Comparison of the Weibull and the Crow-AMSSA model in prediction of early failures among power cables

Tang, Zeyang; Zhou, Wenjun; Zhao, Jiankang; Wang, Dajiang; Zhang, Leiqi; Liu, Haizhi; Yang, Yang; Zhou, Chengke

Published in:
IEEE Transactions on Power Delivery

DOI:
10.1109/TPWRD.2015.2404926

Publication date:
2015

Document Version
Peer reviewed version

Comparison of the Weibull and the Crow-AMSAA Model in Prediction of Early Cable Joint Failures

Zeyang Tang, Student *Member, IEEE*, Wenjun Zhou, Senior *Member, IEEE*, Dajiang Wang, Leiqi Zhang, Haizhi Liu, Yang Yang, Chengke Zhou, Senior *Member, IEEE*

***Abstract*—This paper compares the application of the Weibull distribution and the Crow-AMSAA (C-A) model to the analysis of cable joint failures. The procedures of how to use the two models to analyze failure data and to predict future number of failures have been described before the models are applied to a set of early-failure data. The data which include 16 failures and 1126 suspensions were collected from a regional power supply company in China. It is observed that the Weibull and the C-A model produce opposite results in terms of β value when the dataset contains failures of multiple years where the failure rate in early years differs significantly from those in later period. The paper shows that, when applying the C-A model, separating those data into subsections and analyzing them independently can yield useful information. Recent failure data can better reflect the current state of cable joints. The paper also proves that the Weibull distribution provides more reliable results in the analysis of early-failure data. The results of this paper should help utility asset managers to better analyze their past failure data.**

***Index Terms*—Weibull; Crow-AMSAA; early-failure; failure prediction; power cables, asset management**

I. Introduction

Power cables and cable accessories are subject to electrical, thermal, mechanical, and environmental stresses on a constant basis when in service. These stresses together with poor practice in installation and maintenance often lead to insulation degradation or defects causing cable breakdowns [1]. Like other power systems assets, the lifetime of cable and accessory failures obey the "bathtub curve" [2] which can be divided into "burn-in phase" with a decreasing rate of early failures (0~5 years), "the useful life phase" with a low number of casual failures (5~25years) and "the wear-out phase" with an increasing rate of aging related failures (>25years) [3]. Early failures usually result from imperfections during manufacturing process, defects associated with poor installation practice and third party damages. During the useful life phase, failures happen occasionally due to various reasons such as third party damage, wear-out of components and environmental stress etc. As time progresses, the bulk dielectric strength degrades, and artifacts such as water ingress and detachments at material interfaces raise local stress. The net effect appears as aging, the rate of which depends on many factors such as voltage, thermal stresses, maintenance, system age, cable system technology, and environment [4].

In developing countries, the cable network is relatively new and still growing rapidly. Take China for example, the cables laid down over the last 10 years account for more than half of the total volumes [5]. Most cable failures are due to third party damages, manufacturing and poor installation problems [6]. In contrast, in developed countries such as in the UK, installation peaked in 1950s and 1960s [7]. A large proportion of the cable assets have already expired or are approaching their end of design life, where a higher proportion of age related failures have been reported [8]. Despite the differences in failure mechanisms, failure prediction is important for cable asset managers to arrange appropriate maintenance programs under both situations.

Among statistical models, the Weibull distribution and the Crow-Army Material System Analysis Activity (AMSAA), have been used to carry out failure predictions. The Weibull distribution has been used by R.M. Bucci [9] to make failure prediction of underground distribution feeder cables where data were simply sorted according to the age of failed cables without considering the modes or causes of the failures. John P. Ainscough P. E [10] used the Weibull distribution to predict medium voltage underground distribution cable failures. The C-A model was employed by Yancy Gill [11-12] to establish a maintenance model of aging cable. Paul Barringer, P.E. [13] compared the Weibull distribution with the C-A model and concluded that the C-A model worked well with mixed failure modes while the Weibull distribution was a powerful single failure mode tool. These papers mainly focused on age-related data, while the performances of the Weibull and C-A model on early-failure data have not yet been studied. One of the main objectives of early-failure data analysis is to establish a pattern of early failures [14].

In this paper, early cable failure data, collected from a regional power supply company in China, is divided into groups based on the failure causes. As the work presented here focuses on early-failure data, age related failure data are treated as suspensions. The procedures concerning how to apply the Weibull distribution and the C-A model to predict failures are thoroughly analyzed and described. The performances of two

Manuscript received

Zeyang Tang, Wenjun Zhou and Yang Yang are with the school of Electrical Engineering, Wuhan University, Wuhan, 430072, China (e-mail: whu_tzy@126.com; wjzhou@whu.edu.cn; youngyoung1129@yahoo.cn).

Chengke Zhou is with School of Engineering and Built Environment, Glasgow Caledonian University, G4 0BA UK. (e-mail: c.zhou@gcu.ac.uk).

Dajiang Wang, Leiqi Zhang and Haizhi Liu are with Wuhan power supply company, Wuhan, 430054, China.

models in dealing with early-failure data are investigated and critically compared.

II. STATISTICAL MODELS

A. Weibull Distribution

The Weibull distribution is perhaps the most widely used model in the analysis of reliability and failure data. It gives the lifetime distribution of objects and was originally proposed to quantify fatigue data [15-16], but it is also used in the analysis of systems involving the "weakest link" such as insulations in power plant.

Its flexibility to model all the three phases of a reliability bathtub curve makes it attractive to reliability and maintenance engineers. It is found that it can fit most lifetime data better than other distributions and is particularly valuable for relatively small samples of the data which are often encountered by maintenance engineers.

There are two versions of the Weibull model, namely the two-parameter and the three-parameter models. Mathematically, the cumulative probability of failure of the two-parameter model, as a function of time, is given in Equation (1). The three parameter model, as given in Equation (2), has an introduction of a location parameter into the two-parameter model [17-18].

$$F(t) = 1 - \exp(-(\frac{t}{\eta})^{\beta}) \quad (1)$$

$$F(t) = 1 - \exp(-(\frac{t-g}{\eta})^{\beta}) \quad (2)$$

$$R(t) = 1 - F(t) \quad (3)$$

Where $F(t)$ is the cumulative distribution function or the probability of failure between time 0 and t. $R(t)$ is the reliability or probability of not failing between 0 and time t. η is the scale parameter, β is the shape parameter and g is the location parameter. If β is less than 1, it means that the failure rate is decreasing and the asset group under analysis is in early-failure stage. If β is greater than 1, it indicates an increasing failure rate and that the asset has started to age or has already aged. If β is equal to 1, it stands for a constant failure rate and that the asset group is in a period of useful service age.

The probability density function (PDF) of the two parameter model is the derivative of equation (1), which is given in Equation (4).

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \quad (4)$$

The probability density function defines the life probability distribution of a population. The area under this curve is equal to unity (in terms of probability) or 100% which shows all life possibilities. The probability density function is similar to the normal curve, with a typical bell shape. The only difference which makes the Weibull better for describing life of insulation is that, it has no negative values and can assign a starting point (below which there are no failures) to the life of insulation material. But the normal curve has values from negative to positive infinity [11].

The failure or hazard rate function is given in Equation (5).

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (5)$$

Figure 1 gives an illustration of the Weibull functions which include $f(t)$, $h(t)$ and $F(t)$. Here the unit of horizontal axis is in calendar year. The left vertical corresponds with $f(t)$ and the right vertical signifies $h(t)$ and $F(t)$.

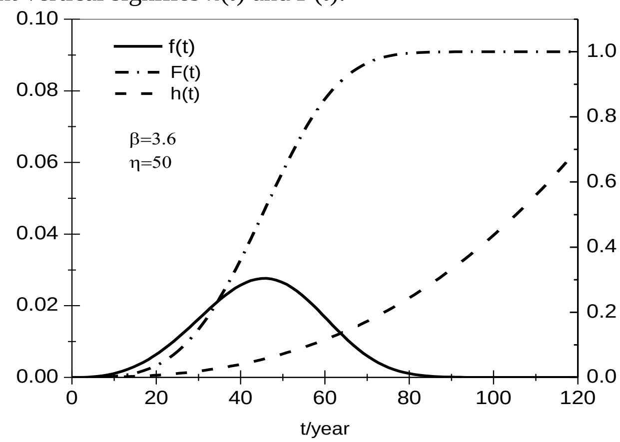

Figure 1 Illustration of the Weibull functions

B. Crow-AMSAA model

The C-A model was originally developed to track and quantify the reliability growth of preliminary product designs or in manufacturing processes to help in production of a product or process when adequate reliability is achieved [16]. However, over the past several years, the C-A model has been used increasingly as a tool to monitor reliability and to forecast failures/faults in field mechanical and electrical systems. The advantage of the C-A model is that it models repairable systems. This is an important distinction, as C-A can model a component that has failed and been repaired multiple times, while the Weibull distribution can only be used to model the first failure. The C-A model is also capable of handling a mixture of failure modes whereas the Weibull model works best with one, perhaps two failure modes only [12]. This reduces the requirement for detailed information of time to first failure. The forecast of overall failures is based on cumulative time against cumulative failures and does not need to consider failure modes.

The process where repairs are assumed to return the equipment to the level at which it was operating before failure is known as the Non-Homogeneous Poisson Process (NHPP)[19]. In this case, the process is only time homogeneous when failure rate is a constant over a specific period of time. It can be shown, however, that, if $t_1 < t_2 <\ldots$, are the time at which failure events occur, then failure rate is a constant between the time t_1 and t_2. The expected number of failures in a selected interval is as following.

$$N(t_2) - N(t_1) = \int_{t_1}^{t_2} \rho(t)dt \quad (6)$$

The failure intensity function of the model is given as:

$$\rho(t) = \lambda\beta t^{\beta-1} (\lambda \text{ and } \beta>0) \quad (7)$$

Therefore, the cumulative number of failures as a function of cumulative failure time can be expressed as:

$$N(t) = \lambda t^{\beta} \quad (8)$$

The reciprocal of $\rho(t)$ is the instantaneous Mean Time Between Failure (MTBF). The logarithm of cumulative failures $N(t)$ plotted against logarithm cumulative time is a linear plot, as given in Equation (9).

$$\log N(t) = \log\lambda + \beta\log t \quad (9)$$

In this model, λ is the scale parameter or the intercept on the y-axis in the linear plot as will be shown in a later section of this paper, and β is the growth parameter which is the slope of the line. Like the Weibull distribution, when β is less than 1, the failure rate is decreasing. The failure rate is increasing when β is greater than 1, and constant when β equals to 1. Figure 2 gives an illustration of the failure rate in relation to the value of β.

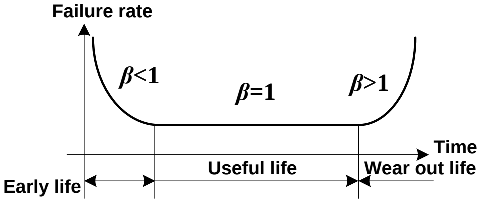

Figure 2 Illustration of the Failure rate

III. FAILURE PREDICTION USING THE WEIBULL AND THE CROW-AMSAA MODELS - A CASE STUDY

Table 1 HV Cable joint failures with a Utility company between 2004 and 2011

No.	Failure causes	Date of Commission	Date of Failure	t/days
1	Quality issue	2010.06.08	2010.06.13	5
2	Installation issue	2007.10.23	2007.12.05	43
3	Installation issue	2007.10.23	2007.12.27	65
4	Quality issue	2009.07.01	2010.01.11	194
5	Quality issue	2008.08.02	2009.04.18	259
6	Quality issue	2007.08.21	2008.05.09	262
7	Quality issue	2008.07.12	2009.07.01	354
8	Quality issue	2003.06.01	2005.02.10	620
9	Quality issue	2008.02.04	2010.09.29	968
10	Quality issue	1998.06.01	2004.03.01	2100
11	Quality issue	2002.07.17	2009.09.27	2629
12	Unknown	1999.09.01	2006.12.29	2676
13	Unknown	1999.09.01	2006.12.29	2676
14	Quality issue	2003.05.30	2010.12.03	2744
15	Unknown	1996.06.01	2008.01.20	4250
16	Quality issue	1996.06.01	2008.01.20	4250

A set of HV cable (rated at 110kV and 220kV) failure data has been collected from a regional power supply company in China. The cable asset involved in the data has a total circuit length of 380km and there were a total of 1142 cable joints. During the period between January 2004 and December 2011, 31 failures were registered. However two of them were registered with an age of 0, and are included as left censored data (Appendix Table A). There were 16 early-failures, all given in Table 1, which will be the focus of this paper. The remaining 13 failures caused by third party damages and aging are not included.

A. *Weibull Distribution*

When the Weibull model is applied to forecast failures, the procedures are as follows.

(1) Calculate the age to failure of all failed items and censored time t (between the date of commissioning and the date of data collection for suspended items), then rank t, for both failed and suspended items, from the smallest to the largest, as shown in Appendix Table A. Here it is very important to include censors because they will provide useful information for the Weibull analysis, which will be illustrated later in this section. The rank of the failed items should be modified by the presence of the censors. The adjusted rank can be calculated by Equation (10) [20].

$$AR_i = \frac{RR \times AR_{i-1} + (n+1)}{RR+1} \quad (10)$$

RR denotes the reverse rank which rank from the largest to smallest. AR_i denotes the ith adjusted rank. n is the total number of sample. $AR_0 = 0$.

(2) Estimate β and η

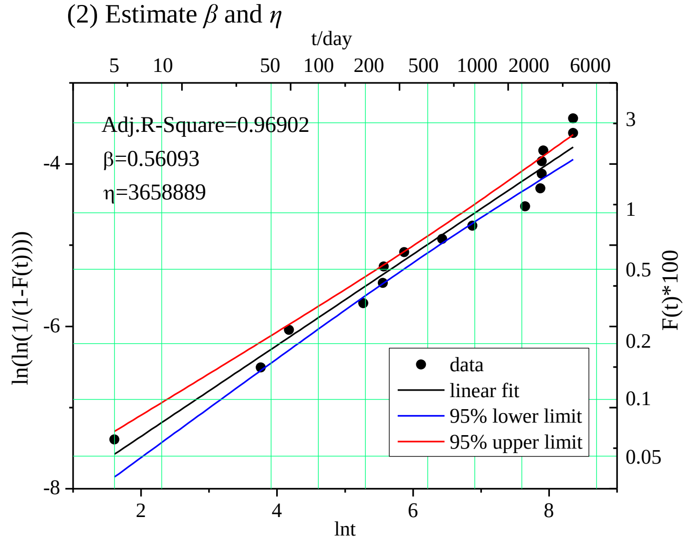

Figure 3 Fitting result of the Weibull distribution. Here the x axis at the top of the Figure denotes the time to failure t. The bottom x axis denotes ln$t$. The left y axis denotes $\ln[\ln\frac{1}{1-F(t)}]$. The right y axis denotes the percentage of the cumulative distribution function.

When the two parameter Weibull model is used for failure prediction, the cumulative distribution function given in Equation (1) is used. If the natural logarithm is taken at both sides of the function, then Equation (11) can be obtained:

$$\ln[\ln\frac{1}{1-F(t)}] = \beta(\ln t - \ln\eta) \quad (11)$$

where $F(t)$ can be calculated by the median rank equation in Equation (12) below:

$$F(t)=\frac{AR-0.3}{n+0.4} \tag{12}$$

let:

$$y=\ln[\ln\frac{1}{1-F(t)}],\ x=\ln t,\ b=\beta,\ a=-\beta\ln\eta \tag{13}$$

then:

$$y=a+b\times x \tag{14}$$

Based on Equations (13) and (14), the software package Origin has been adopted to carry out linear fitting. The fitting results are shown in Figure 3. The Adj.R-Square which measures the quality of the data fitting process is equal to 0.961. The closer Adj.R-Square is to 1, the more accurate the fitting result is. For a particular value x_p, the 95% lower limit and upper limit are obtained using Equations (15) and (16) respectively.

$$\hat{y}-t\left(\alpha/2,\ n^*-1\right)S_E\sqrt{\frac{1}{n}+\frac{(x_p-\overline{x})^2}{SXX}} \tag{15}$$

$$\hat{y}+t\left(\alpha/2,\ n^*-1\right)S_E\sqrt{\frac{1}{n}+\frac{(x_p-\overline{x})^2}{SXX}} \tag{16}$$

Where $\hat{y}=a+x_p\times b$, and the value of $t\left(\alpha/2,\ n^*-1\right)$ can be obtained from the distribution Table. α equals to 5%. n^* equals to n-1. n is number of dataset(in this paper n equals to 16). S_E represents the standard error which can be obtained by $S_E=\frac{\sum_{j=1}^{n}(y_j-\hat{y}_j)^2}{n^*-1}$. $\overline{x}$ can be calculated by $\overline{x}=\frac{1}{n}\sum_{j=1}^{n}x_j$. SXX can be calculated by $SXX=\sum_{j=1}^{n}(x_j-\overline{x})^2$. x_j and y_j are given in Equations (13) and (14).

Based on the results in Figure 3, the Weibull parameter are obtained as β=0.561 and η=3658889 respectively, according to Equation (13). The shape parameter β indicates that the failure rate decreases with age. The scale parameter η is equal to 3658889 meaning that 63.2% faults happened before 3658889 days, or over 10,000 calendar years, meaning that the population can last almost forever if these cable joints can fail only due to manufacturing and installation issues. In fact, when only the 16 failed joints are included in the data (it can also be called complete data), η equals to 1384 which is a pessimistic estimation of the characteristic life of the whole population when censored data are included. The actual life of a cable population will eventually be determined by age related failures.

(3) The time t_c is used to calculate future failures. When the failed joints are replaced or new joints are installed, t_c is the time between the date when cable joints are replaced or installed and the date of data being collected. If a joint is a suspension it means that the joint has not failed, then t_c is equal to t.

(4) Determine the time boundary t_b. For example, in the case of the greatest age to failure of the 16 early cable joint failures, as shown in Table 1, being 4250, 4250 is assumed as the time boundary for early failure. It is assumed that beyond this instant in time, the probability of early failure between t_c and t_b is zero.

(5) Calculate $F(t_{cq})$ and $F(t_{cq}+k)$ for each joint item. Here t_{cq} stands for the t_c of mth item, k is the duration of a period over which failure is to be predicted. It should be noted that $F(t_{cq})$ equals to $F(4250)$ when the time t_c exceeds the time boundary t_b based on the assumption in step (4).

(6) Calculate the expected failures. The expected failures during a period between t_{cq} and $t_{cq}+k$ can be calculated using Equation (17):

$$\text{Expected failures}=\sum_{q=1}^{SN}\left(\frac{F\left(t_{cq}+k\right)-F(t_{cq})}{1-F(t_{cq})}\times JN_q\right) \tag{17}$$

Where $F(t_{cq})$ denotes the accumulated probability of failure for the qth item between the time 0 and t_{cq}. $F(t_{cq}$+k) denotes the accumulated probability of failure for the qth item between the time 0 and $t_{cq}+k$. JN_q denotes the number of joints when the serial number is q.

Assume there are 100 joints whose t_c are 5. How many failures will we have among the 100 joints in the next year? When only the failures are considered in the Weibull analysis (in this situation η equals to 1384, β equals to 0.587), the expected failures will be 36.9 by using Equation (17). While if the suspensions and failures are included, the expected failures will be 0.52. Based on the field experience, it is believed that the results are more reliable when suspensions are included in the Weibull analysis.

Table 2 Failure predictions using the Weibull distribution

Serial number SN	Age to failure or censored time t (day)	Number of joints JN	t_{cq}	F(t_{cq})	F(t_{cq}+365)	F(t_{cq}+365*2)	F(t_{cq}+365*3)	F(t_{cq}+365*4)
1	0	1	1430	0.012174	0.013819	0.015319	0.01671	0.018013
2	0	1	1624	0.013069	0.014632	0.016071	0.017412	0.018675
3	5	1	536	0.007039	0.009408	0.011375	0.0131	0.01466
4	43	1	1457	0.012302	0.013934	0.015426	0.016809	0.018106
5	65	1	1435	0.012198	0.01384	0.015339	0.016728	0.01803
6	97	2	97	0.002704	0.006478	0.008969	0.010999	0.012766
7	125	2	125	0.003116	0.006694	0.009137	0.011142	0.012893
8	131	1	131	0.003199	0.00674	0.009173	0.011173	0.01292
9	144	2	144	0.003373	0.006838	0.00925	0.011239	0.012979
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

298	6119	2	6119	0.022315	0.022315	0.022315	0.022315	0.022315
299	8369	7	8369	0.022315	0.022315	0.022315	0.022315	0.022315
300	8369	5	8369	0.022315	0.022315	0.022315	0.022315	0.022315
301	8369	10	8369	0.022315	0.022315	0.022315	0.022315	0.022315
302	8369	12	8369	0.022315	0.022315	0.022315	0.022315	0.022315

Table 3 Basic failure data required for applying the C-A algorithm

Year	Cumulative time (t)	Number of failures in each year ($N1$)	Cumulative number of failures ($N2$)	Cumulative number of cable joints($TN1$)	Cumulative number of cable joints whose t_c are less than 4250($TN2$)	$\frac{N1}{TN2}*100$	$N=\sum\left(\frac{N1}{TN2}*100\right)$
2004	1	1	1	414	380	0.263	0.263
2005	2	1	2	440	406	0.246	0.509
2006	3	2	4	477	423	0.473	0.982
2007	4	2	6	586	514	0.389	1.371
2008	5	3	9	793	683	0.439	1.811
2009	6	3	12	909	782	0.384	2.194
2010	7	4	16	1029	887	0.451	2.645
2011	8	0	16	1142	953	0	2.645

B. *Crow-AMSAA model*

There are two ways of applying the C-A model. One (Model I) is to analyze failure numbers against time in calendar year as shown in Figure 4, whilst Figure 5 gives the results of C-A model (Model II) analyzing failures against the size of cable joint population.

In this paper, we mainly focus on the early failures. As it has been mentioned before, 4250 days can be assumed as the time boundary for early failure. So it is considered that the cable joint whose age has exceeded 4250 will not suffer from early failures anymore and is not included in censors.

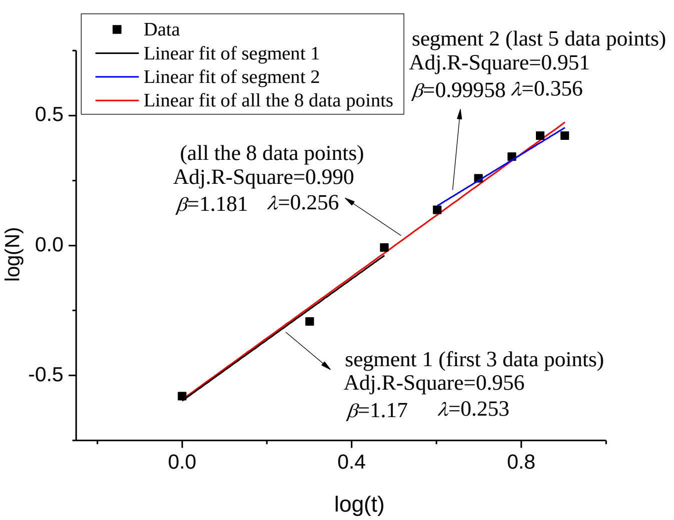

Figure 4 Application of C-A model I

Due to the rapid increase in number of joints since 2007, the data have been divided into two segments. In the first segment there are 3 data points and the second has 5 data, as shown in Figures 4 and 5, where all the 8 data points are also modelled together.

As shown in Figure 4, C-A model I analyzes cumulative time t (log(t) is taken as the x axis) and cumulative number of failures per 100 joints N (log(N) is taken as the y axis).

It can be seen from the results of C-A model I that only the β value of segment 2 is less than one, which indicates that the failure rate (or the number of failures per 100 joints under unit time) is decreasing. It can be concluded that segment 2 can best reflect the current state of the cable. Thus the linear fit result of segment 2 will be used for failure prediction.

When the size of a cable population is still increasing sharply and the age profile of the population changes year-on-year, subsections should be considered in C-A model.

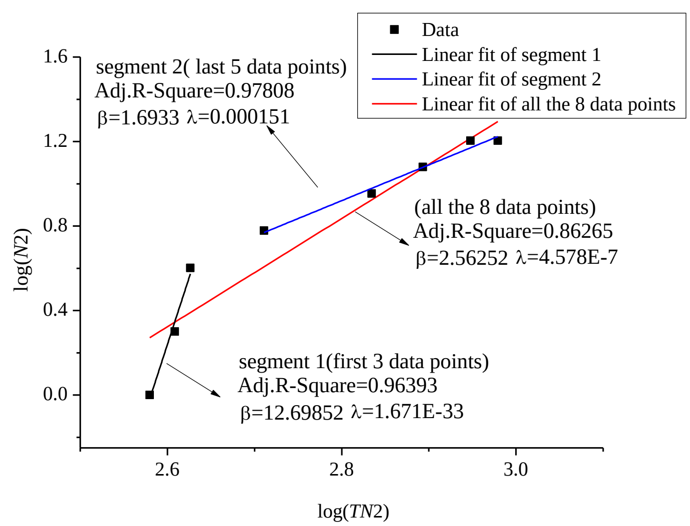

Figure 5 Application of C-A model II

As shown in Figure 5, C-A model II analyzes the cumulative number of joints $TN2$ (log($TN2$) is taken as the x axis) and the cumulative number of failures $N2$ (log($N2$) is taken as y axis).

It can be found in Figure 5, the β value of segment 1, 2 and “all the 8 data points” are greater than 1, which indicates that the failure rate is increasing. The failure rate here is the number of failures under unit number of installed joints according to the physic meaning of C-A model, which is quite different from C-A model I. In this situation, the expected number of failures has little to do with time and is influenced only by the cumulative joints.

When making prediction of the number of early failures, Equation (18) should be used in C-A model I. Although discontinued joints will be replaced by new ones, but the number of the total joints will not be affected by the replaced joints. The total number of joints in this case should be $TN2(t+\Delta t)+TI(t+\Delta t)$ when considering replaced and newly installed joints .

Expected failures=

$$\begin{aligned}&\left(N(t+\Delta t)-N(t)\right)*\left(TN2(t+\Delta t)+TI(t+\Delta t)\right)\\&=\left(\lambda(t+\Delta t)^{\beta}-\lambda t^{\beta}\right)*\left(TN2(t+\Delta t)+TI(t+\Delta t)\right)\end{aligned} \tag{18}$$

Where TI is the number of newly installed cable joints during the period of Δt. It should be noted that during the period of Δt, some joints' t_c will exceed 4250. Thus the number of $TN2$ will decrease as time goes by.

While Equation (19) should be used for failure prediction in C-A model II.

Expected failures=

$$\begin{aligned}&N(t+\Delta t)-N(t)\\&=\lambda(TN2(t+\Delta t)+TI(t+\Delta t))^{\beta}-\lambda(TN2(t)+TI(t))^{\beta}\end{aligned} \tag{19}$$

IV. Results Analysis

As it can be seen in Figure 6, Weibull, C-A model I and II are used for failure prediction. The same dataset has been used in the Weibull and the C-A approaches, but the results have some differences.

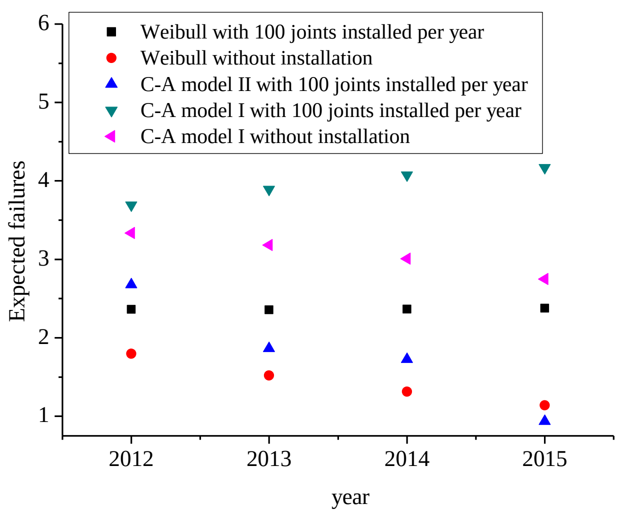

Figure 6 Failure predictions using the C-A model and the Weibull distribution

When the installation is not considered, there is a decrease trend of the expected failures of Weibull and C-A model I. The expected failures of Weibull and C-A model I increase due to the installation is considered and the decrease trend turns into the increase trend. It can be easily concluded that the expected number of failures is relevant both with the past information (failure data which decide the parameters of Weibull and C-A model) and the future information (the number of installed joints per year).

It should be noted when installation is considered, the expected failures of C-A model II has a decrease trend even if the β value of C-A model II is greater than 1. This phenomenon can be explained by the following. As time goes by, the value of $TN2$ decreases due to some of the joints whose t_c have exceed 4250. Despite there are 100 joints installed per year, the increment value of $TN2(t)+TI(t)$ decrease. So the number of predicted failures decreases. It can also be found from Figure 6 that the expected failures of C-A model II are closer to the ones of the Weibull distribution when installation is considered. It can be concluded that C-A model II is more suitable to use when there is a sharp increase in the number of joints.

Clearly when the number of joints does not increase or it actually decrease, the value of $TN2(t+\Delta t)+TI(t+\Delta t)$ is equal or less than $TN2(t)+TI(t)$, the predicted failures of would be zero or minus respectively. In this situation, C-A II model is not suitable for failure predictions.

V. Discussions

Weibull uses the life data, or the details of the dates of commissioning and failure. Weibull applies a failure rate to each individual asset reflecting its real age. Weibull is not straightforward as C-A model when used for failure prediction. It needs to summate the failure probability of each joint. But Weibull can directly and correctly reflect the failure mode.

While C-A model only considers the accumulated failures per year, it does not model the failure rate of the individual asset which changes over time. Although the C-A approach works for data sets that are missing information, which has often been the case with power utilities [21], it does not consider how long a cable has been in service.

Two types of C-A model have been compared in this paper. According to the physic meaning of the C-A model, the failure rate of C-A model I is defined as the number of failures per 100 joints under unit time, while the failure rate of C-A model II is defined as the number of failures under unit number of installed joints. Due to the difference, when C-A model is applied to a situation where there is a sharp change in the number of asset population, C-A model II should be recommended. When the population is relative stable and the failures has much to do with time, C-A model I should be chosen.

It is important for asset managers to be aware of the failure mode and the expected number of failures which are useful to make specific maintenance strategies. It should be very careful to deal with early-failure data when using C-A model. Otherwise, some opposite result will be obtained. The Weibull distribution could be used as a double check of failure mode.

VI. Summary

This paper presented a comparison of the Weibull distribution and the C-A model for prediction of early cable joint failures. The procedures of applying the two models for failure prediction with considerations of installation have been demonstrated. While a case study was carried out, using early cable joint failure data, the expected failures have been compared and analyzed. Further analysis was then conducted to compare the fundamental differences between the two models. Analysis showed that the Weibull distribution, which is based on life data, provides more reliable results about failure mode when the overall population increases rapidly. In the case study, when there is a sharp increase in the number of installed joints during a short period, subsections should be carried out in order to better reflect the current state of cable joints. Despite the limitation that using the cumulative time with cumulative failure does not reveal the change in reliability of the cable joints with their service time, C-A model works with incomplete data and the predictions are more straightforward.

References

[1] X. Duan, M. Tang, F. Lin, and H. Ye, "Analysis on breakdown of main insulation on middle joint of 220 kV power cable," *High Voltage App*., vol. 45, no. 6, pp. 142-144, 2009.

[2] G. A. Klutke, P. C. Kiessler, M. A. Wortman, "A critical look at the bathtub curve," *IEEE Transactions on Reliability*., vol. 52, no. 1, pp. 125-129, 2003.

[3] J. Luo, Y. Qiu and L. Yang, "Operation Fault Analysis of CLPE Power Cable above 10 kV," *High Voltage Eng*., Vol. 29, No. 6, pp. 14-16, 2003.

[4] M. Begovic, "Cable Asset Management Using Failure Forecasting and Statistical Assessment of Diagnostic Testing," http://www.wseas.us/conferences/2007/france/mmactee/Plenary4.htm.

[5] M. Wang, "State grid corporation of China's yearbook 2011," *China Electric Power Press*, 2012: pp. 566.

[6] H. Li, Z. Zhou and P. Chen, "Fault Analysis for 110 kV and the Above HV Cross-Linking Cable System," *Electrical Equipment*., no. 8, pp. 003, 2004.

[7] C. Walton and R. MacKinlay, "PD monitoring—A critical tool for condition-based assessments," *Transmission and Distribution World*, no. 55, pp. 38-46, 2003.

[8] C. Zhou, M. Michel, D. M. Hepburn, X. Song, "On-line partial discharge monitoring in medium voltage underground cables," *IET science, measurement & technology*, 2009, vol. 3, no. 5, pp. 354-363.

[9] R. M. Bucci, R. V. Rebbapragada, A. J. McElroy and E. A. Chebli, "Failure prediction of underground distribution feeder cables," *IEEE Transactions on Power Delivery*, vol. 9, no. 4, pp. 1943-1955, 1994.

[10] J. P. Ainscough P. EandI. W. Forrest P. E,"Predicting medium-voltage underground-distribution cable failures," in Proc. IEEE PES-ICC Fall Meeting, Scottsdale, AZ, USA, Nov. 11, 2009, pp.1–11, AZ—WG C26D.

[11] Y. Gill, "Development of an electrical cable replacement simulation model to aid with the management of aging underground electrical cables," *IEEE Dielectrics and Electrical insulation Systems*, vol. 27, no.1, pp. 31-37, 2011.

[12] Y. Gill, "Forecasting Underground Electric Cable Faults Using the Crow-AMSAA Model," *Reliability hotwire*, no. 115, pp. 1-6, 2010.

[13] P. E. Paul Barringer, "Predict failures: Crow-AMSAA 101 and Weibull 101," *International Mechanical Engineering Conference*, Kuwait, pp. 1-14, 2004.

[14] Y. Wang, Y. Jia, and W. Jiang, "Early failure analysis of machining centers: a case study," *Reliability Engineering & System Safety*, vol. 72, no. 1, pp. 91-97, 2001.

[15] T. R. Moss, "The Reliability Data Handbook," Professional Engineering Publishing Ltd, London and Bury St Edmund, UK, 2005

[16] W. Nelson, "Applied Life Data Analysis," Wiley Series inter-science, A John Wiley & Sons Inc. 1982.

[17] D. N. P. Murthy, M. Xie, R. Jiang, "Weibull models," John Wiley & Sons, 2004.

[18] W. Weibull, "A statistical theory of the strength of materials," *Ingeniors Vetenskaps Akademaiens Handingar*, no. 151, pp. 1-45, 1939.

[19] S. A. Hossain, R. C. Dahiya, "Estimating the parameters of a non-homogeneous Poisson-process model for software reliability," *IEEE Transactions on Reliability*, vol. 42, no. 4, pp. 604-612, 1993.

[20] Robert Abernethy, "The New Weibull Handbook, fifth edition,"

[21] L. H. Crow, A. P. Basu, "Reliability growth estimation with missing data. II," *Reliability and Maintainability Symposium*, *1988. Proceedings.*, pp. 248-253, 1988.

Zeyang Tang was born in Jingzhou, Hubei province, China, in 1987. He received the B.Sc. degree from the School of Electrical Engineering, Tianjin University, China in 2009. He is currently pursuing the Ph.D. degree in School of Electrical Engineering, Wuhan University, China. His research is mainly focused on cable failure data analysis and cable asset management.

Wenjun Zhou (SM'13) was born in China on July 1, 1959. He received the Ph.D. degree in Hydraulic and Electrical Engineering from Wuhan University, Wuhan, China, in 1990.

Currently, he is a Professor in the School of Electrical Engineering, Wuhan University, China. He is also the Vice Director of the Hubei High Voltage Committee. His research interests include lightning protection and the diagnostic techniques for outdoor electrical insulations.

Dr. Zhou is a member of the High Voltage Committee of the Chinese Society of Electrical Engineering (CSEE), the Vice Director of Electro-technical Test and Measurement Committee of the China Electro-technical Society (CES), and the China Lightning Protection Standard Committee.

Chengke Zhou (SM'13) received the B.Sc. and M.Sc. degrees in Hydropower Automation from Huazhong University of Science and Technology, Wuhan, China, in 1983 and 1986, respectively, and the Ph.D. degree in Electrical Science from The University of Manchester, Manchester, U.K., in 1994.

He joined the School of Engineering and Computing, Glasgow Caledonian University (GCU), in Glasgow, U.K. in 1994 and was a Post-doctoral Research Fellow, Lecturer, and Senior Lecturer until 2006 when he joined Heriot-Watt University as a Reader in Edinburgh, Scotland, UK. In 2007, he returned to GCU as a Professor. He has more than 20 years of research experience in power systems and partial-discharge-based high-voltage plant condition monitoring and has acted as a Consultant to EDF Energy, Scottish Power plc, and British Energy.. So far, he has published more than 100 papers.

Dr.Zhou is a Fellow of IET and a Chartered Engineer.

Yang Yang was born in Ezhou, Hubei province, China, in 1991. He received the B.Sc. degree from the School of Automation, Wuhan University of Technology, China in 2009. He is currently pursuing the Master degree in School of Electrical Engineering, Wuhan University, China. His research is mainly focused on high voltage cable online monitoring, cable insulation aging and dielectric loss trend measurements.

Appendix

Table A The data used in paper

SN	*t*	F/S	*JN*	*SN*	*t*	F/S	*JN*	*SN*	*t*	F/S	*JN*	*SN*	*t*	F/S	*JN*	*SN*	*t*	F/S	*JN*
1	0	S	1	62	563	S	5	123	1005	S	3	184	1533	S	2	245	3424	S	2
2	0	S	1	63	563	S	2	124	1005	S	3	185	1587	S	4	246	3424	S	4
3	5	F	1	64	579	S	5	125	1053	S	1	186	1587	S	2	247	3428	S	3
4	43	F	1	65	579	S	5	126	1056	S	1	187	1605	S	2	248	3436	S	5
5	65	F	1	66	582	S	2	127	1064	S	2	188	1605	S	4	249	3440	S	2
6	97	S	2	67	585	S	2	128	1068	S	2	189	1614	S	4	250	3440	S	2
7	125	S	2	68	585	S	2	129	1092	S	3	190	1614	S	4	251	3451	S	7
8	131	S	1	69	587	S	4	130	1125	S	2	191	1624	S	16	252	3487	S	2
9	144	S	2	70	587	S	2	131	1148	S	4	192	1646	S	2	253	3623	S	4
10	144	S	2	71	587	S	7	132	1157	S	3	193	1675	S	2	254	3623	S	4
11	154	S	9	72	587	S	2	133	1186	S	2	194	1675	S	2	255	3638	S	19
12	154	S	9	73	604	S	4	134	1186	S	2	195	1682	S	2	256	3777	S	16
13	157	S	7	74	620	F	1	135	1186	S	2	196	1683	S	2	257	3935	S	3
14	157	S	6	75	627	S	2	136	1186	S	2	197	1736	S	3	258	3935	S	3
15	168	S	7	76	627	S	2	137	1188	S	6	198	1800	S	1	259	4020	S	2
16	170	S	2	77	629	S	2	138	1216	S	2	199	1889	S	12	260	4020	S	2
17	170	S	1	78	629	S	2	139	1217	S	7	200	1939	S	2	261	4084	S	2
18	182	S	4	79	629	S	2	140	1219	S	5	201	1970	S	9	262	4084	S	2
19	182	S	4	80	629	S	2	141	1219	S	5	202	1982	S	9	263	4250	F	1
20	194	F	1	81	629	S	2	142	1231	S	10	203	2049	S	4	264	4250	F	1
21	238	S	2	82	629	S	2	143	1231	S	10	204	2100	F	1	265	4383	S	2
22	238	S	2	83	668	S	1	144	1231	S	2	205	2167	S	3	266	4444	S	10
23	248	S	3	84	668	S	1	145	1231	S	2	206	2175	S	7	267	4474	S	15
24	252	S	3	85	674	S	2	146	1231	S	2	207	2175	S	7	268	4474	S	1
25	252	S	3	86	674	S	2	147	1231	S	2	208	2281	S	2	269	4474	S	4
26	259	F	1	87	674	S	2	148	1231	S	2	209	2282	S	2	270	4474	S	8
27	261	S	2	88	701	S	3	149	1232	S	3	210	2313	S	2	271	4532	S	7
28	261	S	2	89	714	S	2	150	1232	S	3	211	2397	S	5	272	4777	S	7
29	262	S	3	90	714	S	2	151	1232	S	2	212	2549	S	19	273	4901	S	2
30	262	S	3	91	720	S	1	152	1232	S	2	213	2552	S	9	274	4931	S	1
31	262	S	3	92	722	S	4	153	1237	S	12	214	2629	F	1	275	4931	S	5
32	262	F	1	93	722	S	4	154	1246	S	5	215	2676	F	1	276	5174	S	5
33	310	S	15	94	730	S	9	155	1246	S	4	216	2676	F	1	277	5232	S	2
34	317	S	4	95	748	S	2	156	1246	S	4	217	2744	F	1	278	5266	S	2
35	317	S	6	96	748	S	2	157	1248	S	12	218	2815	S	2	279	5296	S	4
36	328	S	10	97	777	S	2	158	1259	S	2	219	2815	S	2	280	5296	S	4
37	328	S	4	98	777	S	9	159	1278	S	2	220	2817	S	2	281	5357	S	2
38	328	S	1	99	797	S	7	160	1278	S	2	221	2817	S	2	282	5508	S	3
39	354	F	1	100	819	S	2	161	1312	S	8	222	2907	S	2	283	5508	S	3
40	442	S	1	101	878	S	11	162	1315	S	11	223	2996	S	4	284	5661	S	4
41	523	S	1	102	883	S	1	163	1328	S	13	224	2996	S	4	285	5661	S	8
42	523	S	1	103	890	S	2	164	1329	S	4	225	2996	S	4	286	5661	S	6
43	528	S	2	104	890	S	2	165	1347	S	5	226	3087	S	2	287	5661	S	8
44	528	S	7	105	901	S	2	166	1351	S	2	227	3087	S	2	288	5661	S	4
45	528	S	3	106	901	S	2	167	1351	S	1	228	3087	S	2	289	6012	S	4
46	528	S	2	107	903	S	2	168	1351	S	1	229	3087	S	2	290	6012	S	2
47	528	S	7	108	903	S	2	169	1396	S	1	230	3087	S	2	291	6012	S	3
48	528	S	3	109	904	S	2	170	1430	S	1	231	3100	S	7	292	6012	S	2
49	541	S	4	110	910	S	2	171	1455	S	7	232	3105	S	16	293	6020	S	3
50	546	S	2	111	914	S	2	172	1464	S	2	233	3107	S	8	294	6027	S	2
51	546	S	2	112	918	S	6	173	1464	S	2	234	3119	S	3	295	6027	S	2
52	547	S	1	113	918	S	6	174	1464	S	2	235	3121	S	2	296	6086	S	16
53	547	S	1	114	921	S	2	175	1464	S	2	236	3121	S	2	297	6119	S	2
54	550	S	2	115	922	S	2	176	1491	S	3	237	3121	S	2	298	6119	S	2
55	550	S	2	116	922	S	9	177	1500	S	8	238	3289	S	5	299	8369	S	7
56	550	S	2	117	923	S	2	178	1500	S	6	239	3290	S	3	300	8369	S	5
57	550	S	2	118	964	S	4	179	1507	S	2	240	3352	S	2	301	8369	S	10
58	550	S	2	119	968	F	1	180	1507	S	2	241	3353	S	3	302	8369	S	12
59	555	S	2	120	981	S	12	181	1508	S	5	242	3353	S	2				
60	555	S	2	121	990	S	2	182	1514	S	7	243	3424	S	5				
61	562	S	4	122	991	S	2	183	1522	S	7	244	3424	S	2				

SN denotes the serial number. *t* denotes age to failure or censored time. F/S denotes the status of cable. F means failure, while S means suspension. *JN* denotes the number of joints.