 \quad (33)$$

The average value *E(t)* according to the log-normal distribution was found as follows when evaluated according to Equation 34:

$$E(t) = e^{\mu+\frac{\sigma^2}{2}} = e^{6.094} = 443.191\ hours \quad (34)$$

When the reliability evaluation was made using the Minitab 19 program, the curves in Figure 5 were obtained.

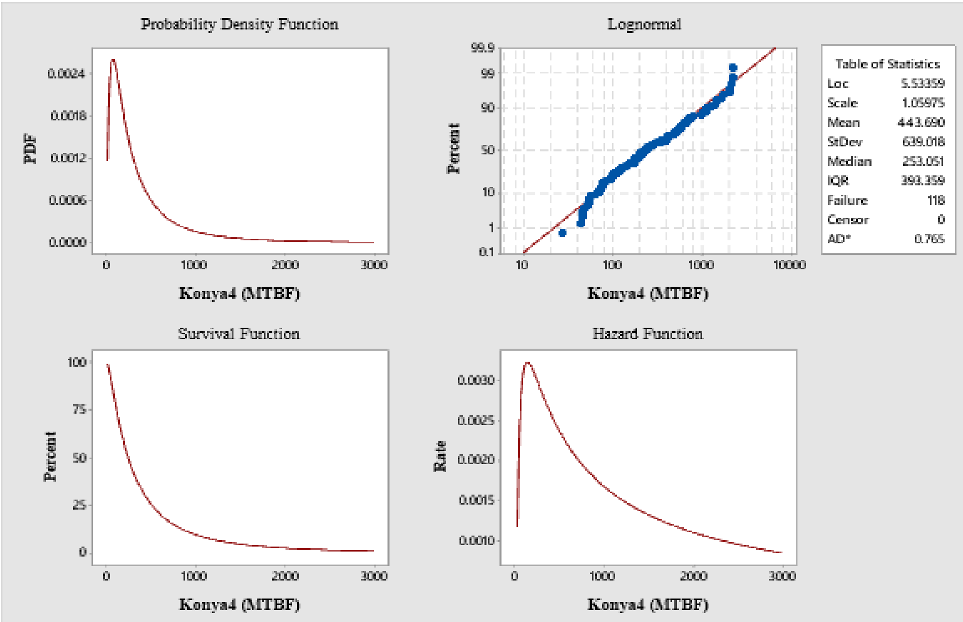

Figure 5. *Konya4 feeder reliability evaluation curves*

As the curves in Figure 5 demonstrate, the location parameter, scale parameter, and average values are equal when the values in the statistical table, which is the program output, and the values caused by the rounding in the calculations are ignored. While the probability density function curve and the hazard function curve should ideally be in the form of a bell curve, it was found as a positively skewed curve in the study. Most of the fault times on a positively skewed curve are less than the average value. The results of the reliability assessment of the Konya4 feeeder are low. Low reliability may cause the transmission line to be interrupted at unexpected times. With the RCM, the reliability value can be raised and the intervention time to the system can be adjusted. If the maintenance work of the Konya4 feeder system had been based on reliability, it could have been interpreted that many failures would occur before the planned intervention and the success of the reliability centered maintenance work would be low when planned intervention to the system at an average value was required.

4. CONCLUSION AND EVALUATION

In this study, the reliability model has been created with the MTBF values obtained by extracting the fault values of the 380 kV electricity transmission line feeder named Konya4 at the Kayseri 11th Regional Directorate Yeşilhisar Transformer Center under General Directorate of Turkish Electricity Transmission Corporation. While creating the reliability model, evaluation is made with parametric methods. The goodness of fit test, which is the most appropriate distribution of probability, is performed using the MINITAB 19 statistical program. According to the result of the goodness of fit test, the log-normal distribution is determined as the most appropriate distribution. Reliability is evaluated using log-normal distribution. As a result of the evaluation, the reliability is found as R(t) = 0.763 and the mean value E(t) = 443.191 hours. The following assessments have been proposed based on these results:

1) The closer the reliability value is to 1, the longer the system may be able to operate safely. Reliability value of Konya4 feeder is far from 1. The most important reason for the low reliability value is that the maintenance and operation activities used are not reliability-centered. When maintenance and operation activities are based on reliability, activity planning can be created more easily. In reliability-centered maintenance activities, system intervention is carried out on average time. When the fault data set is analyzed, it is seen that the majority of the time between failures is smaller than the average value and the number and duration of intervention to the system are high. Since intervention to the system may be planned in reliability-centered maintenance activities, preventive maintenance work can be used more effectively, thus saving on costs, personnel, time, and materials to be kept in the warehouse inventory. In addition, many risks in terms of the safety of employees can be eliminated as the working time under risk can be shortened.

2) When the maintenance and operation activities are based on reliability, the interruption time of the transmission lines can be kept as low as possible. Turkish Electricity Transmission Corporation keeps a regular and careful plan of failure records that occur in the system. General maintenance plans are scheduled by taking into account the workload. There are other factors such as errors caused by meteorological conditions and errors caused by renewed equipment in reliability-centered maintenance activities. Turkish Electricity Transmission Corporation records these entries for different purposes. These recorded data can be easily evaluated within reliability-centered maintenance studies. If the information in the current method can be collected under the reliability-centered maintenance studies without requiring major changes for reliability-centered maintenance studies, the awareness of the advantages of reliability-centered maintenance can be achieved and used correctly.

REFERENCES

[1] Yssaad, B., M. Khiat, and A. Chaker, Reliability centered maintenance optimization for power distribution syst9ems. International Journal of Electrical Power & Energy Systems, 2014. 55: p. 108-115.

[2] Geraci, A., IEEE standard computer dictionary: Compilation of IEEE standard computer glossaries. 1991: IEEE Press.

[3] Verma, A.K., S. Ajit, and D.R. Karanki, Reliability of Complex Systems, in Reliability and Safety Engineering. 2016, Springer. p. 123-159.

[4] Nayak, P., A.N. Padmasali, and S.G. Kini. Life estimation of high power LED using distribution based reliability analysis. in 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). 2017. IEEE.

[5] Modarres, M., M.P. Kaminskiy, and V. Krivtsov, Reliability engineering and risk analysis: a practical guide. 2016: CRC press.

[6] Transmission and D. Committee, IEEE Guide for Electric Power Distribution Reliability Indices. IEEE Std, 2003. 1366: p. 2003.

[7] Shakhatreh, M.K., A.J. Lemonte, and G.M. Arenas, The log-normal modified Weibull distribution and its reliability implications. Reliability Engineering & System Safety, 2019. 188: p. 6-22.

[8] Pham, H., Springer handbook of engineering statistics. 2006: Springer Science & Business Media.

[9] Krit, M., Goodness-of-fit tests in reliability: Weibull distribution and imperfect maintenance models. 2014, Université de Grenoble.

[10] Alptekin, D., New goodness-of-fit tests based on empirical distribution function, Doctoral dissertation, Hacettepe University Graduate School of Sciene and Engineering, Ankara, 2018.

[11] Gorjian Jolfaei, N., et al., Reliability modelling with redundancy—A case study of power generation engines in a wastewater treatment plant. Quality and Reliability Engineering International, 2020. 36(2): p. 784-796.

[12] Kumar, A.R. and V. Krishnan, A Study on System Reliability in Weibull Distribution. Methods, 2017. 5(3).

[13] Li, C., Teng, Y., An, L., & Dan, Q., Imprecise reliability of lifetime data based on three-parameter generalized inverse Weibull distribution. in 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia). 2019. IEEE.

[14] Volkanovski, A., M. Čepin, and B. Mavko, Application of the fault tree analysis for assessment of power system reliability. Reliability Engineering & System Safety, 2009. 94(6): p. 1116-1127.

[15] Tur, M.R., Reliability Assessment of Distribution Power System When Considering Energy Storage Configuration Technique. IEEE Access, 2020. 8: p. 77962-77971.

[16] Deng, J., Y. Liu, and H. Li. Reliability Analysis of Rotating Equipment Based on Weibull Distribution. in 2019 Chinese Control And Decision Conference (CCDC). 2019. IEEE.

[17] Roshan, S., Mohan, B. R., Kumar, M., & Vandanapu, N. Model selection among log-normal, Weibull, Gamma and generalized exponential distributions. in 2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO). 2017. IEEE.

[18] Gupta, R.C., M. Ghitany, and D. Al-Mutairi, Estimation of reliability from a bivariate log-normal data. Journal of Statistical Computation and Simulation, 2013. 83(6): p. 1068-1081.

[19] Raqab, M.Z., S.A. Al-Awadhi, and D. Kundu, Discriminating among Weibull, log-normal, and log-logistic distributions. Communications in Statistics-Simulation and Computation, 2018. 47(5): p. 1397-1419.

[20] Dey, A.K. and D. Kundu, Discriminating among the log-normal, Weibull, and generalized exponential distributions. IEEE Transactions on reliability, 2009. 58(3): p. 416-424.

[21] Fiondella, L. and L. Xing, Discrete and continuous reliability models for systems with identically distributed correlated components. Reliability Engineering & System Safety, 2015. 133: p. 1-10.

[22] Ross, S.M., Introduction to probability and statistics for engineers and scientists (Associated Press). 2009.

[23] Montgomery, D.C., G.C. Runger, and N.F. Hubele, Engineering statistics. 2009. Hoboken, NJ: Weily. 4.

[24] Dufour, J.M., et al., Simulation-based finite sample normality tests in linear regressions. The Econometrics Journal, 1998. 1(1): p. 154-173.

[25] Han, D., X. Tan, and P. Shi. Clutter distribution identification based on anderson-darling test. in 2017 3rd IEEE International Conference on Computer and Communications (ICCC). 2017. IEEE.

[26] Giles, D.E., A saddlepoint approximation to the distribution function of the Anderson-Darling test statistic. Communications in Statistics-Simulation and Computation, 2001. 30(4): p. 899-905.

[27] Tolikas, K. and S. Heravi, The Anderson–Darling goodness-of-fit test statistic for the three-parameter lognormal distribution. Communications in Statistics—Theory and Methods, 2008. 37(19): p. 3135-3143.

[28] Jäntschi, L. and S.D. Bolboacă, Computation of probability associated with Anderson–Darling statistic. Mathematics, 2018. 6(6): p. 88.

[29] Parchami, A., S.M. Taheri, and M. Mashinchi, Testing fuzzy hypotheses based on vague observations: a p-value approach. Statistical Papers, 2012. 53(2): p. 469-484.

[30] Hair, J. F., Black, W., Babin, B., & Anderson, R., Multivariate Data Analysis. Always learning. 2013, Pearson Education Limited.

[31] Tabachnick, B.G. and L.S. Fidell, Using Multivariate Statistics, Always Learning. 2013, Boston, MA: Pearson.