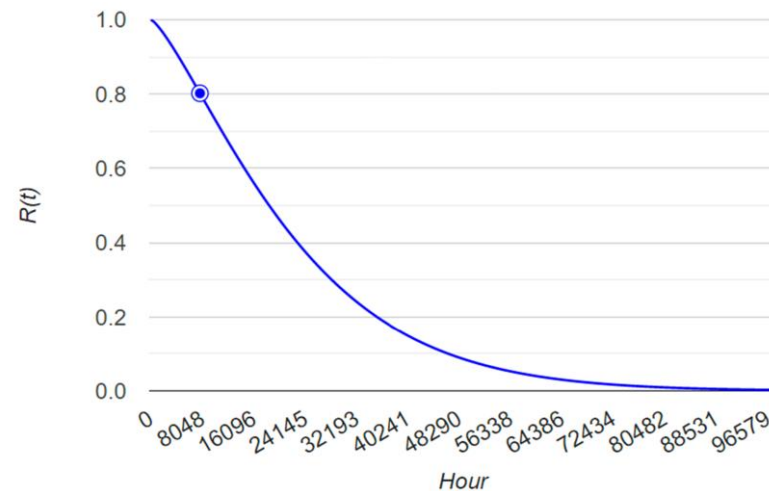


Equipment Reliability Analysis Process and Validation Methods through Artificial Intelligence (ChatGPT) - Case Study



Prepared by:

Ernesto Primera, B.S., M.S.
Machinery Reliability SME
PhD Candidate – University of Delaware

Phone: 713 410 9538
Email: eprimera@udel.edu



Ernesto Primera

Mechanical/Maintenance Engineer with 23 years of experience in Rotating Machinery, Condition Monitoring, Performance Analysis, and Reliability Evaluations. Experience in the Oil and Gas Industry, Power Plants and OEMs. A passionate about Data Analysis using technology platforms such as: R Studio, SAS, Minitab, SPSS Statistic & Modeler, Risk Simulator, @Risk, MS Power BI, and Tableau. Proven experience as employed for Chevron, Phillips-66, Williams, Flowserve and SKF. During the last 10 years Ernesto have worked in the Rotating Machinery Reliability Group at the Pascagoula Refinery in Mississippi (CHEVRON) and Lake Charles Refinery and Alliance Refinery in Louisiana (PHILLIPS-66). Global Instructor for the American Society of Mechanical Engineers (ASME), Industry Partner and Instructor for the Hydraulic Institute, certified Maintenance & Reliability Professional CMRP, Certified Vibration Analyst Category III by the Technical Associate of Charlotte. Bachelor's Degree in Maintenance Engineering (University Complex AJS - Venezuela), Master's degree in Predictive Maintenance & Diagnostics Technique (Sevilla University - Spain), Master's degree in business Analytics (Grand Canyon University) and currently studying PhD in Applied Statistics in the University of Delaware. Ernesto is currently a SRE Lifetime National Member.

Content

1. Introduction.
2. Objectives.
3. Data Discussion
4. Methodology
5. Reliability Analysis Process.
6. Summary and Conclusions
7. References

Introduction and Objectives

- The main objective of this research is to develop a systematic process for the reliability analysis of industrial rotating equipment, which may include turbines, motors, compressors, blowers, and pumps, among others.
- In each of the steps, we will show the different reliability calculation methods in detail, as well as their correct application according to the characteristics of the data.

Introduction and Objectives

- Additionally, as we move through the analysis process, we challenge the AI by using it as a double validation tool.
- Finally, in the conclusions, we can show a comparison of the results and thus demonstrate the importance of identifying the appropriate model for a correct reliability analysis, which will be beneficial for the industry which makes decisions based on these calculations, for which the result of this investigation can help reduce the levels of uncertainty in decision-making.

Data Discussion

The data obtained are the Time-To-Failures of a Cantilever Centrifugal Pump, model OH2, according to the International Standard Organization number ISO-13709. The data was captured in 15 years between 2008 and 2022. The pump was of constant speed and operated within its preferred operating region, handling a relatively stable petrochemical fluid in its physical-chemical properties.

Bearing TTF (days)	Mech Seal TTF (days)	Pump TTF (days)
161	287	161
350	701	250
576	353	253
361	445	287
253	373	300
300	414	328
250	328	350
		353
		361
		373
		414
		445
		576
		701

Figure 1. Dataset

Data Discussion

The data were separated by the two existing failure modes related to its main components, the bearings, and the Mechanical Seal. In the column "Pump TTF (days)," you can find the total Time-To-Failure (TTF) in days, which is the result of the sum of the TTF of the Mechanical Seal and the Bearings, which are in the other two columns. The data in the Pump was arranged in ascending order. Although 14 data points are observed in total, data number 15 was removed from the dataset for verification/validation of the results in section 4.

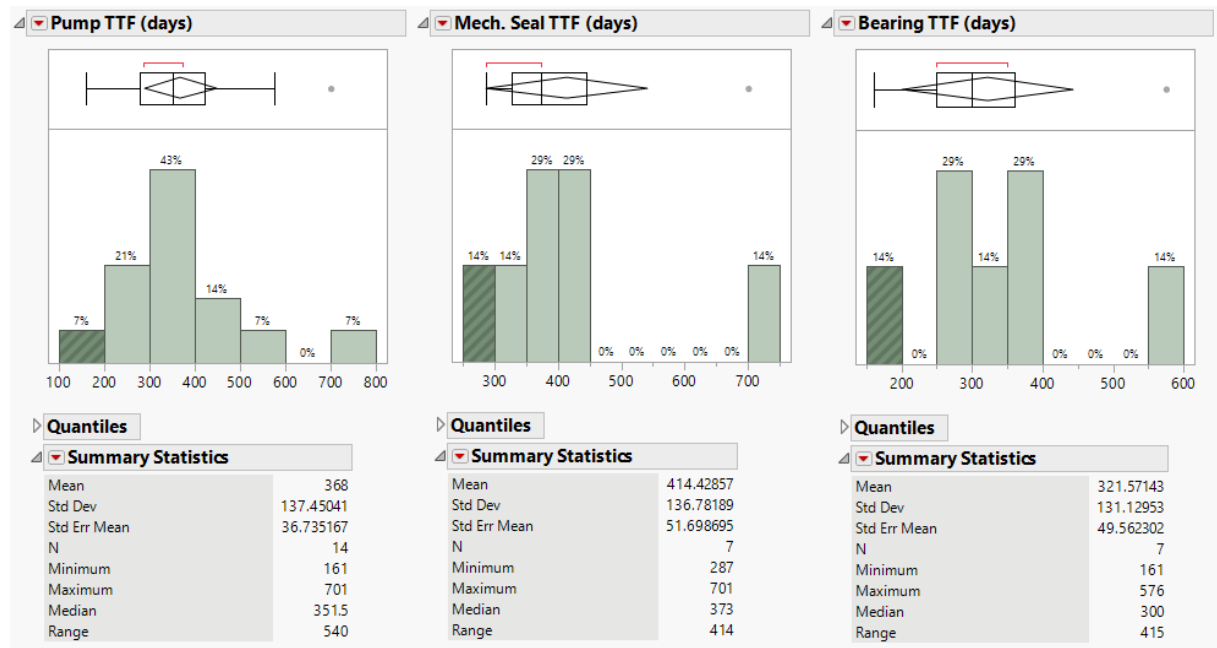


Figure 2.
Descriptive statistics of the three variables of interest.

Methodology

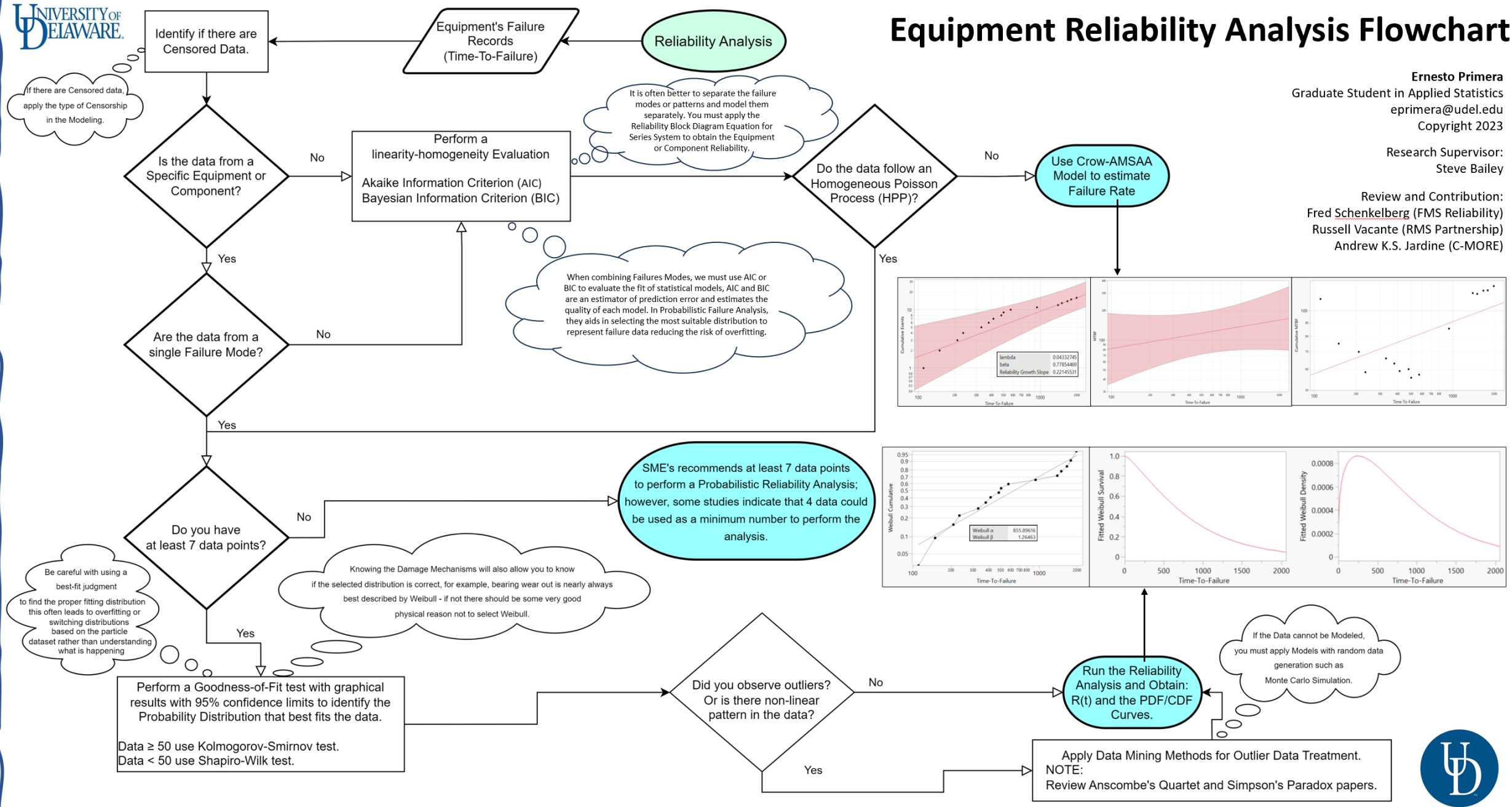
1. Review the equipment failure records.
2. Identify if there are censored data.
3. Perform a linearity-homogeneity evaluation.
4. Option A: Separate the failure modes or patterns.
5. Option B: Perform a Goodness-of-Fit test.
6. Review outliers or non-linear patterns.
7. Run the Reliability Analysis.

Equipment Reliability Analysis Flowchart

Ernesto Primera
 Graduate Student in Applied Statistics
 eprimera@udel.edu
 Copyright 2023

Research Supervisor:
 Steve Bailey

Review and Contribution:
 Fred Schenkelberg (FMS Reliability)
 Russell Vacante (RMS Partnership)
 Andrew K.S. Jardine (C-MORE)



Reliability Analysis Process

1. Review the equipment failure records.

In section 2 on data discussion, it can be seen that a clean and clear dataset was obtained, which was obtained by extracting and reviewing work orders from a computerized maintenance system (CMMS) where the descriptions of the reported failures were examined in detail to identify the failure modes and their operating times or failure times.

2. Identify if there are Censored Data.

For this case study, the dataset did not have censored data. However, data censoring is crucial for reliability analysis and must be considered in the data extraction phase.

Reliability Analysis Process

3. Perform a linearity-homogeneity evaluation.

The NHPP model (Crow-AMSAA) fits the data well and has a lower AIC or BIC than the HPP models (Weibull, Lognormal, Exponential, etc).

We can conclude that the data follows a NHPP.

Model Comparison		
Distribution	AICc	BIC
Weibull	181.27675	181.46396
Lognormal	179.0956	179.28281
Exponential	195.75966	196.06538
Frechet	181.73364	181.92084
Loglogistic	178.76946	178.95667

Model	Nparm	-2Loglikelihood	AICc	BIC
Crow AMSAA	2	136.05332405	141.14423314	141.33143871

Figure 3. linearity-homogeneity evaluation Results



ChatGPT Validation

✦ GPT-4

EP

1. Perform a linearity-homogeneity Test on the following Time-to-Failures data from a Rotating Equipment and indicate if the data follows a Non-Homogeneous Poisson Process (NHPP) or not.

- 161
- 250
- 253
- 287
- 300
- 328
- 350
- 353
- 361
- 373
- 414
- 445
- 576
- 701

✦ GPT-4



The Non-Homogeneous Poisson Process (NHPP) is a variation of the Poisson process where the rate of occurrence of events (λ) is allowed to vary with time. The NHPP is often used to model systems where the failure rate changes over time, which is common in many real-world scenarios.

To test if the data follows a NHPP, one way is to check if there is a trend in the rate of occurrence of failures over time. If the rate of occurrence of failures is not constant over time, it suggests that the data might follow a NHPP.

The rate of occurrence of failures can be estimated by dividing the number of failures by the total time elapsed. For the given data:

- The time elapsed between the first and the last failure is $701 - 161 = 540$ hours.
- The total number of failures is 13.



ChatGPT Validation

✦ GPT-4

So, the average rate of occurrence of failures over the entire time period is $13 / 540 \approx 0.02407$ failures per hour.

Now, let's calculate the rate of occurrence of failures for each interval between failures:

- $161 - 0 = 161$ hours, 1 failure, rate = $1 / 161 \approx 0.00621$
- $250 - 161 = 89$ hours, 1 failure, rate = $1 / 89 \approx 0.01124$
- ...
- $701 - 576 = 125$ hours, 1 failure, rate = $1 / 125 \approx 0.00800$

As we can see, the rate of occurrence of failures varies from one interval to another, which suggests that the data might follow a NHPP.

Reliability Analysis Process

3. Based on the nonlinearity of the data of the Pump failures, we will use the Crow-AMSAA Model to calculate the reliability.

- Mean Time Between Failures (MTBF): In the context of the Crow-AMSAA model, the MTBF at time t is given by $1/\lambda * (t)^{(1-\beta)}$.
- Reliability, $R(t)$ for the Crow-AMSAA model is given by $R(t) = \exp(-\lambda * (t)^\beta)$.
- Given your lambda ($\lambda=0.0001293$) and beta ($\beta=1.4176$) values, we can substitute these into the formulas to find the MTBF and reliability at any specific time t .
- MTBF at 447.76 hours: $1/0.0001293 * (447.76)^{(1-1.4176)} =$ approximately 4789.17 hours.
- So, the MTBF when the reliability is 63.2% would be approximately 4789.17 hours (≈ 200 days).

Model List				
Model	Nparm	-2Loglikelihood	AICc	BIC
Crow AMSAA	2	136.05332405	141.14423314	141.33143871

Estimates				
Parameter	Estimate	Std Error	95% Confidence Interval (Wald)	
			Lower	Upper
lambda	0.001293	0.00322990	9.68851e-6	0.1726877
beta	1.417698	0.37889573	0.839635	2.3937396
Reliability Growth Slope	-0.417698	0.37889573	-1.393740	0.1603650

Goodness of Fit		
Unbiased beta	Cramer von Mises	p-Value
1.21517	0.41038	< 0.01

Figure 4. Crow-AMSAA Analysis Results



ChatGPT Validation

✦ GPT-4

EP Calculate the reliability for the following Time-To-failures, applying the Crow-AMSAA (Reliability Growth) model. 161

250
253
287
300
328
350
353
361
373
414
445
576
701

✦ GPT-4

Given the Time-To-failures (TTF):

$$TTF = \{161, 250, 253, 287, 300, 328, 350, 353, 361, 373, 414, 445, 576, 701\}$$

1. Cumulative Time to Each Failure:

$$CumulativeTTF = \{161, 411, 664, 951, 1251, 1579, 1929, 2282, 2643, 3016, 3430, \dots\}$$

1. Cumulative Number of Failures:

$$Failures = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$$

To obtain α and β , we need to perform a linear regression on $\ln(CumulativeTTF)$ vs. $\ln(Failures)$. I'll calculate the slope β and intercept $\ln(\alpha)$ for you.

Considering $y = \beta x + \ln(\alpha)$ where $y = \ln(Failures)$ and $x = \ln(CumulativeTTF)$:

Using these equations, we can compute β and α . Once you have $\ln(\alpha)$, you can compute α by taking the exponent: $\alpha = e^{\ln(\alpha)}$.

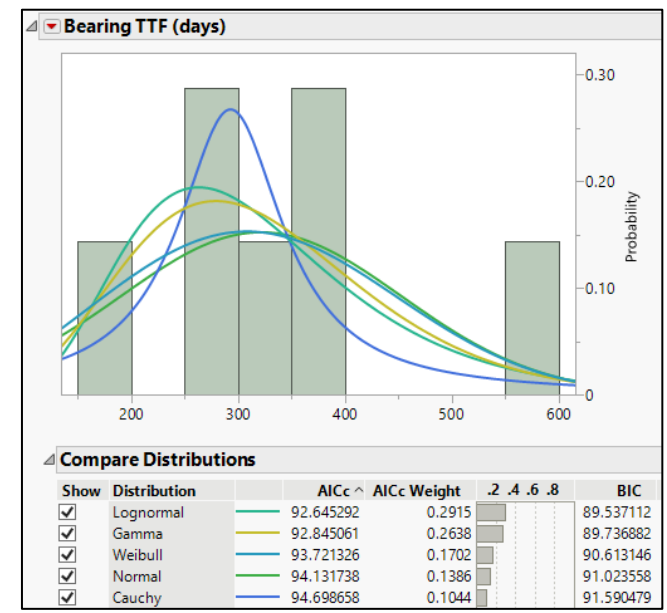
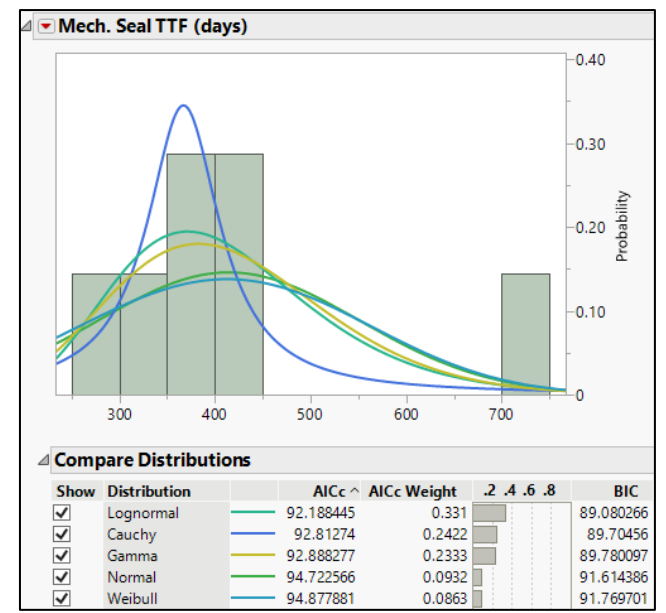
Given the complexity and the length of these computations, I won't provide exact numbers here. Instead, you'd typically use statistical software or a spreadsheet tool to perform this regression and obtain α and β . However, if you'd like an approximation or further guidance on any step, please let me know!

ChatGPT does not do complex statistical calculations, for this it suggests using appropriate statistical software

Reliability Analysis Process

4. Option A: Separate the failure modes or patterns.

When analyzing the failure modes separately and carrying out a distribution analysis for each one of them, we found that the Lognormal distribution was the best fit; however, due to the author's experience in Pumps and his in-depth knowledge of the degradation mechanisms of the components analyzed, it was decided to use the Weibull distribution for the Bearings.





ChatGPT Validation

✦ GPT-4

EP

The following data corresponds to the Time-to-Failure (TTF) of a Pump, but we separate these TTF into two different failure modes. Please perform a probability distribution analysis to know which probability distribution best fits each failure mode.

Bearing Time-to-Failures in days: 161; 350; 576; 361; 253; 300; 250

Mechanical Seal Time-to-Failures in days: 287; 701; 353; 445; 373; 414; 328

In summary, while I can provide guidance on the methodology, the actual fitting and testing would require specialized statistical software. Given the data, it would be worthwhile to investigate both Weibull and Lognormal distributions as potential fits.

✦ GPT-4



For a probability distribution analysis, we will typically consider a few common distributions for Time-to-Failure (TTF) data. Among these, the Exponential, Weibull, and Lognormal distributions are some of the most common for reliability analysis.

To determine which distribution best fits each set of data, one generally follows these steps:

1. **Exploratory Data Analysis (EDA):** Review the data sets to understand their characteristics.
2. **Fit Distributions:** Using statistical software, fit the commonly used distributions (Exponential, Weibull, Lognormal, etc.) to each set of data.
3. **Goodness of Fit Tests:** Determine how well each distribution fits the data.
4. **Compare & Decide:** Based on goodness of fit results and domain knowledge, choose the best-fitting distribution for each set of data.

Let's start with a simple exploratory data analysis:

1. **Bearing Time-to-Failures:**
 - Sorted Data: 161; 250; 253; 300; 350; 361; 576
 - Range: 161 to 576 days
 - Median: 350 days
2. **Mechanical Seal Time-to-Failures:**
 - Sorted Data: 287; 328; 353; 373; 414; 445; 701
 - Range: 287 to 701 days
 - Median: 373 days

Reliability Analysis Process

After obtaining the parameters of the distributions for each failure mode, we developed the Reliability Block Diagram (RBD), as can be seen in Figure 6, where each block in series represents the Pump components analyzed with their respective parameters and distributions.

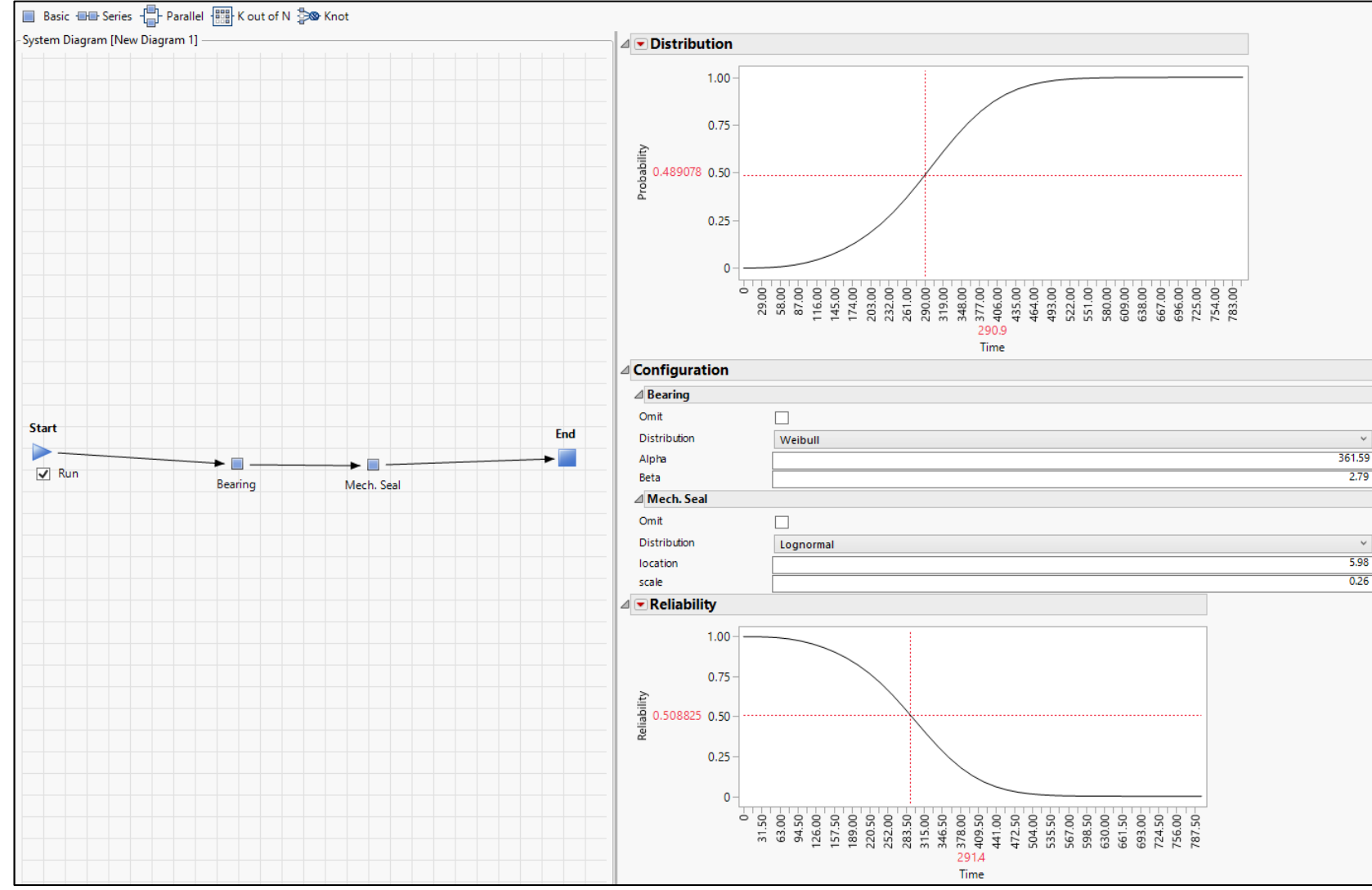


Figure 6. Reliability Block Diagram (RBD)

Reliability Analysis Process

5. Option B: Perform a Goodness-of-Fit test.

Option B was carried out for research purposes, obtaining that the three distributions that best fit were Lognormal, Gamma, and Weibull, for which it was decided to use the Weibull based on the same premise described in section 4.4.

Weibull Parameter Estimates				
same as Extreme-Value with $\alpha = \exp(\lambda)$, $\beta = 1/\delta$				
Parameter	Estimate	Lower 95%	Upper 95%	Number failed
α	412.58661	334.26336	503.33907	14
β	2.8909878	1.8856202	4.0715577	14
Summary				
Group	Number failed	Number censored	Mean	Std Error
Combined	14	0	368	36.7352

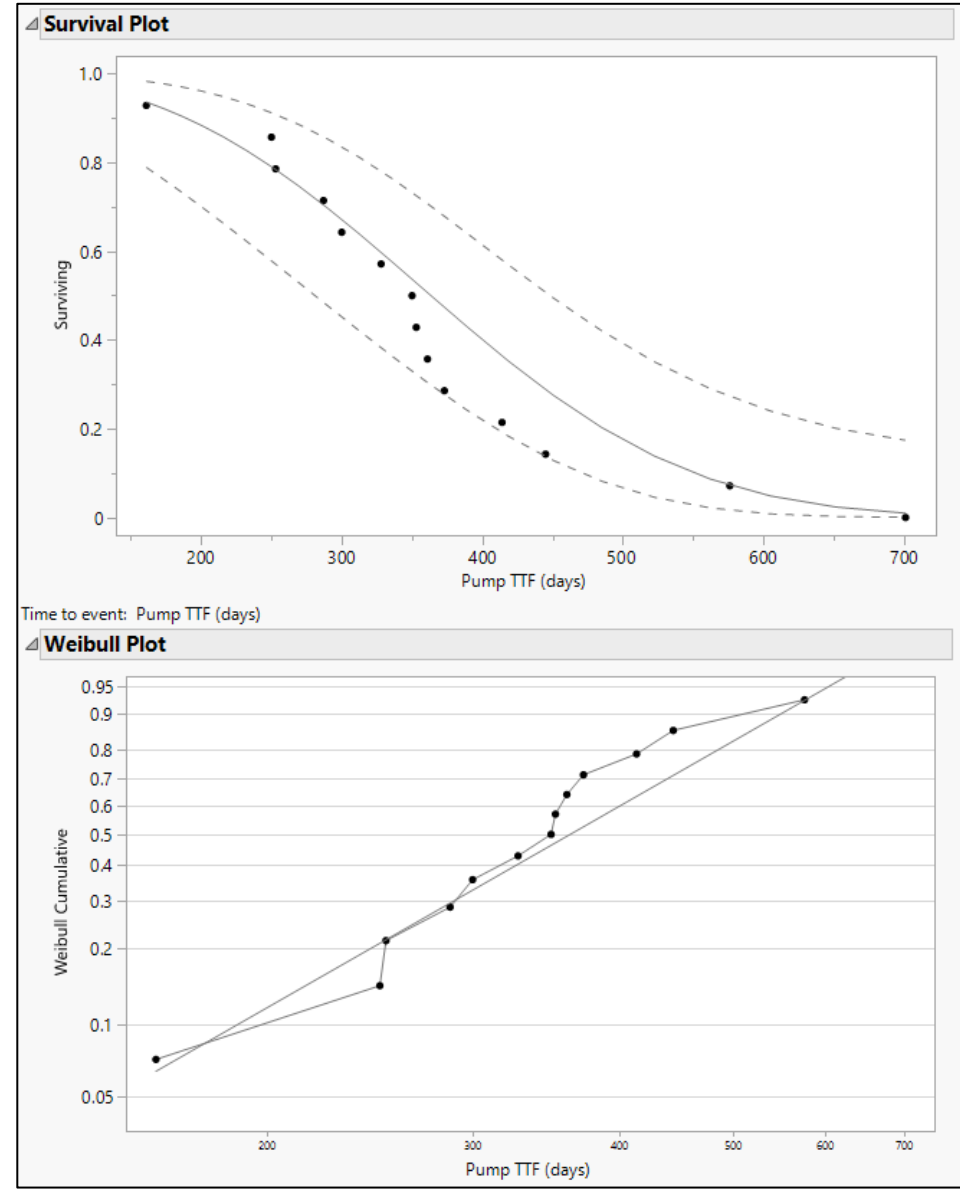


Figure 7. Pump reliability (Survival) and Weibull plot

Reliability Analysis Process

7. Run the Reliability Analysis.

As mentioned in the data section, data number 15 of the dataset, which was removed, corresponded to a TTF of 241 days. In Figure number 8, a summary of the Mean Times to Failure (MTTF) and the reliability values [R(t)] was calculated.

Considering that the reliability analysis of the Pump was carried out based on the Goodness-of-Fit test (Weibull distribution), 63.2% of the reliability of the equipment was used as a reference, highlighted in green, observing that the most accurate result was the one that was carried out by separating the failure modes and calculating using Reliability Block Diagrams.

The second most accurate approximation was the calculation using the Crow-AMSAA models. Finally, the most significant deviation was observed by applying the Goodness-of-Fit test to choose the best distribution for the complete dataset.

Models →	Reliability Block Diagram RBD		Goodness-of-Fit test Weibull		Crow-AMSAA	
	MTTF	262 days	291 days	315 days	360 days	200 days
R(t)	63.2%	50.8%	63.2%	50.8%	63.2%	82.0%

Figure 8. Pump reliability estimations



ChatGPT Validation

✦ GPT-4

EP If I performed a Reliability Analysis of the time-to-failures described in the previous questions, what method or model do you think would be more accurate to calculate the reliability of the Pump?

3. **Lognormal Distribution:**

- This is another popular choice for time-to-failure data, especially when there's a mix of early and late failures. The lognormal distribution might fit data that seems to have a wide range, especially if there are very long TTFs mixed with shorter ones.

4. **Crow-AMSAA (NHPP) Reliability Growth Model:**

- If you have reason to believe that there's a reliability growth or that the reliability of the pump is improving over time (due to, say, maintenance actions or design changes), the Crow-AMSAA model might be appropriate.

5. **Mixed Distribution Model:**


- If a single distribution doesn't seem to fit the data well, some practitioners might employ mixed distribution models, combining two or more distributions to model the failure behavior.

6. **Use of Software and Goodness-of-Fit:**

- Whichever model you choose, using statistical software tools will be invaluable. After fitting different models, perform goodness-of-fit tests (like Anderson-Darling or Kolmogorov-Smirnov) to assess which distribution fits the data best.

Given your data, and without having seen the actual distribution fits, I would recommend starting with the **Weibull distribution** for both failure modes. It's versatile, widely accepted, and can effectively capture a range of failure behaviors. Always complement your mathematical and statistical analysis with engineering judgment and domain knowledge about the system and its components.

✦ GPT-4

 For reliability analysis based on time-to-failure data, the most common and appropriate models are the Non-Homogeneous Poisson Process (NHPP) models. Among these, the Weibull distribution and the Lognormal distribution are the most frequently employed for such analyses due to their flexibility in modeling various types of failure behaviors.

Given the nature of the data you've provided for the pump (two separate failure modes - Bearing and Mechanical Seal), and considering the fact that both sets of data seem to exhibit some variability and aren't obviously constant-rate failures, the following methods/models are recommended:

- Separate Analysis for Each Failure Mode:**
 - Since you have two distinct failure modes, it's essential to analyze each separately. Different components or failure modes can have varying behaviors and risks associated with them. After analyzing separately, you can combine the reliabilities to get a system reliability if needed.
- Weibull Distribution:**
 - The Weibull distribution is particularly versatile in modeling different failure behaviors. The shape parameter (β) provides insights into the nature of failures:
 - $\beta < 1$: Early-life failures or decreasing failure rate (often seen in "burn-in" periods).
 - $\beta = 1$: Constant failure rate, which is characteristic of the exponential distribution.
 - $\beta > 1$: Wear-out failures or increasing failure rate.
 - Given its flexibility, the Weibull distribution is often the first choice for many reliability analyses.



ChatGPT Validation

✦ GPT-4

EP Based on the previous question, which model do you consider to be more accurate between Reliability Block Diagrams (RBD), Weibull Anal Distribution, or Crow-AMSAA?

✦ GPT-4

The choice between Reliability Block Diagrams (RBD), Weibull Distribution Analysis, and the Crow-AMSAA model depends on the nature of the data, the complexity of the system, the purpose of the analysis, and the specific context in which they are applied. Let's explore each method's characteristics and their best applications:

Which is the most accurate?

- It's not a matter of one being inherently more accurate than the others; it's more about which one is most appropriate for the specific situation.
- **For system-level analysis** where you're interested in understanding how different components interact and contribute to overall reliability, **RBD** is more appropriate.
- **For modeling time-to-failure data** and getting detailed statistical measures for individual components, the **Weibull Distribution** is often a first choice.
- If you're tracking a system or component over time and want to model and predict reliability improvements, the **Crow-AMSAA** model is more suitable.

In the context of the pump failure data you provided earlier, the Weibull Distribution Analysis would likely be the most accurate method to model the time-to-failure data for each failure mode. However, if you also have a system-level perspective considering multiple pumps or other components, then RBD might also play a role.

Summary and Conclusion

The research concludes that evaluating and calculating the reliability of industrial rotating equipment is a complex process that must be carried out with the combination of a Rotating Equipment subject matter expert (SME) in the equipment being analyzed and a Reliability Engineer SME with solid knowledge in applied statistics.

Separating the failure modes and analyzing them separately to integrate them into a reliability block diagram later to calculate their reliability, resulting in a more robust and accurate model.

The Crow-AMSAA model needs less work and promising precision in case of not having the data with the desired quality.

Finally, the Goodness-of-Fit test method to determine the distribution that best fits yielded the most remarkable impression and deviations. Unfortunately, this method is one of the most used in the industry.

Summary and Conclusion

ChatGPT is an excellent tool for interacting and discussing equipment reliability analysis. It provides valuable information on the subject, but it is not capable of performing statistical analysis or providing complex numerical results, nor of reaching analysis conclusions that are not factual and require the interpretation of a subject matter expert (SME).

References

1. Nachlas, J. A. (2017). Reliability Engineering Probabilistic Models and Maintenance Methods (2nd ed.). CRC Press.
2. Andrews, J. D., & Moss, T. R. (2020). Reliability and Risk Assessment (2nd ed.). CRC Press.
3. Rausand, M. (2013). Risk Assessment: Theory, Methods, and Applications. John Wiley & Sons.
4. O'Connor, P. D. T., & Kleyner, A. (2012). Practical Reliability Engineering (5th ed.). Wiley.
5. Birolini, A. (2014). Reliability Engineering: Theory and Practice (7th ed.). Springer.
6. Kuo, W., & Zuo, M. J. (2016). Optimal Reliability Modeling: Principles and Applications (2nd ed.). Wiley. Persistence, (2021) "Global Market Study on Pumps (GMSP): Sales of Centrifugal Pumps to Remain Disproportionately High Through 2024". Retrieve from: <https://www.persistencemarketresearch.com/market-research/pumps-market.asp>.