Applied Mathematical Sciences, Vol. 6, 2012, no. 127, 6327 - 6339

The P.d.f Fitting to Time between Failure for High Power Stations

Faris Mahdi Alwan

School of Mathematical Sciences
University Sains Malaysia
11800 USM, Penang, Malaysia
faris.or@gmail.com

Adam Baharum

School of Mathematical Sciences
University Sains Malaysia
11800 USM, Penang, Malaysia

Saad Talib Hasson

Computer Science Department
College of Science
University of Babylon, Iraq

Abstract

Reliability is one of the essential factors that affects the performance of power stations. The current study deals with step down station transformers that transform electricity from 33000 KV to 11000 KV. The process of the fitting model requires data collection and analysis for use in estimating the parameters required to ascertain the suitable probability distributions. In the present paper, we determined the distribution fitting to time between failure (TBF). The model data were collected from ten stations from an electricity distribution company in Iraq. This paper describes the distribution fitting for one station based on failure data collection, calculated TBF, plotted the histogram for TBF, and matched the plot on the continuous distributions' functions. After conducting data analysis, the most valid distribution was found to be the Exponential distribution.

Keywords: Reliability function, hazard function, failure rate, distribution fitting

1 Introduction

The term 'reliability' was first used in scientific research during World War I to describe engine failures. Thereafter, the term was used to describe the level of safety in any device or system with reference to failure rate. Studying the concept of reliability and failure of devices started in1930. The issue of maintenance or replacement of complex systems has become crucial in recent era. The performance reliability associated with decision makers using scientific tools involves either maintenances or replacements, including systems/subsystems depending on the optimization of cost and time of the maintenance or replacement and system operation efficiency and its lifetime. Among the three known types of connections in system reliability *(series, parallel, and mixed)*, the difficulty lies in mixed connection. Calculating the reliability of an individual part is easier than calculating the reliability of the system as a whole. With the expansion of various industries and multinational businesses, it is necessary to get estimations of the efficiency of the performance of each system and their reliability. Calculating system reliability may be done by separating each system into subsystems (units or elements). Their failure rates may be estimated based on the test or simulated data to reach a system failure distribution.

Maintenance, which is one of the main factors affecting system reliability, is also called preventive maintenance. The preventive maintenance function has increased rapidly due to its important role in keeping and improving the availability, product quality, safety requirements, and plant cost-effectiveness levels. Maintenance costs represent an ideal part of the operating budget of manufacturing firms [1]. In 1981, *Nguyen* and *Murthy* [12] studied problems related to the subject of "optimal preventive maintenance policies for viable reform systems." In their paper, they discussed two kinds of optimal preventive maintenance policies to reduce the expected cost for each time unit for an unfinished period of time; they also provided an algorithm to calculate the optimal solutions. In 1984, *Winterbottom* [14] summarized the difficulties faced by anyone who tried to estimate the period for reliability of the system from data of components test either by statistical or mechanical methods. He provided a test for various statistical methods, which can be used to achieve this purpose. Meanwhile, *Pensky* and *Singh* [13] and *El-Gohary* [4] used Bayes function to estimate the reliability, survival average time, and variance for Exponential, Gamma and Weibull distributions, respectively, by presenting a recursive system for finding the probability of ruin and the distribution of severity of ruin in a particular annual based on the hybrid hazard rate recursion rule. In 2006, *Xiaojaun*, *Lifeng* and *Jay* [15] developed a dynamic opportunistic maintenance policy for a continuously monitored multi-unit series system integrating imperfect effect into maintenance activities. In 2007, the same researchers tried to integrate sequential imperfect maintenance policy into condition-based predictive maintenance (CBPM). Another study subjected a system to degradation and then monitored it continuously and perfectly. When the system reliability reached the threshold R, they then performed an imperfect predictive maintenance (PM), leading to a gradual decrease in maintenance [16]. In 2008, *Guo* and *Yang* submitted a new technique to automatically create Markov's model for the reliability assessment of safety-instrumented systems. They generated a framework on voting and failure modes, and then incorporated this framework into repairs and common-cause failures to build the complete Markov model [6].

This paper is arranged as follows. In section 2, we introduce the failure functions, and show some of the basic concepts of failure functions. The relationships between the failure functions have been studied in Section 3. Section 4 is dedicated a problem statement which illustrating the case study of a power station type 33/11 KV for electric distribution. This station as described in the records of the maintenance department stations from an electricity distribution company in Baghdad, Iraq. The goodness of fit for statistical distribution, and the distribution fitting software "*Statgraphics*" commentary obtained in the research methodology (Section 5). In section 6, collected and analysis of data is revealed. Finally, the discussion of the results and conclusions will be presented in section 7.

2 Failure Functions

There are two general cases for every machine: it may be operating well (working) or it may not be operating at all (fail). The failure case has three possible situations: waiting for the repair, the repair process, and achieving complete repair. To build and estimate a good reliable system that may be used to enhance a power station's operations and optimize its operation times, some important required data must be collected or estimated. These data may be in the form of probability distributions, such as probability density function (*p.d.f*), denoted by f(t); cumulative distribution function (*C.D.F*), denoted by F(t); reliability function, denoted by *R(t)*; and hazard function, denoted as *h(t)* and others. These functions are called "failure functions."

2.1 Basic Concepts of Failure Functions

There are many factors and definitions related to reliability. The most important of these are the following:

I-***Failure***: It is defined as disabilities of the system (subsystem or one of its components) to perform its job [5], or the "inability of the item to meet the requirements of the work" [3].

II-***Available***: The state of an item, such that it can perform its function

under stated conditions of use and maintenance in the required location [2].
III-***Availability***: Most researchers define availability as the probability that an item, will be available [2] or the probability that the system is operating satisfactorily at any point in time when operating under a specified condition [10]. It can be expressed mathematically as follows [2, 10]:

$$Availability = \frac{Up\ time\ (Operating\ Time)}{Up\ time\ (Operating\ Time) + down\ time\ (Excluding\ Free\ Time\ )}$$

For more accurate quantity, "inherent availability" is defined as [2, 10]:

$$Inherentavailability = \frac{Up\ time\ (Operating\ Time)}{Up\ time\ (Operating\ Time) + Active\ Repair\ Time}$$

IV-***Maintainability***: Maintainability can be defined as the probability that a failed system can be restored to operating conditions in a specified interval of downtime, which excludes any free downtime [10], or the probability that the equipment will be restored to specified conditions within a given period of time T when the maintenance action is performed in accordance with prescribed procedures and resources [8].

There are many ways to define *reliability*. Colloquially, reliability is the property in which equipment works when we want to use it. By necessity, more formal definitions of reliability must account for whether or not an item performs at or above a specified standard, how long it is able to perform at that standard, and the conditions under which it is operated. The reliability of an electrical switch, for example, may be defined as the probability that it successfully functions under a specified load and at a particular temperature [7].

3 Relationships between Failure Functions

Suppose a nonnegative random variable T ($T \geq 0$) that denotes the failure time of a unit has a cumulative probability distribution $F_T(t) = P(T \leq t)$ with right continuous, and a probability density function $f_T(t)$ $(0 \leq t < \infty)$; i.e., $f_T(t) = dF_T(t)/dt$ and $F_T(t) = \int_0^t f(u)\ du$. They are called *failure time distribution* and *failure density function* in reliability theory, and are sometimes simply called failure distribution $F_T(t)$ and density function $f_T(t)$ [11].
Failure density function $f_T(t)$ is the probability of machine work time from period (t) to $(t + \Delta t)$, which is also sometimes called unconditional function [8] expressed by:

$$f_T(t) = \lim_{\Delta t \to 0} \frac{P(t < T < t + \Delta t)}{\Delta t} \quad t \geq 0$$

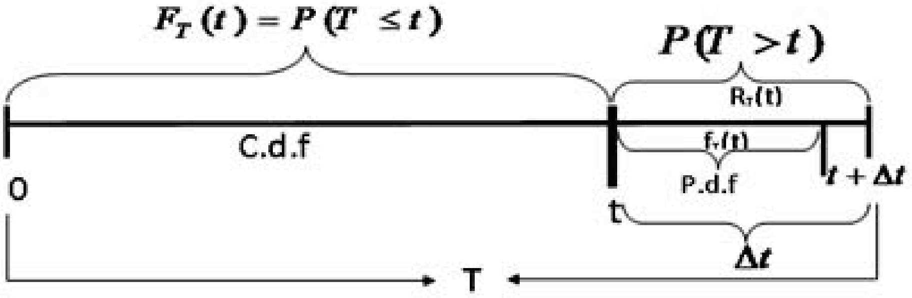

Figure 1: Relationships among $f_T(t)$, $F_T(t)$, and $R_T(t)$

The survival distribution of T is [11] given by:

$$R_T(t) = P(T > t) = 1 - F_T(t) = \int_t^\infty f_T(u)\, du = \overline{F_T(t)} \tag{1}$$

It means that:

$$MTBF = E(T) = \int_0^\infty t\ .\ f_T(t)\, dt = \int_0^\infty R_T(t) = \mu \tag{2}$$

In which, μ and E(T) are the mean lifetime. The failure distribution $F_T(t)$ increases as time increases, i.e., from 0 to 1, $F_T(0) = 0$, and $F_T(\infty) = \lim_{t\to\infty} F_T(t) = 1$ [11]. The reliability function $R_T(t)$ is complementary to the cumulative distribution function, where $R_T(t)$ is decreasing from 0 to 1, i.e., $R_T(0) = 1$ and $R_T(\infty) = 0$. A simple graphical representation can be developed to indicate the relationships among $f_T(t)$, $F_T(t)$, and $R_T(t)$, which are shown in Fig. (1). This representation can help the decision makers decide on the correct design and development of their power stations and equipment.

3.1 Failure Rate Function

Hazard function $h_T(t)$ is the further method of specifying the characteristics of random variable. It is called the conditional probability, because it studies the probability that an item will fail in the time interval $[t, t + \Delta t]$ when that item is working at time t [7]. The probability that an equipment in the system will fail in time interval $[t, t+\Delta t]$, given that the equipment is working at time t, can be written as [7]:

$$P(t < T \leq t + \Delta t \mid T > t) = \frac{P(t < T \leq t + \Delta t)}{P(T > t)} = \frac{F(t + \Delta t) - F(t)}{R(t)}$$

To find and estimate the failure rate, this equipment must be divided by the length of the interval Δt, and letting $\Delta t \to 0$. This gives [7]:

$$h_T(t) = \lim_{\Delta t \to 0} \frac{P(t < T < t + \Delta t \mid T > t)}{\Delta t}$$

$$h_T(t) = \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \frac{1}{R(t)}$$

The first term on the right-hand side is the derivative of the cumulative distribution function, $F(t)$, which is the probability density function, $f(t)$ [7]. Therefore,

$$h_T(t) = \frac{f_T(t)}{R_T(t)} \tag{3}$$

4 Problem Statement

The current paper represents a case study of a power station type 33/11 KV for electric distribution. The data were collated from the principal records of the maintenance department stations. The main problem faced was that the failure data were record manually. To address this, we wrote the dates of breakdowns for these stations and calculated them together with the TBF for the period under a case study. For example, the first breakdown was on 21st Jan, and the second breakdown was on 27th Apr; the operation time TBF was equal to 94 days. The period was for five years. The sample for a case study included ten stations, which we referred to as Station 1, Station 2 and so on, because the data were confidential.

We studied and analyzed the TBF from an electricity distribution company in Baghdad, Iraq, where we visited the maintenance department and met with the engineers and technicians. We also listened to the primary technical explanations necessary to understand the nature of the work of stations in terms of engineering side. These meetings allowed us to study the reliability of these stations and find the optimal method to maintain them. This list was also needed for further study and analysis, in light of the difficult conditions and scarcity of electric power in Iraq. Furthermore, the meetings took place for several days, accompanied by the codification of technical notes and the experiences of workers repairing these stations to aid our study of these phenomena.

The power stations under a case study included *Two/Three Transformers* depending on the type source. Each one of these transformers had a circuit breaker with limited capacities (1200 A) that acted as the main circuit breaker for the transformers. Connected between the conduction pieces are the Bas-Bar, which are linked with a group of feeders to each of the transformers. The first, second, and third transformers are separated by circuit breakers with limited capacities of 800 A, called the Bas-Section circuit breaker. Each feeder has a circuit breaker with a capacity of 400 A. The main circuit breaker should be switched ON, and the Bas-Section circuit breaker should be switched OFF, if the transformers are operating. However, if one of these transformers stops due to any failure, the circuit breakers for these transformers should have to be switched OFF, and the Bas-Section circuit breaker is switched ON to provide electricity to the broken transformer feeders.Fig. 2 shows the station as described in the records of the chamber referred to above. Through study and analysis, we created a representation of Fig. 2 for scientific verification and analysis. This is shown in Fig (3).

5 Research Methodology

The probability distribution of failure time is an important aspect for a failure rate. The failure rate for any equipment during the lifetime period has two characteristics: 1- it is constant with time (constant failure rate), and 2- it varies with time (failure rate changing with time). If the hazard function is invariable, then the failures occur with the same frequency during any equal period of time. The exponential distribution has a constant hazard rate. For another distribution, the hazard function is variable, so the failure rate is not constant with time.

In the current paper, the main focus is on the performance of a station component that fails randomly, i.e., the TBF is a random variable. In this case, a statistical function to identify a statistical distribution to TBF was studied. A goodness of fit for this statistical distribution was tested, including the use of the Kolmogorov-Smirnov, and Chi-square test. We also used the distribution fitting software "*Statgraphics*" to display the goodness of fit reports, including the test statistics and critical values calculated for significance level (0.01). The histogram was based on sample data. To define the number of vertical bars based on the total number of observations, we used the equation, $Q = 1 + log_2 N$, where N is the total number of TBF, and Q is the resulting number of classes. The height of each histogram bar indicates how many of the data points fall into that class. Distribution graph are used to support the result of goodness of fit.

6 Data Collection and Analysis

Data for TBF were collected from an electricity distribution company in Baghdad, Iraq. The sample included ten stations and the study period was for five years. After analysis and testing the data under many distributions using the Statgraphics software, we found through the optimal analysis of the data that

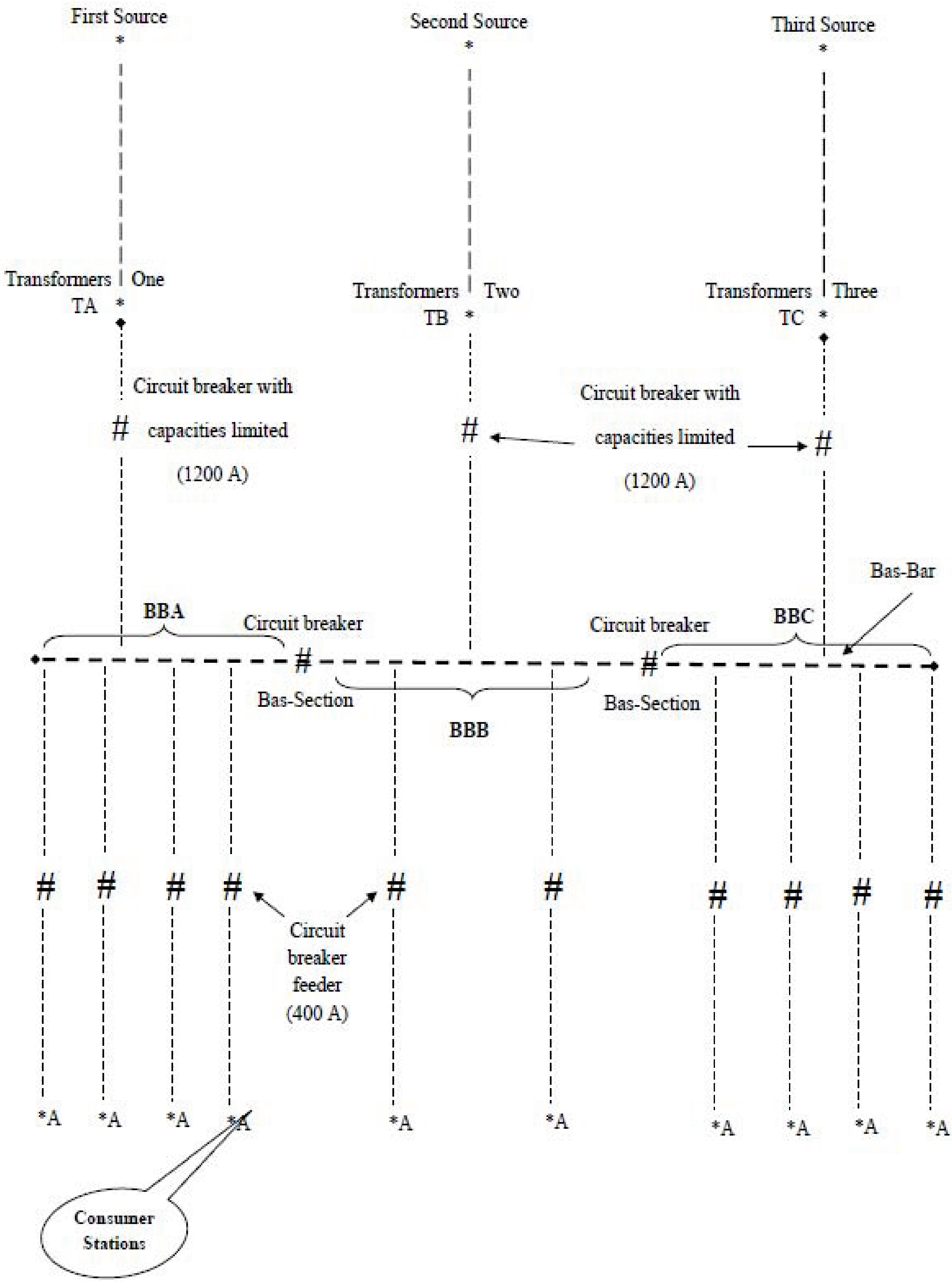

Figure 2: Actual high-power station plan located in the records of the Department of Electricity Distribution Baghdad

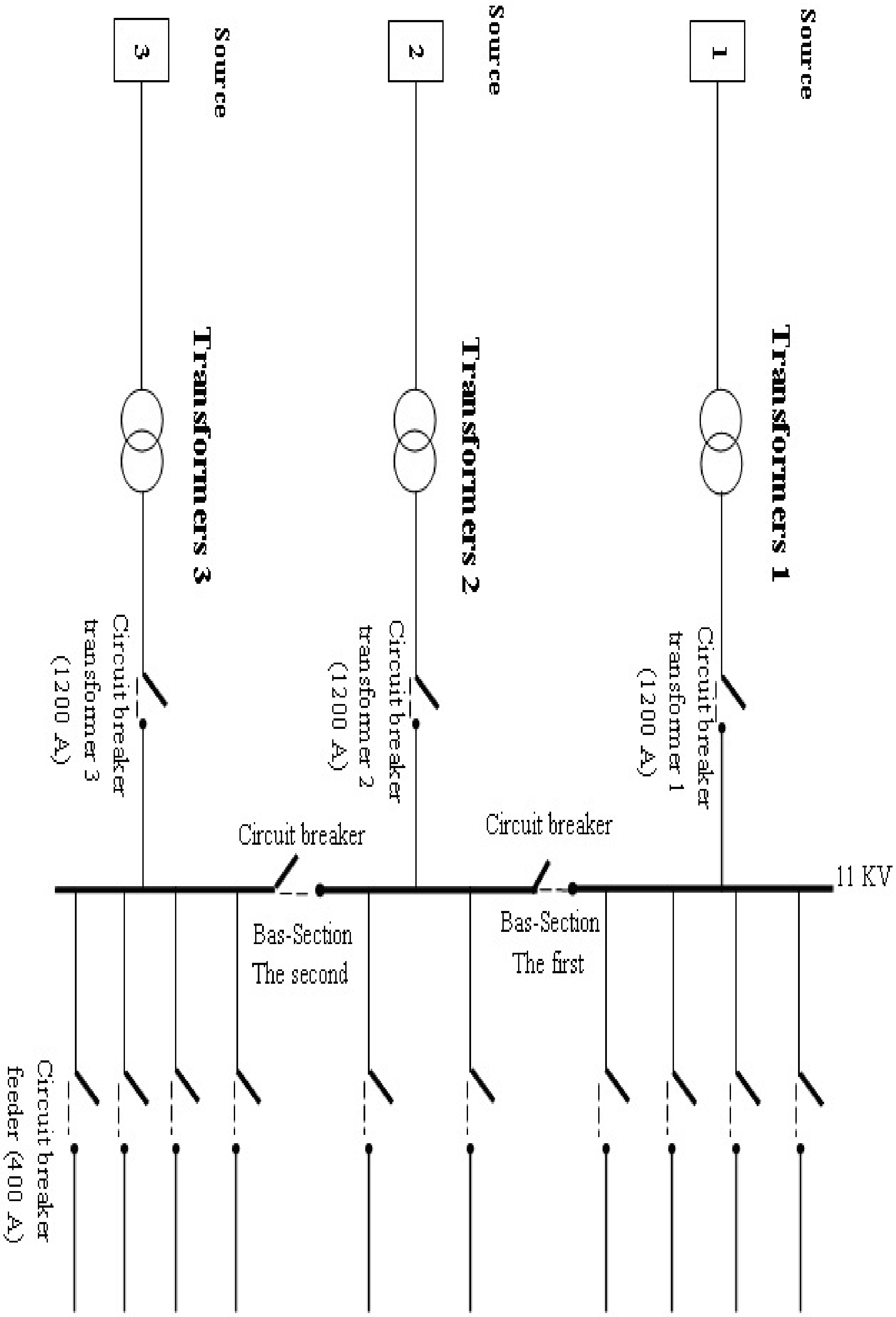

Figure 3: Illustration of the high-power station 33/11 KV

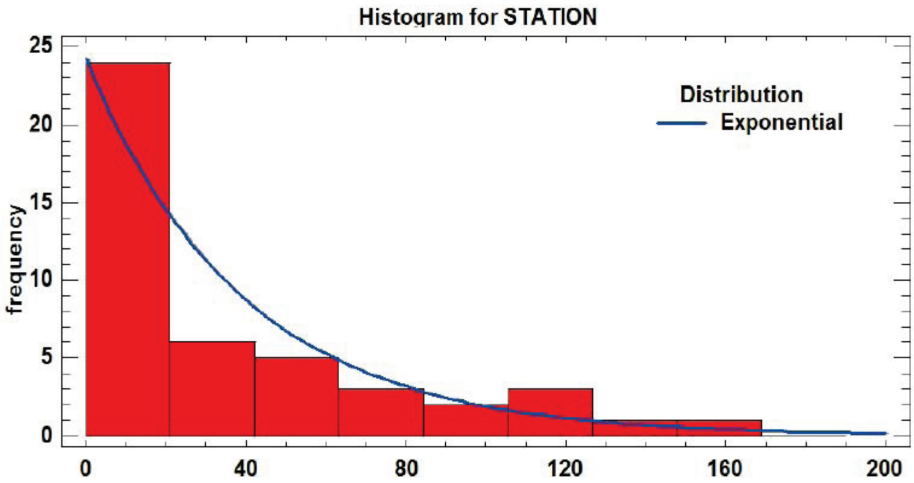

Figure 4: *p.d.f* for the Exponential distribution and TBF for Station one

they follow the Exponential distribution. The idea underlying the goodness of fit tests is to measure the "distance" between the data and the distribution being tested, and then comparing that distance to some threshold value. The goodness of fit reports that involve the *P-Value* and *chi-square* value calculated for significance level of 0.01. Furthermore, the goodness of fit test statistics indicates the distance between the data and the provided distributions. It is obvious that the distribution with the lowest statistic value is the best-fitting model.The P-value can be helpful specifically when the null hypothesis is rejected at all selected significance levels, and we need to know at which level it could be accepted [9]. The results of the analysis for much closer distributions are shown in Figs. 4. The results for goodness of fit are shown in Tables 1. The *hypotheses* are as follows:
H_0: The data follow exponential distribution.
H_1: The data not follow exponential distribution.

Chi-Square = 7.6 with 7 d.f. P-Value = 0.369182
Estimated Kolmogorov statistic DPLUS = 0.138757
Estimated Kolmogorov statistic DMINUS = 0.0771397
Estimated overall statistic DN = 0.138757
Approximate P-Value = 0.354545

Table 1: **The details for goodness of fit (Exponential distribution)**

Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chi-Square
below	4.6135	8	5.00	1.80
4.6135	9.84385	5	5.00	0.00
9.84385	15.8818	6	5.00	0.20
15.8818	23.0233	6	5.00	0.20
23.0233	31.7637	2	5.00	1.80
31.7637	43.032	3	5.00	0.80
43.032	58.9139	2	5.00	1.80
58.9139	86.0641	6	5.00	0.20
above	86.0641	7	5.00	0.80

7 Result and Conclusion

In this work, the TBF has been analyzed in order to find the best fitting distribution. We calculated the number of failures based on original failure data from Station one. After running the software and recording the optimal distribution, the values of *chi-square calculated* (χ^2_{Cal}) and *chi-square table* (χ^2_T) with 7 degrees of freedom at 0.01 significant level were compared. The value of χ^2_T was 18.48, and the value of χ^2_{Cal} was 7.6, this means that the value of χ^2_{Cal} was less than the value of χ^2_T, thus the null hypothesis can not be reject. We found that the Exponential distribution is the best distribution. Exponential distribution is one of the most important distributions for reliability. This function can be easily derived, and it shows the constant failure rate. The probability density function $f_T(t)$ for an exponential failure time t is $f_T(t;\lambda) = \lambda e^{-\lambda t} \;\; t > 0, \; \lambda > 0$; the reliability function is $R_T(t;\lambda) = e^{-\lambda t}$ the hazard function is $h_T(t;\lambda) = \lambda$. The exponential distribution is the only continuous distribution with the memoryless property, which means that the past has no influence on the future. If the equipment is working successfully in time t, the probability that it will work after an additional Δt time period is the same as the (unconditional) probability that it will work after time Δt, forgetting that it has made past time t.

$$P(x \geq t + \Delta t \mid x \geq t) = \frac{\int_{t+\Delta t}^{\infty} \lambda e^{-\lambda t}\, dt}{\int_{t}^{\infty} \lambda e^{-\lambda t}\, dt}$$

$$P(x \geq t + \Delta t \mid x \geq t) = e^{-\lambda t} = P(x \geq \Delta t)$$

This unique property of the exponential distribution denotes that the system for this station has a stable failure rate. Thus, the maintenance program is applied; the equipment will work like new.

In future our research will aim will be to find the reliability for each component of the station. We will also calculate the total reliability of the station regardless of whether the station has *sequential, parallel, or mixed* system. We will then develop a mathematical maintenance model for the station.

ACKNOWLEDGEMENTS. The researcher would like to thank the school of Mathematical Sciences, USM for financial support.

References

[1] I. A. B. Al-Najjar, Selecting the most efficient maintenance approach using fuzzy multiple criteria decision making, *Int. J. Production Economics,* **84** (2003), 85 - 100.

[2] A. D. S. Carter, *Mechanical Reliability,* macmilan Education Ltd, London, 1986.

[3] A. D. S. Carter, *Mechanical Reliability and Design,* John Wiley and Sons , New York, 1997.

[4] A. El-Gohary, Estimations Of Parameters In A Three State Reliability Semi-Markov Model, *Applied Mathematics and Computation,* **154** (2004), 389 - 403.

[5] E. G. Frankel, *Systems Reliability and Risk Analysis,* Springer, New York, 1988.

[6] H. Guo and X. Yang, Automatic Creation Of Markov Models For Reliability Assessment Of Safety Instrumented Systems, *Reliability Engineering and System Safety,* **93** (2008), 807 - 815.

[7] M.S. Hamada, et al., *Bayesian Reliability,* Springer Science+Business Media LLC, New York, 2008.

[8] A. K. S. Jardine, *Maintenance, Replacement And Reliability,* Pitman, New York, 1973.

[9] S. K. Jha, A. K. D. Dwivedi and A.Tiwari, Necessity of Goodness of Fit Tests in Research and Development, *INTERNATIONAL JOURNAL oF COMPUTER SCIENCE and TECHNOLOGY,* **2** (2011), 135 - 141.

[10] H. F. Martz and R. A. Waller, *Bayesian Reliability Analysis,* John Wiley and Sons, United States of America, 1982.

[11] T. Nakagawa, *Maintenance Theory of Reliability,* : Springer-Verlag London Limited, London, 2005.

[12] D. G. Nguyen and D. N. P. Murthy, Optimal Preventive Maintenance Policies for Repairable Systems, *Operations Research,* **29** (1981), 1181 - 1194.

[13] M. Pensky and R. S. Singh, Empirical Bayes Estimation of Reliability Characteristics For an Exponential Family, *The Canadian Journal of Statistics,* **27** (1999), 127 - 136.

[14] A. Winterbottom, The Interval Estimation Of System Reliability From Component Test Data, *Operations Research,* **32** (1984), 628 - 640.

[15] X. Zhou, L. Xi and J. Lee, A Dynamic Opportunistic Maintenance Policy For Continuously Monitored Systems, *Journal of Quality in Maintenance Engineering,* **12** (2006), 294 - 305.

[16] X. Zhou, L. Xi and J. Lee, Reliability-Centred Predictive Maintenance Scheduling For A Continuously Monitored System Subject To Degradation, *Reliability Engineering and System Safety,* **92** (2007), 530 - 534.

Received: July, 2012