1. Report No. RailTEAM UD-2	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle On the Development of a Predictive Rail Maintenance Planning Methodology Utilizing Parametric Weibull Prediction Methods		5. Report Date September 1, 2020 6. Performing Organization Code:
7. Author(s) John J. Cronin and Allan Zarembski https://orcid.org/0000-0002-4282-9330		8. Performing Organization Report No. UD-2
9. Performing Organization Name and Address Department of Civil & Environmental Engineering University of Delaware 301 DuPont Hall Newark, DE 19716		10. Work Unit No. 11. Contract or Grant No. 69A3551747132
12. Sponsoring Agency Name and Address Office of Research, Development and Technology (RD&T) US Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The railroad industry has used for the past 50 years the 2-Parameter Weibull equation to determine the rate of rail fatigue defect occurrences and to forecast the fatigue life of railroad rail. With the advent of more powerful computers, more frequent data collection and new techniques to analyze data, a new field of data analysis has been created, Data Analytics, sometimes referred to as "Big Data". This report makes use of this new area of Data Analytics to develop and implement an improved rail defect forecasting approach building upon the traditional Weibull equation to overcome many of its limitations and problems.

This report presents a novel Data Analytic method developed from Parametric Bootstrapping. This approach is designed to provide for an application of the Parametric Bootstrapping modified Weibull forecasting to rail segments with insufficient numbers of defects to allow for the traditional Weibull forecasting analysis. Thus, the Bootstrapping method provides reasonable estimates of the rate of defects for track segments that have little or no prior defect data, which allows far more track to be analyzed, and to be accounted for in maintenance planning efforts. In addition, there is a range of values used in the prediction, allowing for an estimate of "best case" and "worst case" scenario. This approach results in an ability to forecast the probability of rail defect occurrence as a function of cumulative tonnage experienced by the rail as well as other key track and traffic parameters that affect the development of fatigue defects.

The results presented here show that the Parametric Bootstrapping Weibull Analysis approach offers a more accurate and effective approach to determining the probability of developing future defects with an overall benefit to the railroads in their maintenance of an expensive rail asset.

17. Key Words	18. Distribution Statement		
Parameter Welbull equation, parametric bootstrapping, probability of rail defect occurrence	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 112	22. Price

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

USDOT Tier 1
University Transportation Center
on Improving Rail Transportation
Infrastructure Sustainability and Durability

Final Report UD-2

ON THE DEVELOPMENT OF A PREDICTIVE RAIL MAINTENANCE PLANNING METHODOLOGY UTILIZING PARAMETRIC WEIBULL PREDICTION METHODS

John J. Cronin, PhD
Graduate Research Assistant
Department of Civil Engineering
University of Delaware
jjcronin@udel.edu

and

Allan Zarembski PhD, PE, FASME, Hon. Mbr. AREMA
Professor and Director of Railroad Engineering and Safety Program
Department of Civil Engineering
University of Delaware
dramz@udel.edu

September 1, 2020

Grant Number: 69A3551747132

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The railroad industry has used for the past 50 years the 2-Parameter Weibull equation to determine the rate of rail fatigue defect occurrences and to forecast the fatigue life of railroad rail. With the advent of more powerful computers, more frequent data collection and new techniques to analyze data, a new field of data analysis has been created, Data Analytics, sometimes referred to as "Big Data". This report makes use of this new area of Data Analytics to develop and implement an improved rail defect forecasting approach building upon the traditional Weibull equation to overcome many of its limitations and problems.

Because of the serious nature of broken rail defects and the approximately 100 broken rail derailments that occur in the US each year, railroads continue to improve rail inspection techniques, rail maintenance techniques and its rail defect data collection process. The railways industry currently collects data on the occurrence of defects, rail inspection results, rail maintenance techniques such as rail grinding and a broad range of rail statistics, which have the potential to provide increased insight into the rate of occurrence of rail defects. The Weibull Equation, while giving a basic forecast capability, does not explicitly account for many of the key operating and maintenance variables that affect the development of rail defects. As such, using traditional Weibull analysis techniques, it is not possible to predict what the effects of differing maintenance operating or material parameters would be on the rate of defects development. Thus, while the current use of the 2-Parameter Weibull equation is adequate for its current limited use in rail life forecasting, it appears to be possible to improve on the rail life forecasting and prediction of defects through the use of new Data Analytic techniques which make more aggressive use of the extensive rail defect data available. These improvements can lead to a more accurate prediction method with practical implications in rail life forecasting, maintenance management and replacement planning.

This report a novel Data Analytic method developed from Parametric Bootstrapping. This approach is designed to provide for an application of the Parametric Bootstrapping modified Weibull forecasting to rail segments with insufficient numbers of defects to allow for the traditional Weibull forecasting analysis. Thus, the Bootstrapping method provides reasonable estimates of the rate of defects for track segments that have little or no prior defect data, which allows far more track to be analyzed, and to be accounted for in maintenance planning efforts. In addition, there is a range of values used in the prediction, allowing for an estimate of "best case" and "worst case" scenario. This approach results in an ability to forecast the probability of rail defect occurrence as a function of cumulative tonnage experienced by the rail as well as other key track and traffic parameters that affect the development of fatigue defects.

The results presented here show that the Parametric Bootstrapping Weibull Analysis approach offers a more accurate and effective approach to determining the probability of developing future defects with an overall benefit to the railroads in their maintenance of an expensive rail asset.

CHAPTER 1 INTRODUCTION

Rail Transportation in the United States of America plays a vital role in the transportation of goods, as well as passengers, across the vast distances that make up the nation. Railroads have had an overall average share of 31.8% of all freight ton-miles from 2007 to 2016; averaging over 1,712,000 ton-miles[1] per year (1). In specific corridors, such as the New York to Washington metropolitan area, passenger rail transportation can rival airlines, and in some areas has grown to over 11 million passengers per year (2). Helping to drive this increase in usage is the increased loading on the rails themselves; for example, freight axle loads have increased over the past decades, from pre-1970 limits of 27 tons, to current limits of 36 tons. Likewise, higher-speed passenger lines generate higher dynamic loads, resulting in more stress on the rails than previously seen. One of the major factors in handling this increased stress is the maintenance work done to ensure that the rails themselves are in a state of good repair, which demands both prompt alerts to unsafe conditions, and the ability to do the work without undue delay on revenue operations. This is handled two ways; through the constant monitoring of the railway network's rails using various non-destructive testing, such as ultrasonic testing and remote sensing technologies, and through the use of preventative maintenance planning, which allows maintenance work to be done during expected downtimes prior to when issues arise, instead of when the issues arise during revenue operations. Thus, for example, avoiding broken rails in service, which have the potential for highly undesirable broken rail derailments.

In order to improve maintenance activities, railroads have been developing analytical techniques, which forecast the rate of rail defect development and the potential for future failure, e.g. broken rail. Since the late 1970's, when research was published showing that rail defect occurrences follows a Weibull Probability Distribution, railroads have used the Weibull equations to determine the current and future useful lifetime of rail (3). This directly influences the decision on when and where to focus maintenance efforts, and directly impacts when and where defects are more likely to turn into accidents, track failures, and other incidents.

1.1 Statement of Problem

The problem that will be examined and addressed is the development of an improved, more accurate, and more encompassing methodology for the prediction of rail fatigue defect growth and occurrence with respect to the currently collected rail fatigue defect data. While railroads have collected rail defect data for over a hundred years, it is only within the last few decades that the use of this defect data to predict rail life has been used (3, 6, 7, 8) and that with limited accuracy. The focus of this defect forecasting and prediction was on the use of the Weibull approach. However, this methodology only allows analysis of track segments which have had sufficient prior defects identified to allow for this traditional Weibull analysis. Analysis of large-scale defect data bases from a major Class 1 railroad, shows that there are many track segments with few or no defects, which does not allow for the traditional Weibull analysis approach on these segments. By developing a methodology which can account for the lack of defect data in these track segments,

[1] A ton-mile represents one ton of goods moved by rail one mile.

it becomes possible to develop a more extensive forecasting model which in turn allows for more effective scheduling of maintenance efforts based upon the forecast life of the rail.

1.2 Objective of Research

The objective of this research is to develop an advanced fatigue defect prediction methodology which mitigates the issues of the source data. The second objective is to make sure that this methodology is expandable and accessible to railway maintenance planners, to ease the introduction of, and acceptance of, this new methodology. This will then allow the maintenance planners to create predictions for their track in question, allowing them to draw conclusions from the data as to favorable times for maintenance or rail replacement.

The objectives for developing this methodology include:

- Clean and Combine all the source data into one easy-to-use database.
- Develop a baseline Weibull analysis from which to compare new methodologies to.
- Investigate Machine Learning techniques, as they have had success in producing results from limited or obfuscated datasets.
- Approach the problem from a Weibull point of view; expand upon the function in some way to improve results
- Condense what has been learned from both approaches into a single methodology that is easily used.
- Apply the methodology to a sample of track, and detail the process for repeatability.

1.3 Research Approach

This research was conducted in three stages. The first stage dealt with finding out which methods would or would not work with the data to produce a viable result. Initially hampered by the failure to find a baseline from which other analyses could be compared to, this stage expanded into machine learning and improvements on the Weibull methodology. Based upon the results from this first stage, the second stage, a development of a methodology was accomplished by combining what was learned from the machine learning and failures of the Weibull analyses, into a new methodology that applied parametric bootstrapping to the Weibull analysis. This suddenly allowed all the track in the data to have relatively quickly computed Weibull distributions developed, removing issues where few track segments had enough data to develop reasonable Weibull outputs. The third stage of the research focused on applying this parametric bootstrapped Weibull method into a full methodology, taking into account how the railway maintenance planners operate and what they would find useful, and developing ways to provide that information. This was done by showing how the parametric bootstrapped Weibull results could be used to predict the estimated failure range for a broad range of track lengths ranging from a single piece of track, all the way up to any group or collection of track segments going to division or even system level with each made up of their own data.

1.4 Rail Defects

Discussion of methods to improve the accuracy of predicting fatigue defects requires at least some discussion of the rail and track structure, as well as the mechanics of defect growth. Figure 1 shows an idealized track cross section, consisting of the rail, track, ties, ballast, and sub-ballast. The rails support the train, and dissipate the contact forces into the ties, which then spread the forces over a larger area into the ballast, sub-ballast and subgrade.

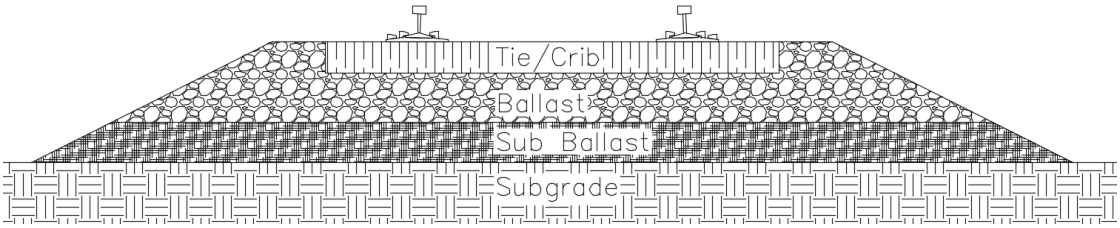


Figure 1: Track Cross-section showing Ballast, Sub-ballast, and Subgrade

Focusing on the rail itself, Figure 2 provides a more detailed view. Of interest is the head of the rail, which is where many fatigue defects originate. This is due to the concentration of stress from the wheel being focused in the area, which can result in an internal flaw or discrepancy in the metal of the rail. As more wheels travel over the rail in this area, that flaw is subject to repeated loading and unloading, and undergoes a fatigue-based failure growth. This growth, if left unchecked, can proceed to expand and remove substantial area from the rail that would go into mitigating the stresses involved, which further exacerbates the failure of the rail.

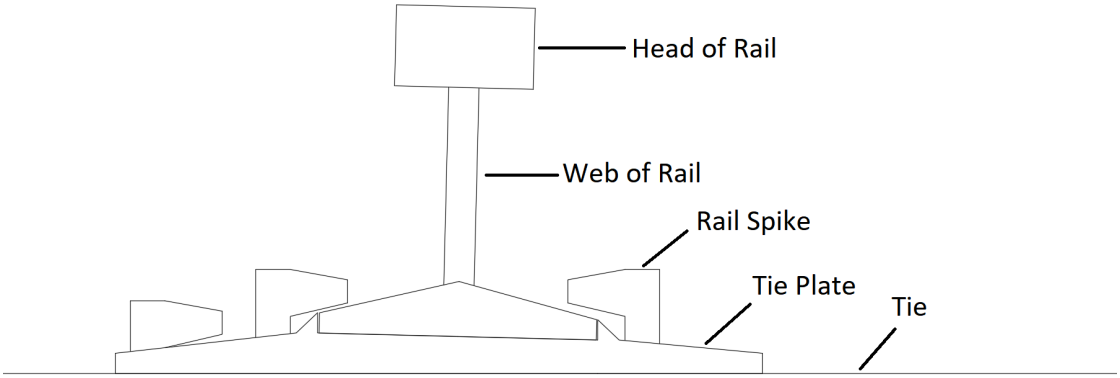


Figure 2: Cross section of the Rail on a Tie

Defects themselves can be found in multiple ways. Historically, defects were found when the rail itself broke, but advances in non-destructive testing have allowed newer technologies such as ultrasonic scanning, to identify defects.

1.5 Weibull Distribution

The Weibull Distribution is a log-log representation of the probability of an event occurring as a function of a defined input. Thus, in its fatigue applications, the Weibull distribution would represent the probability of a defect occurring as a function of cumulative loading. In other

applications, it is the probability of a specified event or observation as a function of a defined input variable (9).

1.5.1 History of Weibull Distribution

The Weibull distribution was first identified by Frechet (10), and applied by Rosin & Rammler (11) on particulate matter, but achieved popularity after Waloddi Weibull described it in detail in 1951 (9) and hence its name. Waloddi Weibull's paper showed how the formula, Equation 1, satisfied the distribution of a variety of observations; yield strength of Bofors steel, size distribution of fly ash, length of Cyrtoideae[2], and more. He explicitly states that the equation did not have a theoretical basis, however that many such distribution functions often do not have theoretical relations with their populations in question.

$$F(x) = 1 - e^{-\varphi(x)} \qquad \text{Equation 1}$$

Weibull initially describes the formula with regard to the form shown in Equation 2, then transformed it from a single case, to the inverse; all but one case, as shown in Equation 3.

$$(1 - P)^n = e^{-n\varphi(x)} \qquad \text{Equation 2}$$

$$P_n = 1 - e^{-n\varphi(x)} \qquad \text{Equation 3}$$

This change shifted the goal of the equation from finding out the probability of a single cause of failure, to the probability that it will not fail from any causes; as Weibull describes it, the failure of a link in a chain. Of note is that this analogy focuses on just a single failure, which will come up later.

1.5.2 Weibull Distribution

The Weibull Distribution, Equation 4, is normally defined by two parameters, Alpha(α), and Beta (β), which are analogous to Slope and Intercept respectively as illustrated in Figure 3. Note, the vertical axis of Figure 3 is the probability of a defect in a single rail (cumulative probability) and the horizontal axis is cumulative loading as defined by Millions of gross Tons of traffic (MGT).

$$f(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}} \qquad \text{Equation 4}$$

[2] Also called Radiolaria or Radiozoa, are protozoa of sub-millimeter size which produce skeletons; often used as a diagnostic fossil.

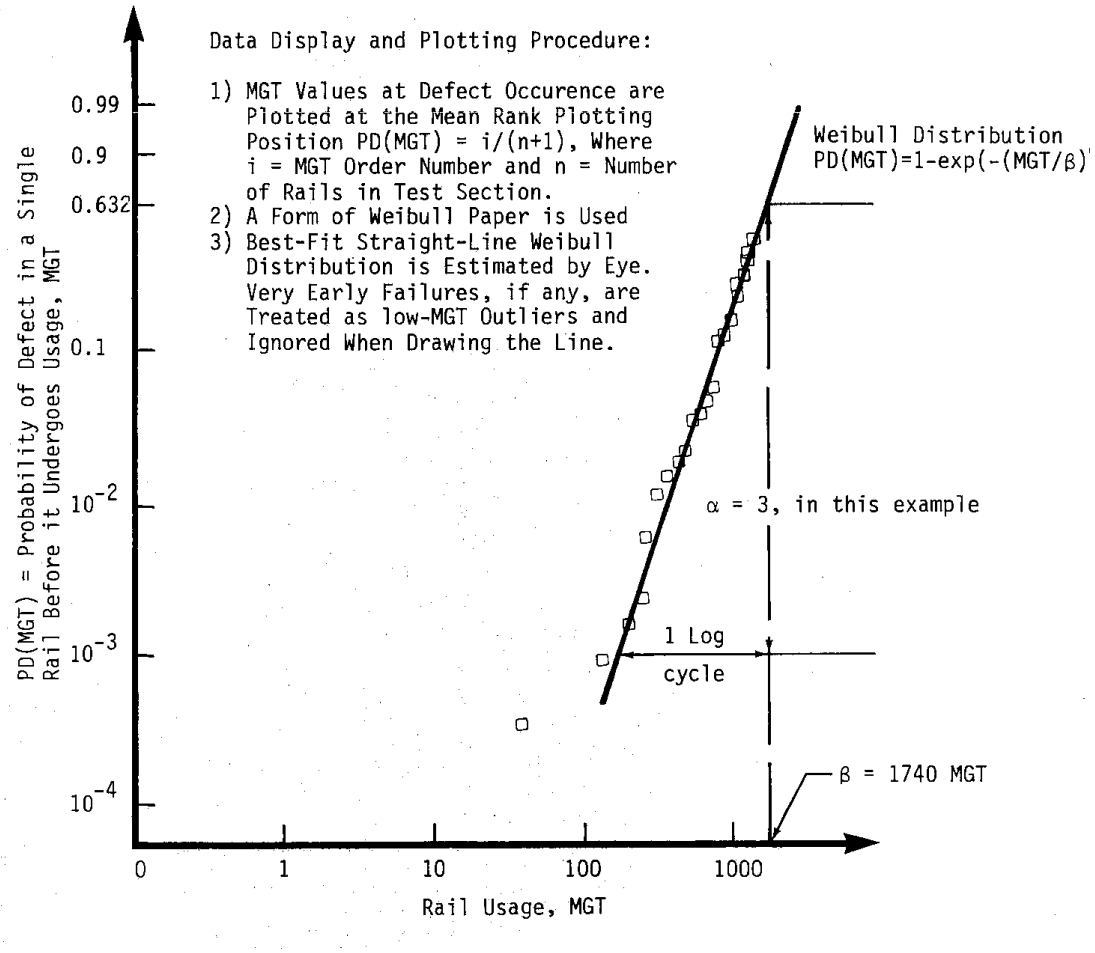


Figure 3: Original example of Weibull plot used in (3)

Different Alpha and Betas will correspond to different probability distributions as shown in Figure 4 and Figure 5.

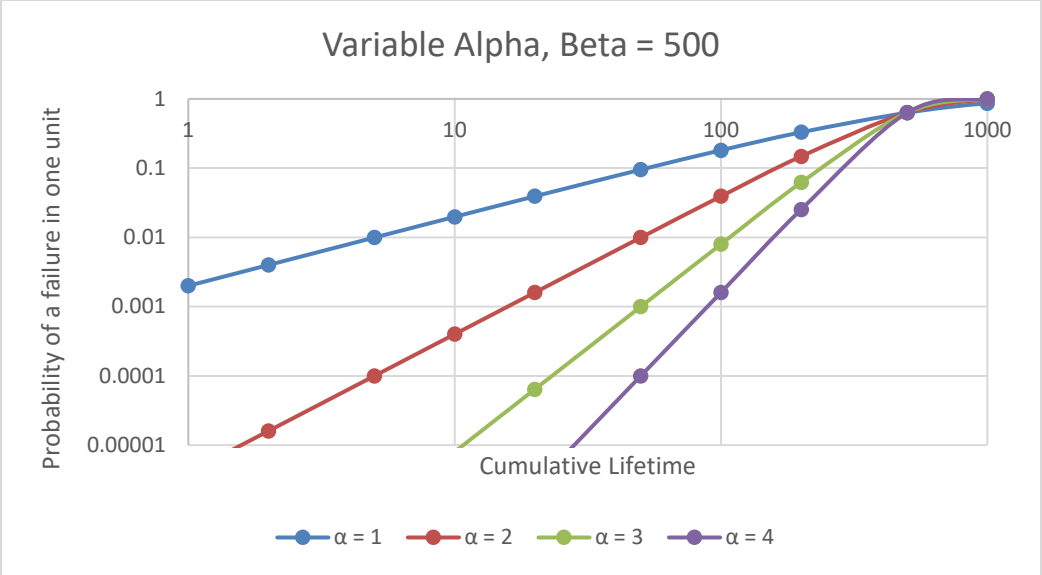


Figure 4: Representative Graph of Varying Alpha values of the Weibull equation

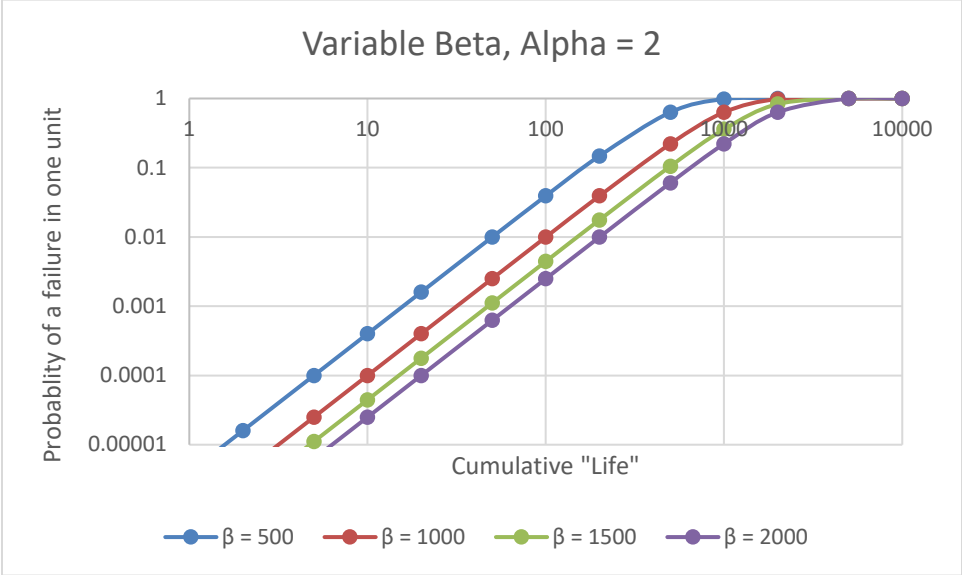


Figure 5: Representative Graph of varying Beta values of the Weibull equation

While the Weibull equation can help determine the chance of failure of a single unit, that poses an issue when there are a large number of units over a long time period. In the same way that a probability can be converted into an expected value, it is possible to convert the Weibull equation into outputting a rate of defects per group of units per year. This is done by differentiating Equation 4, which results in a defect rate function, as shown in Equation 5. This equation allows the estimation of the number of failures per unit of loading (or other input variable) for a specific group of units, allowing the planning to focus more on the overall “health” of the units, instead of

individual units. As shown in Figure 6, varying Alpha results in significant changes to the shape of the Weibull Rate Distribution, while Beta changes the intercept, as shown in Figure 7.

$$f(x) = \left(\frac{Alpha}{Beta^{Alpha}}\right) * \left(x^{(Alpha-1)}\right) \qquad \text{Equation 5}$$

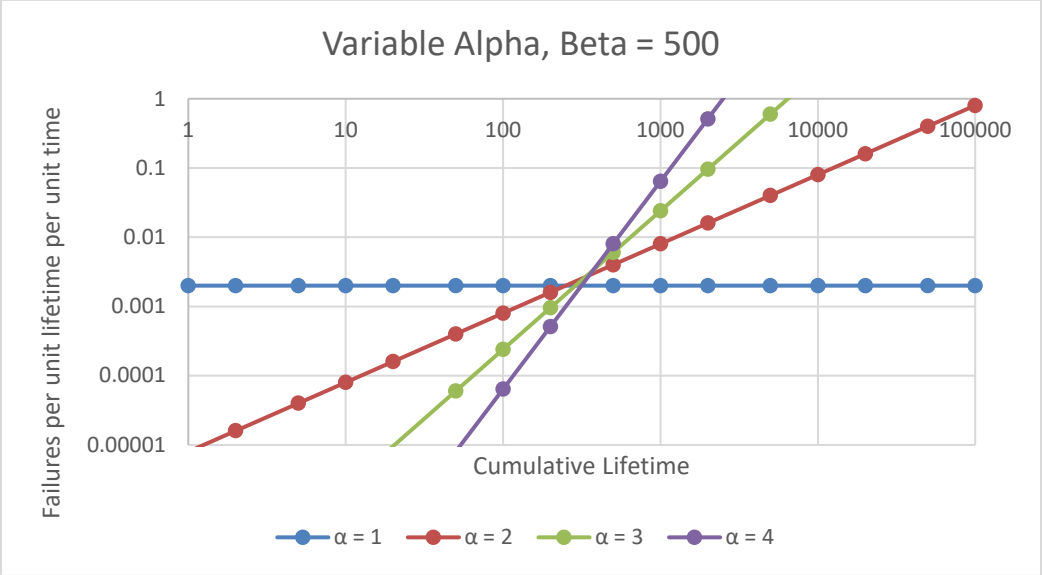


Figure 6: Representation of varying Alpha values on the Weibull Rate function

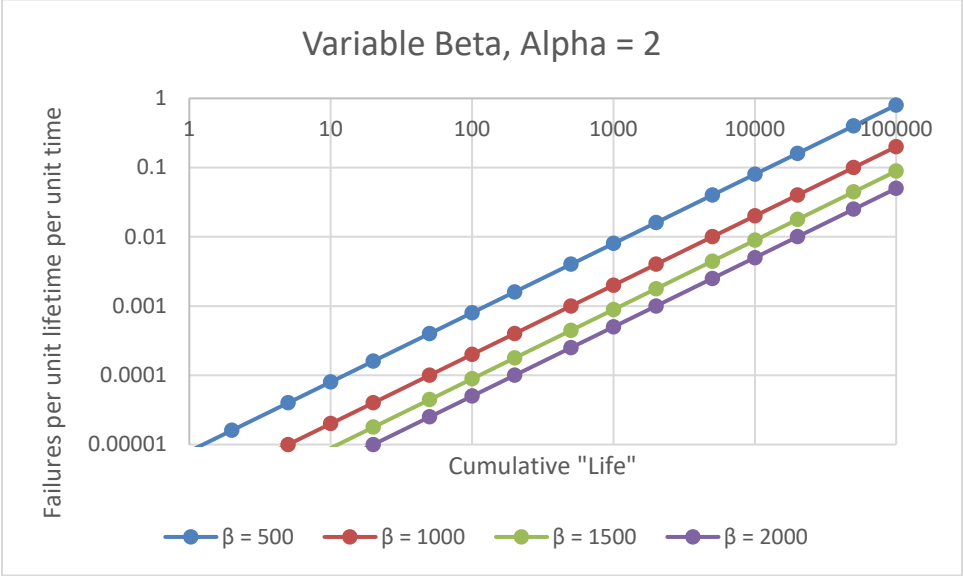


Figure 7: Representation of varying Beta values on the Weibull Rate function

1.5.3 Railroad use of Weibull Distribution

In the late 1970's, the Association of American Railroads conducted research to see how the rate of defect development in freight railroad track could be best modelled. The result was the application of the Weibull distribution, specifically, the two-parameter Weibull distribution, which appeared to fit the rail defect development data (3, 8). This research initially focused on six sections of track totaling about 270 miles of track and 1160 defects, with the rail laid between 8 and 20 years previously (3) and was later extended to several different railroads in the US and Canada (8), as well as varying car loadings (12).

These research efforts worked to develop a relationship between known variables, in this case the cumulative load applied, as well as the when and where defects were found, and the unknown probability of a defect occurring in a unit of rail.

Initially, the research suggested the use of a relationship based on an "Effective MGT" value based on the crack propagation rate, which is tied to the fourth power of stress, as shown in Equation 6, over using plain cumulative MGT[3]. However, the report admits that this would require detailed knowledge of the wheel loadings and load spectra data, which was still in the developmental stages at the time.
\

$$MGT_{effective} = \left[\sum_{all\ wheel\ loads}^{\square}(MGT_{each\ wheel}^{4})\right]^{1/4} \qquad \text{Equation 6}$$

As a result, the research then focused on using the defect data directly, within an application of the Weibull parameters, as defined previously, and specifically the two factor Weibull parameter which appeared to fit the cumulative defect growth data, as presented in references 3 and 8.

Based on this work, the railroad industry started to use the Weibull method for calculating the useful lifetime of the rail, and rail components (13), and has done so for the past 40 years (14). This method allowed railroads to shift from using a years-to-failure approach, to one that focused on the actual use of the rail, which tied in better with the actual physical processes going on that cause defects and failures (15).

Of interest is that many of the research applications mentioned used a single 39-foot rail as the base structural unit, determining that it was a natural unit since rails at the time were manufactured in 39-foot lengths, and that the probability of more than one defect in the same unit would be negligible compared to no defect, or one defect. In addition, infant mortality[4] defects, are often ignored in the analysis, as the Weibull method is not suited to predict them (8, 15, 16).

1.5.4 Current Weibull Distribution methods

---

[3] The sum total of Million Gross Tons of traffic that has passed over the rail in question since it was installed. Used as a lifetime variable, similar to airplane pressurizations.

[4] Failures at the very beginning of the life of the object, usually due to creation or instillation errors, such as inclusions in the metal or being damaged during instillation.

In the decades since the Weibull Distribution was introduced, novel methods of expanding it have been developed in order to better fit the data being examined, as well as provide representation of physical behaviors in the distribution. Many of these, such as (17, 18, 19, 20, 21, 22), include new variables, with (23), a Beta Modified Weibull (BMW) Distribution, appearing to cover the most sub-models by having a 5-parameter (A, B, Alpha, Gamma, Lambda) expansion along with the inclusion of the Beta function. Equation 7 and Figure 8 show the Beta Modified Weibull Distribution Density function, and the relationship with the various sub models going back to the original Weibull, Exponential, and Rayleigh distributions.

$$f(x) = \frac{\alpha x^{\gamma-1} * (\gamma + \lambda x) * e^{\lambda x}}{B(a,b)} * \left[1 - e^{\{-\alpha x^{\lambda} * e^{\lambda x}\}}\right]^{a-1} * e^{\{-b\alpha x^{\gamma} * e^{\lambda x}\}} \qquad \text{Equation 7}$$

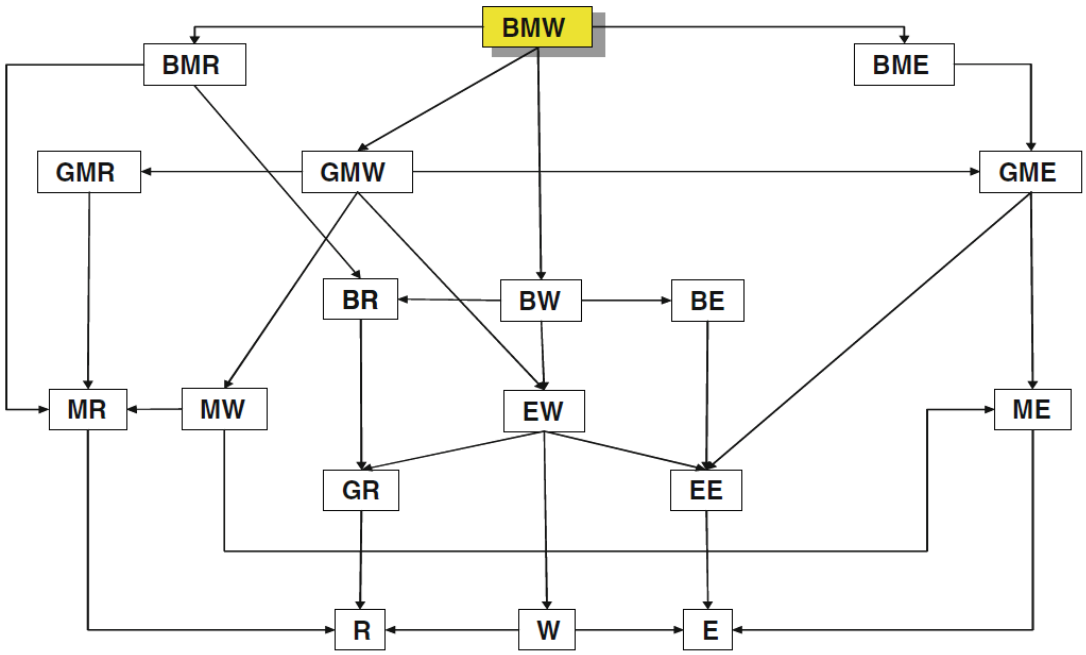


Figure 8: Relationships of the BMW sub models; adapted from (22)

Where:

W = Weibull distribution
R = Rayleigh distribution
E = Exponential distribution
E- = Exponential, such as Exponentiated Exponential (EE)
G- = Generalized, such as Generalized Rayleigh (GR)
B- = Beta modified, such as Beta Weibull (BW)
M- = Modified, such as Modified Exponential (ME)

As for actual uses, several industries still use the standard 2-parameter form, such as the automotive industry, which uses it to determine the failure rates of components (24, 25, 26), and the airplane industry (27, 28). However, some fields have seen issues with using the Weibull

distribution, where conditions do not follow the basis of the formulization, such as: brittle ceramic failures (29), the tensile strength of composites (30). Other industries use various failure approximations, including but not limited to logistic regression (31, 32) and K-Nearest Neighbors (33, 34), to calculate a total risk assessment (35, 36, 37, 38). In other fields, when a crack or defect is not deemed as failure critical, methods which estimate the time until the crack or defect becomes failure critical are used, such as in (39, 40, 41, 42, 43, 44). This stimulates the movement away from Weibull, towards other methods which fit the empirical evidence.

1.5.5 Going beyond Weibull Distribution for Failure Analysis

Several industries are shifting towards Finite Element Analysis models for determination of failure rates, computationally simulating the object of interest over various inputs, and observing when and where material limits are exceeded, as in (35, 45). This sees use in multiple fields where safety is of utmost importance, and when the object in question is of reasonable size, such as aircraft and automobiles. But, while the growth of computational power and methods of mimicking, or creating a "digital twin" of objects has increased dramatically in the past ten years, the ability to accurately process tens of thousands of miles of track is still some time away. As reference 46 mentions, even for smaller objects, high accuracy requires considerable effort in both modeling the object in question correctly, and the computation of the simulations.

In some cases, mainly with track geometry defects, the railway industry has moved to other methods of prediction, such as Multivariate Regression Splines (MARS) (47), using Naïve Bayes and Bayesian networks methods to link geometry defects to rail defects (48), and Logistic Regression methods to link Ground Penetrating Radar data to geometry defects (49, 50, 51). These advances have been spurred on by the development and use of "Big Data" analytics (52), where the advances in recording and storing data about track conditions, defect locations, and usage statistics have developed large databases that are ripe for analysis.

As mentioned in the previous section, modifications to the Weibull method, such as the Beta-Modified Weibull distribution, are being developed for use in failure analysis. These newer, more complex, distributions can help to find a better formula to fit the known data points. Of course, these new methods come with caveats, such as having a parameter translate to a physical state that may not exist, such as a failure-free time period.

1.5.6 Failings of Weibull Distribution

The Weibull Distribution, as used by the railroad industry, reports the probability of a single defect in a 39-foot length of rail. However, as data collected has shown, the assumption that a 39-foot length of rail will see only one or no defect, and the probability of multiple defects are negligible, does not hold true. This stems, paradoxically, from the increase in safety measures and monitoring of the rail for defects; instead of only finding defects when they cause the rail to fail, it is now possible to detect defects before the rail fails through non-destructive methods, such as ultrasonic testing. However, since such testing cannot be done continuously, defects can originate and grow in the periods between testing, usually several months, at which point several of them are found at once. In addition, since these testing methods are not perfect, sometimes defects missed in one run are found at a later date, when their increase in size makes them more apparent. This results in

multiple defects being recorded on the same day, essentially the same Cumulative MGT, and thus violates the assumption that the Weibull method is based upon.

Another shortcoming of the Weibull method is that for rail that has been in track for a very long time, 30 years or more, data on the earlier year's defects is often incomplete or not available. This then does not allow for an accurate compilation of cumulative defects and introduces a potentially large error in the analysis. This has been seen in numerous track segments analyzed in this study.

In addition, the Weibull method is bound by probability on the 0 to 1 range, while the Weibull Rate method, commonly used to set thresholds for maintenance, is unbounded positively. This, as shown in the research performed, leads to situations where backtracking from a set rate of defects leads to the Weibull method reporting "100% chance of defect". Since this only means a single defect, issues arise when trying to convert from a Weibull probability to an expected number of defects.

1.5.7 Bootstrapping

Bootstrapping is any test or metric that relies on random sampling with replacement. Bootstrapping allows assigning measures of accuracy (defined in terms of bias, variance, confidence intervals, prediction error or some other such measure) to sample estimates. This technique allows estimation of the sampling distribution of almost any statistic using random sampling methods. Generally, it falls in the broader class of resampling methods (53).

Bootstrapping is the practice of estimating properties of an estimator (such as its variance) by measuring those properties when sampling from an approximating distribution. One standard choice for an approximating distribution is the empirical distribution function of the observed data. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of resamples with replacement, of the observed data set (and of equal size to the observed data set).

It may also be used for constructing hyporeport tests. It is often used as an alternative to statistical inference based on the assumption of a parametric model when that assumption is in doubt, or where parametric inference is impossible or requires complicated formulas for the calculation of standard errors.

In the process of classifying the output for the methodology developed, it falls into a confidence interval use of bootstrapping. The output itself is saying that, given similar rail sections, there is a confidence interval of some probability that the track will report back a number between the upper and lower bounds.

1.5.7.1 Methodology of Bootstrapping

For normal bootstrapping, the "output" is developed from sequentially picking, with replacement[5], the known observations. Given a set consisting of {1,2,3}, bootstrapping will create new sets from those values, such as {1,1,2}, {1,2,3}, or {3,3,3}. Individually, these sets don't provide the insight into the makeup of the population of the dataset, but together, the sets can provide some insight into the distribution of the data while also providing a higher number of samples to work with. If the samples are the same length as the source data, or there are enough samples taken, then these samples tend to mimic the same distribution as the initial data set.

1.5.7.2 Parametric Bootstrapping

Parametric Bootstrapping operates in a similar way to normal Bootstrapping, but exchanges the use of the exact values, for the use of a probability function. This allows more variation in the output, along with allowing values to exceed the known observation values. Instead of picking sample values based on the source data values, the source data is transformed into a distribution function, which is then used to develop the samples. For example, if the source data was approximated by a normal distribution centered around 0, with a standard deviation of 1, we would expect that some 99% of all samples would occur within [-3,3], but there lies the possibility of having extreme values, such as 5 or -8. Of course, these values do not need to be whole numbers; given that the distribution is continuous, the resulting samples will be as well.

[5] The new observation is returned to the set of known observations.

CHAPTER 2 DATA

2.1 Source of Data

The data used for this analysis came from a Class 1 Freight Railroad operating in the United States of America, as well as companies contracted to provide maintenance operations, such as ultrasonic testing. This data was collected during routine operations and maintenance efforts, not as part of this research. It represents 20 years of data collected on a 20,000+ mile railroad, with over 200,000 defects in the data set.

This had two major impacts on the research; the data was "real-world" data, as would be seen by the railroads during normal operations, and that the collection of the data did not require the research to be delayed. The use of real-world data posed several issues, the most common being data quality, a common real-world problem, where the data had gaps in it for various reasons; data collection errors, administrative changes, or otherwise unavailable. However, dealing with these issues only helps to reinforce the usefulness of the results built from it, as it shows that even badly-behaved real-world data can provide useful results. In fact, the resulting analysis approach is based on use of such real-world data with associated issues in data quality.

2.2 Data Received

As detailed in the table below, the data was received over the course of the research, which resulted in multiple analyses, with each new analysis incorporating the additional data set. The initial data was received directly from the Class 1 railroad as part of the University Transportation Center's (UTC) research for railroads, and portions of the data had been used previously in earlier research projects. Later data was received both from the railroad itself and from contractors who had worked with the railroad, and were allowed to share the data under the prior UTC agreement. Table 2 presents a summary of the data received.

Table 1: Summary of File Acquisitions

File Name	Description	Date Acquired
MGT <Year>, MGT 2010	Three Excel files covering 2010, 2011, and 2012 reported Annual MGT	Jan 2015
Rail Grind Milepost	Four Excel files covering 2011, 2012, 2013, and 2014 Rail Grinding Operations	Jan 2015
All Curves	Single Excel file with all reported Curves	Jan 2015
All Defects by Year	Three Excel files containing all Defects reported for 2010, 2011, and 2012	Jan 2015
Detected Rail Defects	Single Excel file containing all reported Defects between 1/1/2014 and 12/31/2016	Mar 2017

Rail in Track	Single Excel file containing all reported Rail locations in the network for 2017	Sep 2017
MGT <Year>, MGT 1988	Individual Excel files containing the Annual MGT data for the years of 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 2005, 2007, 2009, 2010, 2011, 2012	Sep 2017
Defects All	Originally a single Excel file, split into two due to data limitations; Contains all reported defects from 1/1/1999 to 5/31/2017	Dec 2017

As newer or more complete data was acquired from the railroad, as shown by the date received column in Table 2, previous data was removed and/or consolidated resulting in a single complete set of data. This was done at the starting points of the analyses, as changing the source data midway through the analyses would delay the work, requiring the re-computation of analyses already done.

2.3 Structure of Data

The data was primarily received as Microsoft Excel files (.xls, .xlsx), or was extracted from a Microsoft Access database and then saved as an Excel file. In order to import the data into the R Programming Language, the file formats were changed to the Comma Separated Value (.csv) format, which is handled by R's data handling routines, whereas importing the .xls/.xlsx files would require various workarounds and plugins that did not seem to be suitable at the time. With the exception of converting Microsoft Dates to standardized dates, the data itself was not changed by this conversion. Microsoft Excel stores dates as the number of days since 1/1/1900, instead of as a plaintext MM/DD/YYYY format, such as reporting 1/1/2000 as 36526.

2.4 Ranges of Data Inputs

The date was introduced into a master data base as a series of files. The following tables detail the contents of the files; what the data represented, and the range of that data.

Table 2:INPUT.MGT file data ranges

ID	Numeric	Number specifying the row the data was in
Prefix	Alphanumeric	3 characters corresponding to the location of the track
MP.From	Numeric	The starting milepost value
MP.To	Numeric	The ending milepost value
MP.Tot	Numeric	The total number of miles in the track segment
MILES	Numeric	The total number of miles in the track segment; not the same as MP.Tot

Table 3: INPUT.DEFECT file data ranges

Division	Text	Broad location identifier
Subdivision	Text	Midrange location identifier
Prefix	Alphanumeric	Narrow location identifier
Milepost	Numeric	Milepost of the defect
Track_type	Alphanumeric	What the track was labeled as (Main, track 1, track 2…)
Track_code	Text	Identifier for Main, Siding, or Other
Side	Text	Which side of the track the defect occurred on
Defect_type	Text	Shorthand for the type of Defect
Size	Numeric	Size of defect
Date_Found	Date	Date the defect was found
Date_FD	Date	Duplicate? of Date_Found
Car_Name	Alphanumeric	Identifier of the track inspection car which found the defect, or In-Service defect
Prepared.By	Text	Who prepared the defect report
Curve.Tang	Text	Track type where the defect was found
Roadmaster	Text	Roadmaster of the track where defect was found
Joint.Weld	Text/numeric	Primarily W, J, or “blank”, but values from next column blended in
Rolled.Year	Date	Year the rail was rolled
Mill	Text	Which mill cast the track
Weight	Numeric	The weight of the rail in lbs/yard

For the INPUT.DEFECT file, the Joint.Weld data had to be double checked, as the rolled year data had somehow merged into the column. While the Joint.Weld data was mostly (~75%) blank, 8896 entries were dates, which necessitated checking and moving those values to the appropriate column.

Table 4: INPUT.RAIL file data ranges

ID	Numeric	Row ID of the track segment
Prefix	Alphanumeric	3-character location identifier
From.Milepost	Numeric	Starting milepost for the track segment
To.Milepost	Numeric	Ending milepost for the track segment
Track	Alphanumeric	Identifier for the track number (Main, Track 1, Siding…)
Left.Rail.Weight	Numeric	Weight of the Left Rail, in lbs/yds
Right.Rail.Weight	Numeric	Weight of the Right Rail, in lbs/yds
Left.Rail.Laid	Date	Date the Left Rail was laid

Right.Rail.Laid	Date	Date the Right Rail was laid
New.Relay.Left	Text	Was the Left rail a relay?
New.Relay.Right	Text	Was the Right rail a relay?
Joint.Weld.Left	Text	The type of joint used for the Left rail
Joint.Weld.Right	Text	The type of join used for the Right rail
Curve.Degrees	Numeric	Degree of curvature, for curves
Curve.Direction	Text	Handedness of the curve
Division	Text	Division in which the track was located
Subdivision	Text	Subdivision in which the track was located
LR_Date	Date	Left Raid Laid Date
RR_Date	Date	Right Rail Laid Date

The INPUT.Rail file had a few interesting quirks to it. First, the Left/Right Laid Date, and the later LR/RR Date were duplicates, which meant that one pair of them could be removed to compact the dataset. But, the dates themselves tended to always be January 1st, possibly due to a lack of definite knowledge of when the rail was laid. As for deciding on what counted as Left and Right, it was assumed that positive increase in milepost would be the direction in which left and right would correspond to, and it was assumed that this held true for all cases where Left/Right were used.

2.5 Cleaning and Joining of Data sets

While the process varied between Weibull Analyses, the cleaning and joining of the data sets followed a similar pattern across all of the analyses performed. Initially, columns which contained extraneous information were removed, such as rarely filled-in Latitude and Longitude data. Next, the columns were formatted to the appropriate data type. When loading the .csv[6] files into R (54), the data file assumes a data-type called "Factor", which is restrictive to what can be done to it; by changing the data from "Factor" to Numeric or Character, the data assumes the form of numbers or a text string, allowing easier manipulation of the data. It is at this point that the track data had the length of the track computed and bound to the dataset as well, by taking the absolute difference between the start and end mileposts listed. For text/string values, there was additional cleaning done, in order to correct issues in the data due to transcription errors, non-standard word choice, and shortening full names. This consisted of things such as converting a track value from "SG", "S1", and "Main", to "1"[7].

Once all of the input data was cleaned and properly formatted, the track and MGT data was bound together. This process consisted of taking each individually reported track segment, finding all of the breaks in homogeneity in that segment, such as changes in curvature or annual MGT, and then creating new track segments based on those breaks. As a part of the creation of the new track

---

[6] Comma Separated Value format, a common method of saving plaintext data with delimiters which is easy to load into many programs

[7] Respectively, Single Track(SG), Singe One (S1), Main Track (Main), and Track 1(1)

segments, the track segments would have the correct data applied to them, instead of requiring another pass to apply that data.

Starting with the Third Weibull Analysis, the track data was noted to have gaps existing between reported segments, apparently as a byproduct of the format the data was stored in, which truncated the milepost information. As such, synthetic track segments were created, based on the data of the neighboring track segments. This process essentially cloned the neighboring lower-milepost track data, but changed the Milepost and derived data.

This new dataset using the combined data would then be used in any further analyses. At this point, it was also necessary to separate the track into two rails, left and right rail, thus shifting the analyses from a "track" to a "rail" analysis. This split was based upon the differences in rails for the same segment of track. Rails on the left of a track segment may be a different weight, or more often, laid at different dates. By splitting the data into rail segments, it became possible to calculate Cumulative MGT, the summation of annual MGT since the rail was laid, based on the individual rail's Laid Date, the date the rail was said to be installed in operational service, instead of one based on both rails' laid date. While this doubled the size of the datasets used, and thus their computational requirements, it allowed finer control over the data. It also allowed for more careful collection of homogeneous data. It was expected that this would help with the issues found in prior analyses.

2.6 Lessons Learned for Manipulating Data

The first lesson learned for manipulating the data was that care needed to be taken when cleaning the data. The varying input values, including unexpected values such as text when the value should be a number, can easily mess with the structure of the dataset, which results in errors and erroneous results when forced through computations.

The second lesson learned was that for adjusting/correcting variables that were collected, it was far more effective to go by columns than by rows. By inspecting each individual row for an offending variable, time is spent pulling that row out, checking it, and then putting the row back, with corrections if needed. By wholesale adjusting columns, complicated IF/THEN statements were avoided, drastically shortening the required code, as well as making it far easier to understand.

A third lesson learned for manipulating the data was to frequently clean the working memory of unused variables. Over time, variables that were used once or twice as a stepping stone to reach a further endpoint ended up making up a large part of the working memory in use by the program. These unused variables were often created in order to debug errors that cropped up, such as a faulty adjustment, or a data error, without requiring the whole process to go back to the beginning. While the completed code could be condensed and eliminate the need for such temporary variables, the utility to use them to help track down any errors that may crop up with new datasets lends the advice to just make sure the variables are cleaned up after the data input process is completed.

CHAPTER 3 WEIBULL ANALYSIS

3.1 Traditional and Expanded Weibull Analysis

The initial 2-Parameter Weibull Analysis was run for two primary reasons: to provide a baseline in which all other analyses can be compared to, and to provide information on the data. The choice to use the 2-Parameter Weibull method stemmed from its use in the railway industry already; mimicking the same way they would do the analysis on their own. After combining the track data along geographic location (Prefix) values, the defect data was then processed and bound based upon matching locations given in the track data. Initially, the process of correlating defects with track segments was based on the unique combinations of Division, Subdivision, and Prefix, but was changed to use just the Prefix due to the general uniqueness of the Prefix data compared to the additional uniqueness from Division and Subdivision. Later, this was changed to only Division, as Prefix was shown to have issues across years, where Prefixes changed year-to-year, eliminating their use as unique identifiers for track location. As such, Division was then chosen, with more dependence on milepost values, to bind track together, as shown in the grouping of track function. Concurrent with this, the Defect dataset also had the track data bound to it, developing a parallel dataset that was focused on the defect data. As the defects were bound together, their Cumulative MGT was calculated, based on the Annual MGT data and the age of the rail containing the defect, as shown in Equation 10. The age of the rail was taken from the Defect data's Year Laid, compared to the current year of interest, and taking into account a 2% annual rate of change of MGT.

$$Estimated\ Cumulative\ MGT = AGE * \left(1 - (2 * \left(\frac{\frac{AGE}{2}}{100}\right))\right) * Annual\ MGT \qquad \text{Equation 8}$$

Where:

AGE = Current year – Installation year.
Annual MGT = Annual traffic in Million gross tons (MGT)
Cumulative MGT is the sum total Annual MGT from installation to current year.

As shown in Figure 10, this equation works well when AGE is a small value, but as AGE increases, eventually the Estimated Cumulative MGT starts to decrease, eventually becoming negative after 100 years. While main track should have been replaced in the 50 year range, siding and yard track that saw low usage, or in cases where the date laid was in error, ended up with Cumulative MGT values which did not make sense, such as the right hand side of Figure 9.

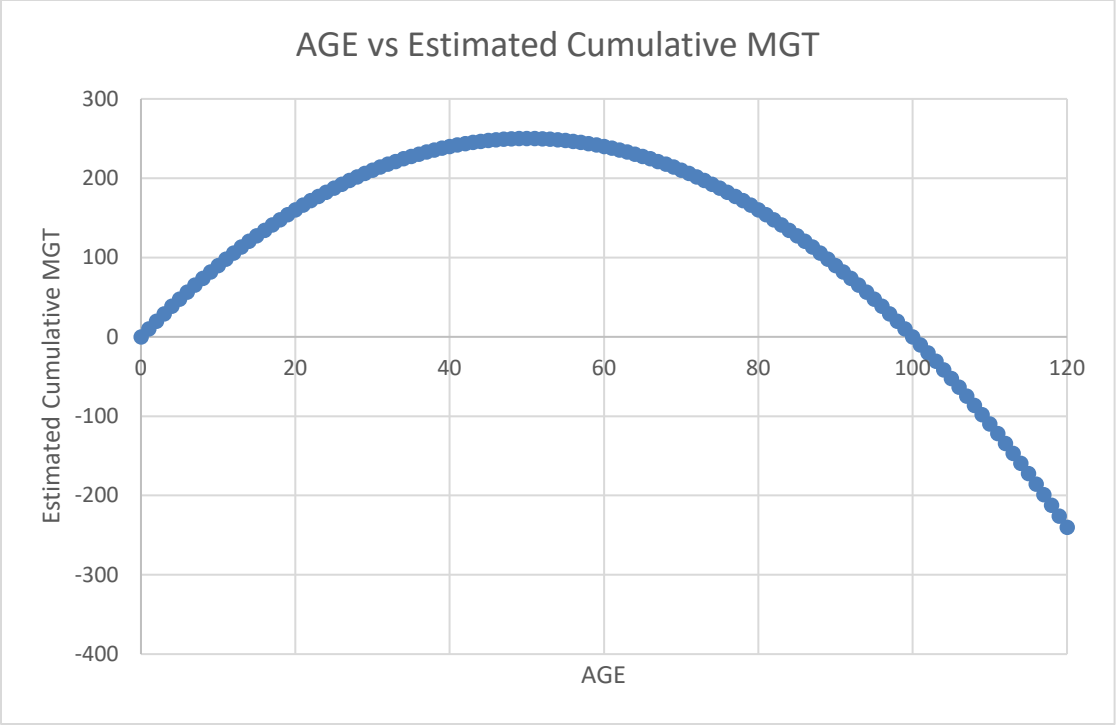


Figure 9: Age vs Estimated Cumulative MGT per Previous Equation

At this time though, there was no other option for developing a way to calculate Cumulative MGT based on the few data points on-hand. However, once more data was acquired later on in the analysis, the cumulative MGT would be recalculated based on historical trends.

At this point, the datasets were bundled together based on their Division specification, primarily to reduce the computational effort needed for analysis until a promising avenue of development was found. It was also thought of at this point that the varying effects of the location conditions could become apparent through comparing the outcomes of the analyses, such as seeing higher occurrences of one type of defect over another, that could be traced back to environmental factors.

3.1.1 Basic Weibull Analysis

The process of computing the Weibull Parameters started by pulling all of the defects that fit the same Division/Subdivision/Prefix location identifiers up to a nominal 30-mile total length, and ordering them based on their Cumulative MGT. Next, the defect data is cleaned of any defects that may have an MGT Age which is less than 0, indicating that the defect occurred in rail that had not yet been laid (most likely a data error), and defects which did not have a Rolled Year, indicating that the age of the Rail could not be defined by the data on-hand. The next batch of code applied the Median Ranking method to the defect data to determine what each defect's probability would be when plotting on the Weibull graphs. As shown in Equation 11, the Ranking depended on the total number of rails being observed, and the ordinal index number of the case being looked at.

$$Rank = \frac{Ordered\ Index\ of\ Current\ Observation}{Total\ Number\ of\ Observations} \qquad \text{Equation 9}$$

With the probabilities and the corresponding cumulative MGT, the data points could now be plotted and have a line fit to them, as shown in Figure 10.

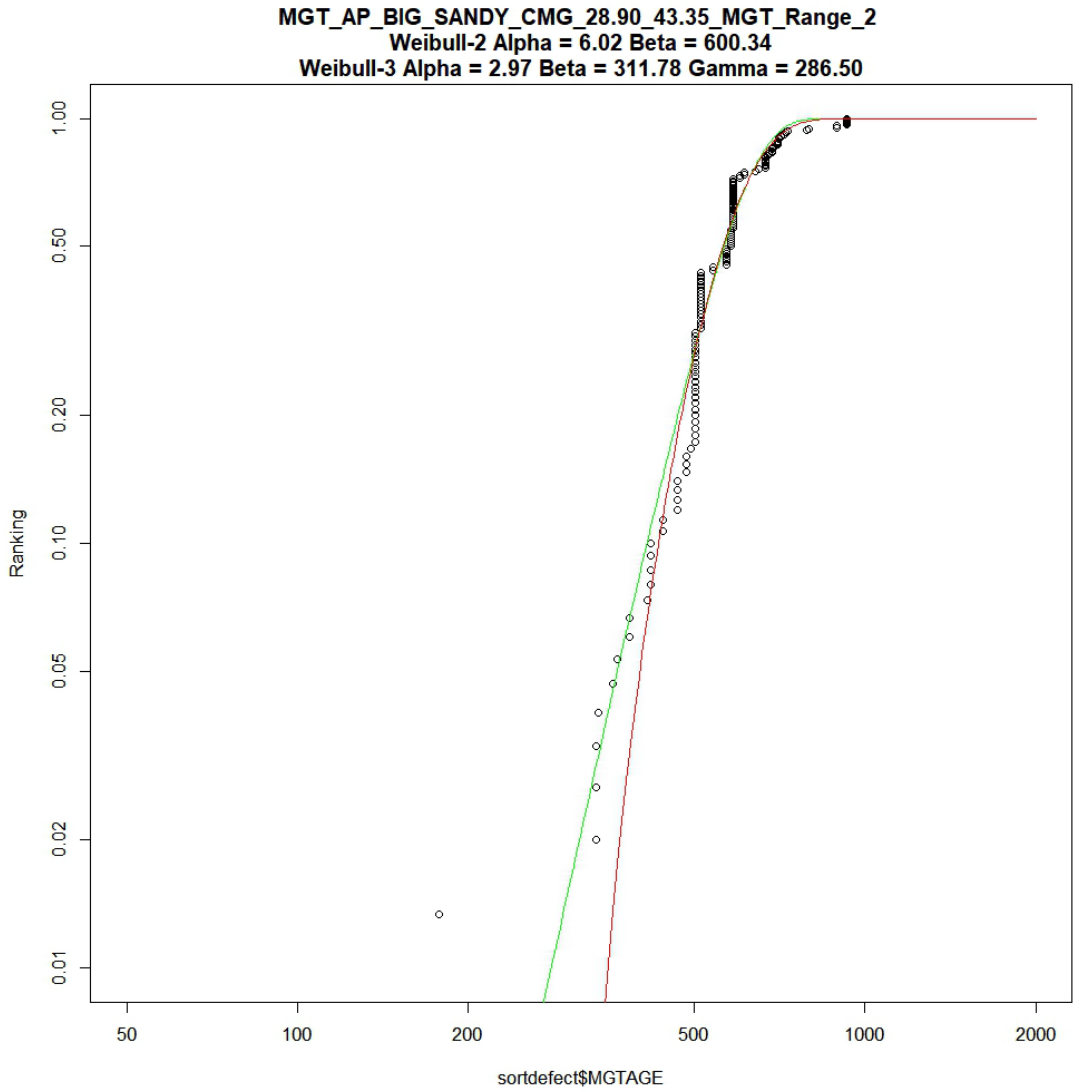


Figure 10: Early Analysis Weibull Plot; Green is the 2-parameter fit, Red is the 3-paramter fit

As shown, there is a strong vertical grouping of defects, which came about due to the collection methods employed by the railroad. Instead of spotting each defect when it became a break, non-destructive testing spotted the defects prior to them breaking the rail. Due to the railways operating the non-destructive testing on an interval basis and not continuously, there is a pronounced clustering of the defects on Cumulative MGT values, as any defect which could be seen was recorded, and then removed from the track. As for the points themselves, future analyses condensed the points together on the highest Ranking value of the cluster at each reported Cumulative MGT. This was done to compute a cautious value for the Weibull analysis, given that the exact MGTs that the defects would have been found through service failures was unknown.

In the very first few Weibull functions fit and plotted, defects which had identical cumulative MGT were not consolidated but treated as unique points. This was later changed once it became apparent that many defects were "stacked" due to the way the defects were recorded. Fitting a line to the data points was done using R's Nonlinear Least Squares function (nls). The first of these, Equation 12, is the fit to the basic 2-Parameter Weibull equation, solving for Alpha and Beta values. Equation 13 is fitting the 3-Parameter Weibull equation, solving for Alpha, Beta, and Gamma. The 3-Parameter Weibull differs from the normal 2-Parameter model by including an offset value, Gamma. While this offset value often improves the fit of the Weibull curve, it corresponds physically to a "defect free" period, which as numerous following figures show, was often not the case. However, due to the simplicity of adding a second fitted line to the analyses, it was often included in calculating and plotting of the data.

$$P(D) = 1 - e^{-\left(\left(\frac{MGT}{Beta}\right)^{Alpha}\right)} \quad \text{Equation 10}$$

$$P(D) = 1 - e^{-\left(\left(\frac{MGT-Gamma}{Beta}\right)^{Alpha}\right)} \quad \text{Equation 11}$$

These fits are then graphed, using the same Alpha, Beta, and Gamma values found, to generate a smooth line showing the fit, as well as the defect points overlaid on the same plot. This allows a visual check of the fit, similar to the process originally done in (3). Initially, the Ranking values were normalized between 0 and 1, due to an error in coding. These graphs and results, as shown in Figure 10 did provide some insight into some problems with the data.

3.1.2 Alterations to Basic Weibull Analysis

After the initial run through the data, seeing the issues presented, and fixing errors in the code itself, further analysis was deemed necessary in order to find the cause of the wide variation in Weibull Parameters. Prior work in both industry and research had tended to show an Alpha range of 2 to 4, and a Beta range of 1500 to 3000, while the current data did not fit with the prior results. This led to several iterations of the Weibull Analysis, varying the grouping of the data, in order to see if there was some unreported step or behavior that was done with prior work.

The first alternation was shifting to whole-division grouping. Instead of subdividing the data into groups of a nominal 30-mile length, the data was bound based on just the Division value. This resulted in far fewer results, 61, due to some divisions not having enough data points for a Weibull curve to be fit, and parameters found. In addition, these analyses saw the first reduction in data points due to identical Cumulative MGT values. The process removed all duplicate values except for the one with the highest Ranking, as shown in the differences between Figure 11 and Figure 12. The difference in Alpha and Beta values were usually minor, as shown in the reported Alpha and Beta values for the two following figures.

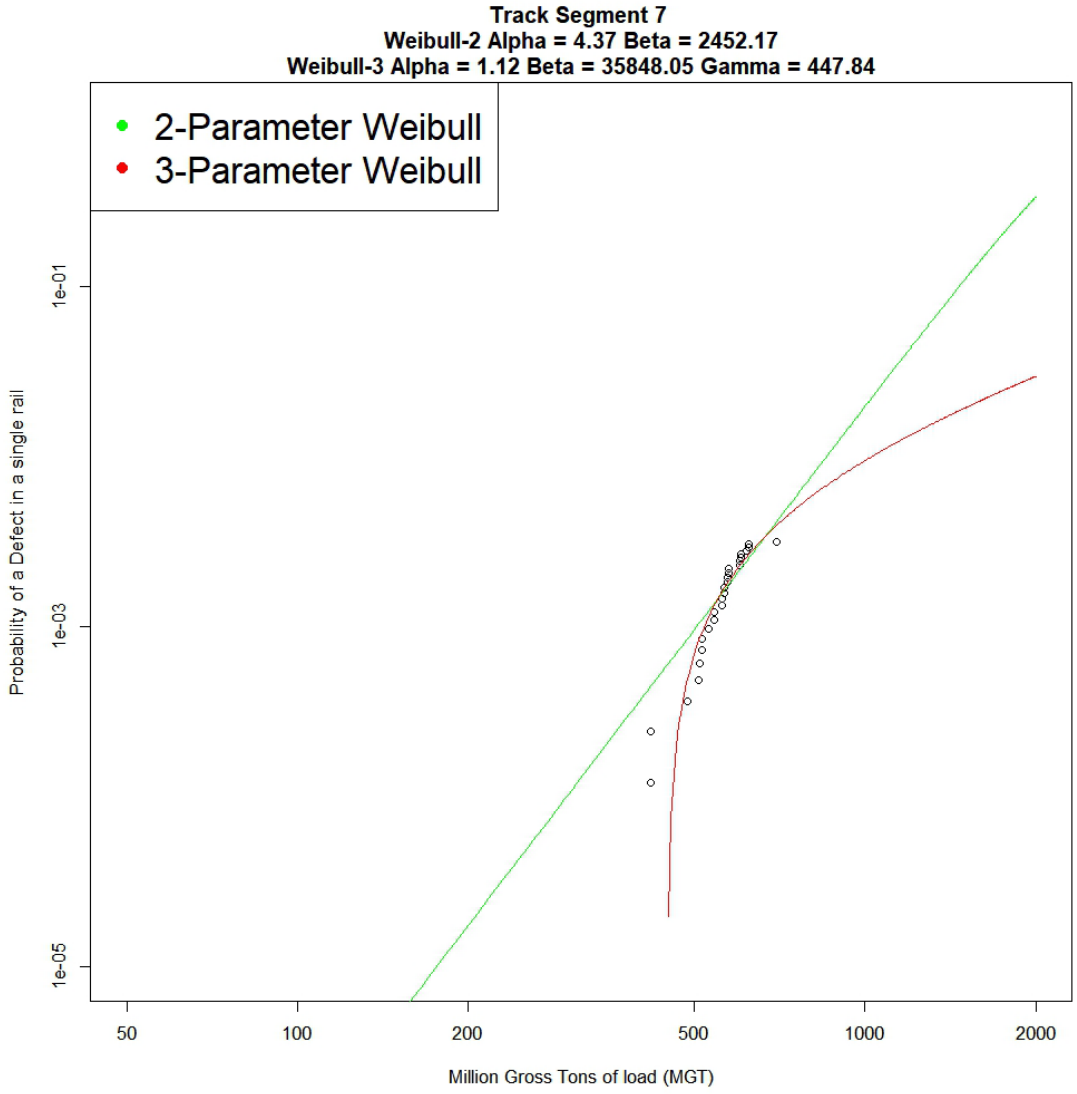


Figure 11: 2 and 3 Parameter Weibull with duplicate Cum. MGT

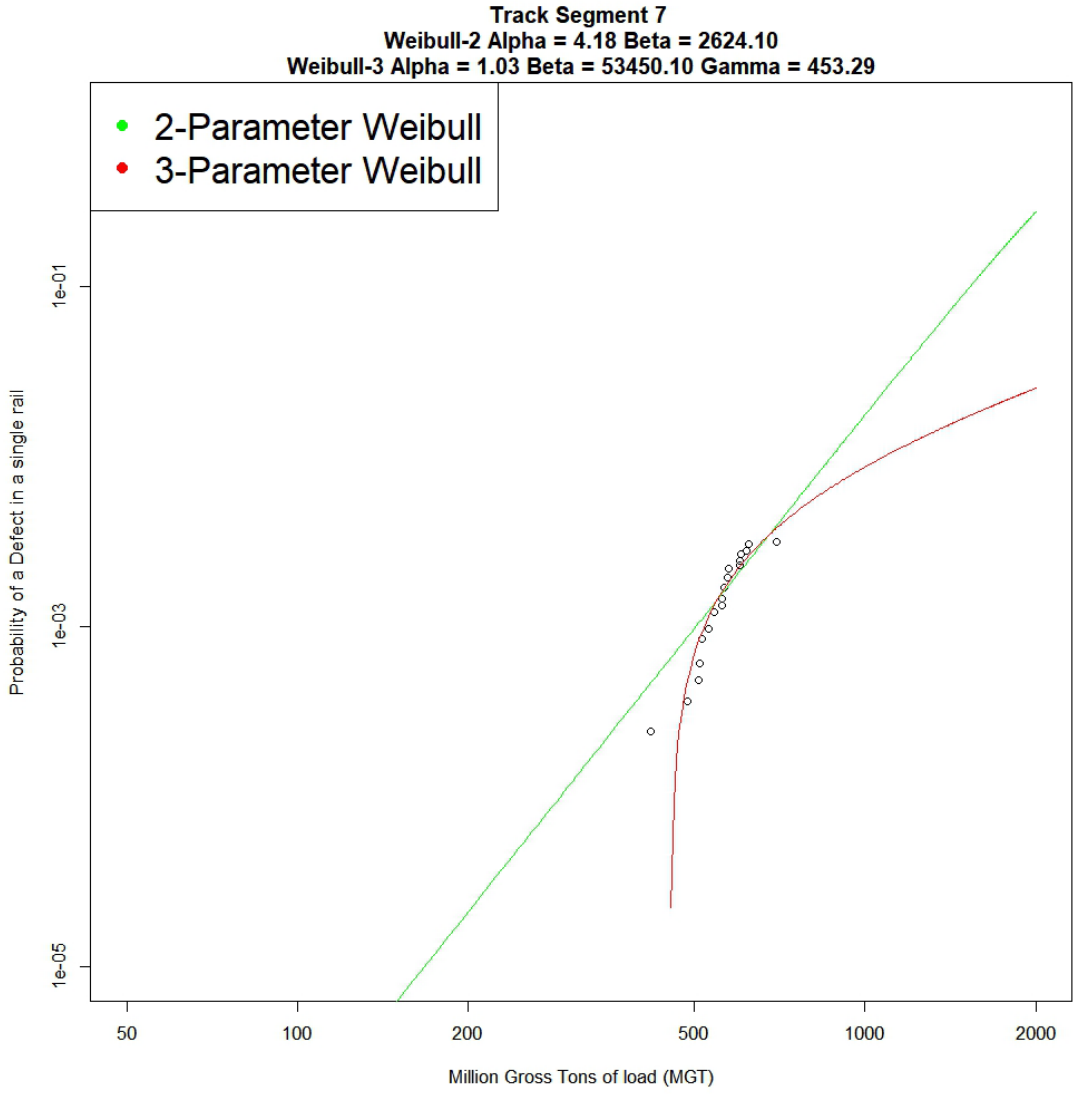


Figure 12: 2 and 3 Parameter Weibull without duplicate Cum. MGT

The second alteration looked at restricting the grouping based on Annual MGT, and length of track. The track saw the removal of all segments below the mean, and then anything beyond 25% of the new mean. This had the result of cutting off low Annual MGT trackage, which would tend to see fewer defects anyway, and the very high MGT outliers which would see many defects. In addition, the track was then grouped into nominal 30-mile segments, reducing the impact that some track might have on the overall results. An issue with this analysis is that there were far fewer track groups that had enough datapoints to compute a Weibull Alpha, Beta, and in the case of the 3-parameter Weibull, Gamma values. There were 57 sets of valid values compared to the first Weibull analysis's 352 sets of values.

The next major alteration was the restriction of defect and track data based on full history. As the data only contained defects going back to 1999, any track that was laid prior to that date was removed from consideration. In addition, defects with a track rolled date prior to 1999 were also

removed. This had the result of severely limiting the possible datapoints for a fit, oftentimes to the minimum 3 points required, and resulted in only 32 sets of Weibull values. In addition, these values tended to be more extreme than prior results, mostly due to the reduction in track length raising the Ranking values, as shown in Figure 13.

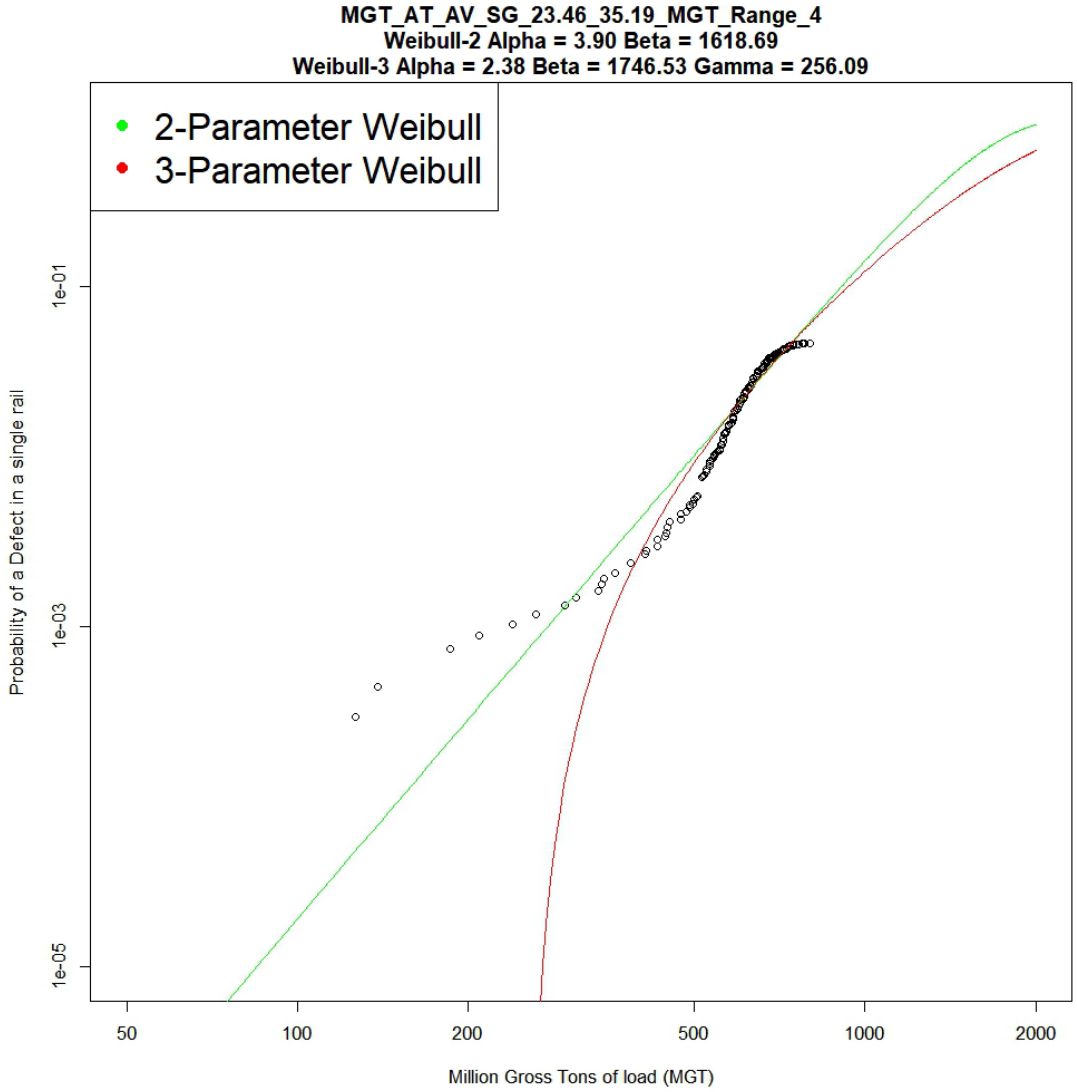


Figure 13: 2 and 3 Parameter Weibull plot using Full History

Following the lack of results with the focus on "full history", the approach was reversed back to the Division grouping at 30-mile lengths, but this time was split based on the type of defect found. As shown in Table 6, a majority, over 60%, of the defects were of the Detail Fracture classification, with split heads, both vertical and horizontal, making up another 20% together.

Table 5: Fatigue Defect Frequencies

Defect Shorthand	Long name	Frequency

FH	Flat Head	4288; 3.7%
HSH	Horizonal Split Head	10597; 9.1%
SD	Shelly Spots	7967; 6.8%
TDC	Compound Fissure	1111; 0.9%
TDD	Detail Fracture	71633; 61.8%
TDT	Transverse Fissure	4392; 3.8%
VSH	Vertical Split Head	15865; 13.7%

While the results, Figure 14, Figure 15, Figure 16, and Figure 17, were similar to previous analyses in that few segments showed reasonable results, the plotting of multiple defect types over one another, Figure 18, and in comparison to the combined case, provided some interesting thoughts. For one, that some defects, such as Shelly Spots, may only occur later on in the rail's life, but have a higher rate of defect accumulative, whereas previously strong defect types started to decrease in appearance. However, since the Weibull Parameters themselves were in question, further examination was not done.

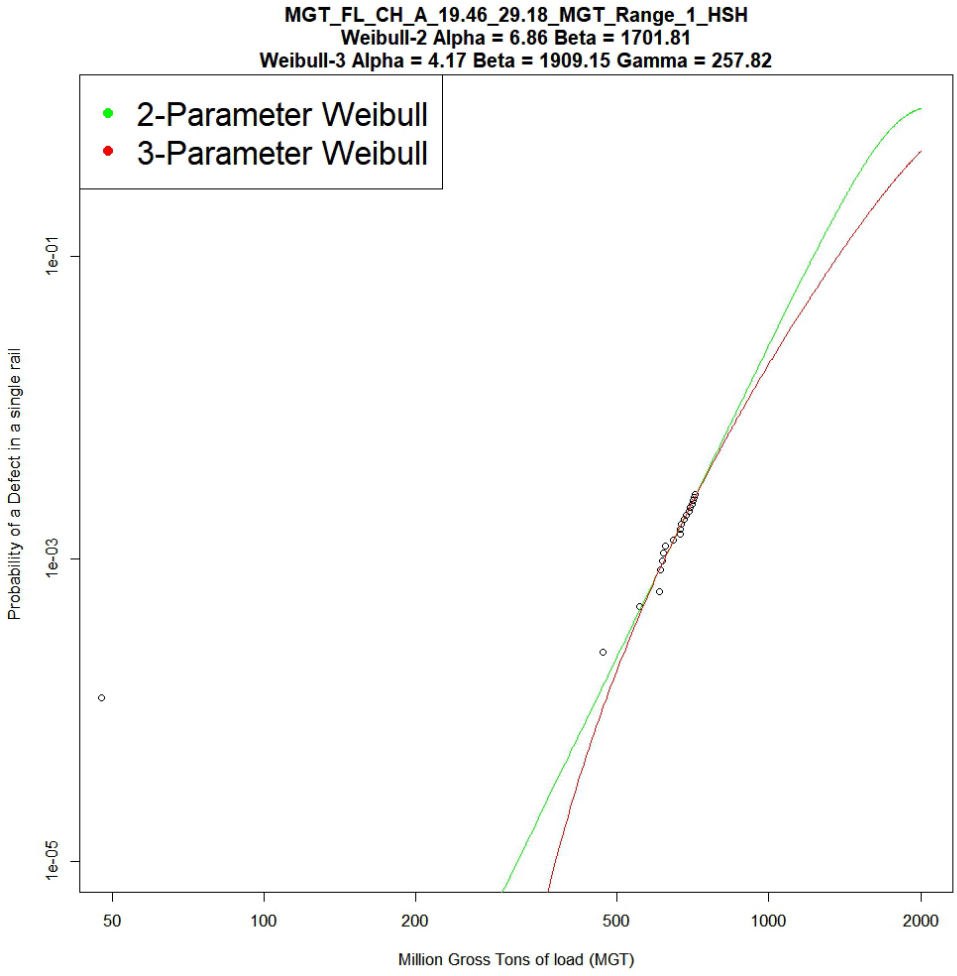


Figure 14: 2- and 3- Parameter Weibull fits for Horizontal Split Head defects in a designated rail

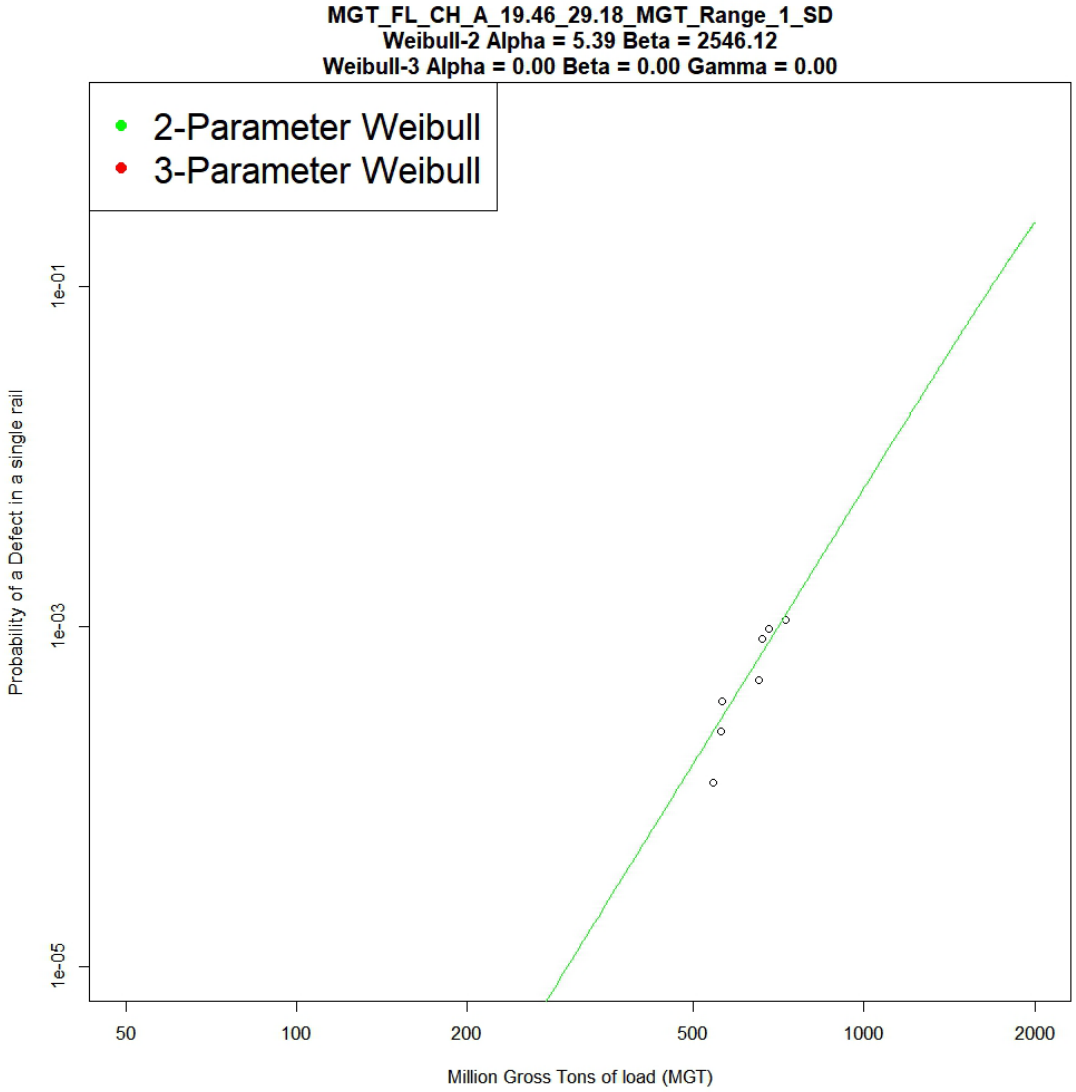


Figure 15: 2- and 3- Parameter Weibull fits for Shelly Spot defects in a designated rail

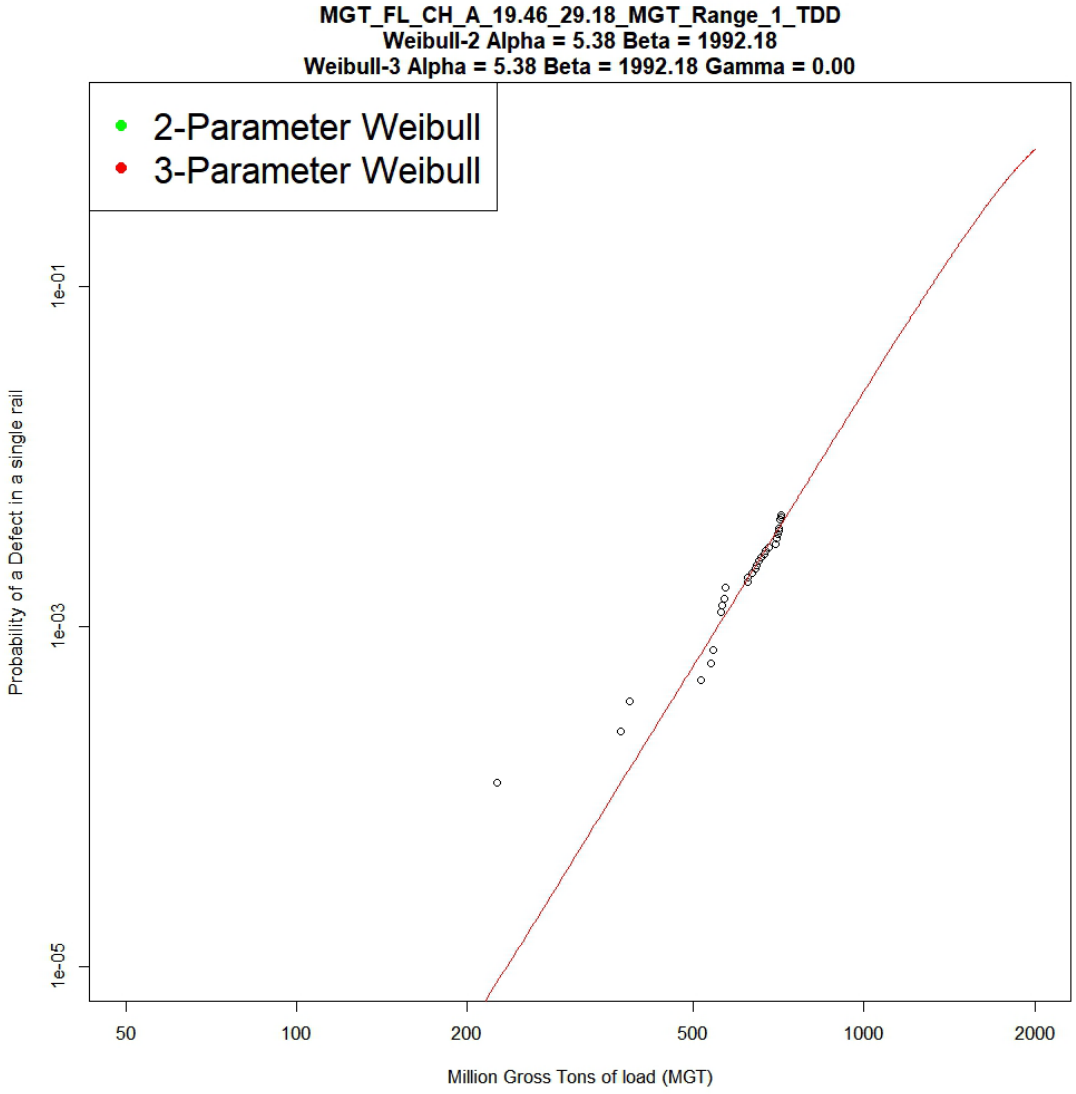


Figure 16: 2- and 3- Parameter Weibull fits for Detail Fracture defects in a designated rail

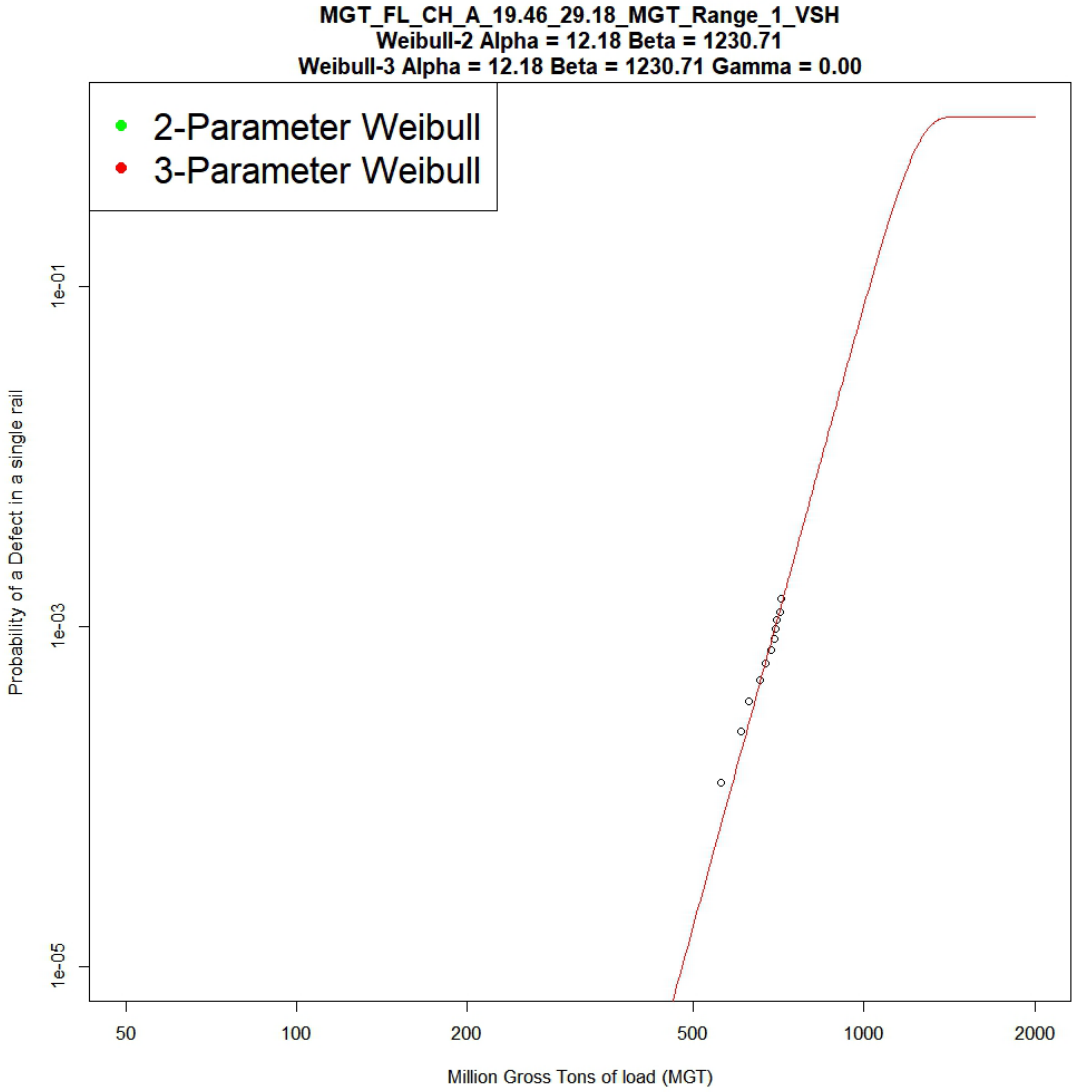


Figure 17: 2- and 3- Parameter Weibull fits for Vertical Split Head defects in a designated rail

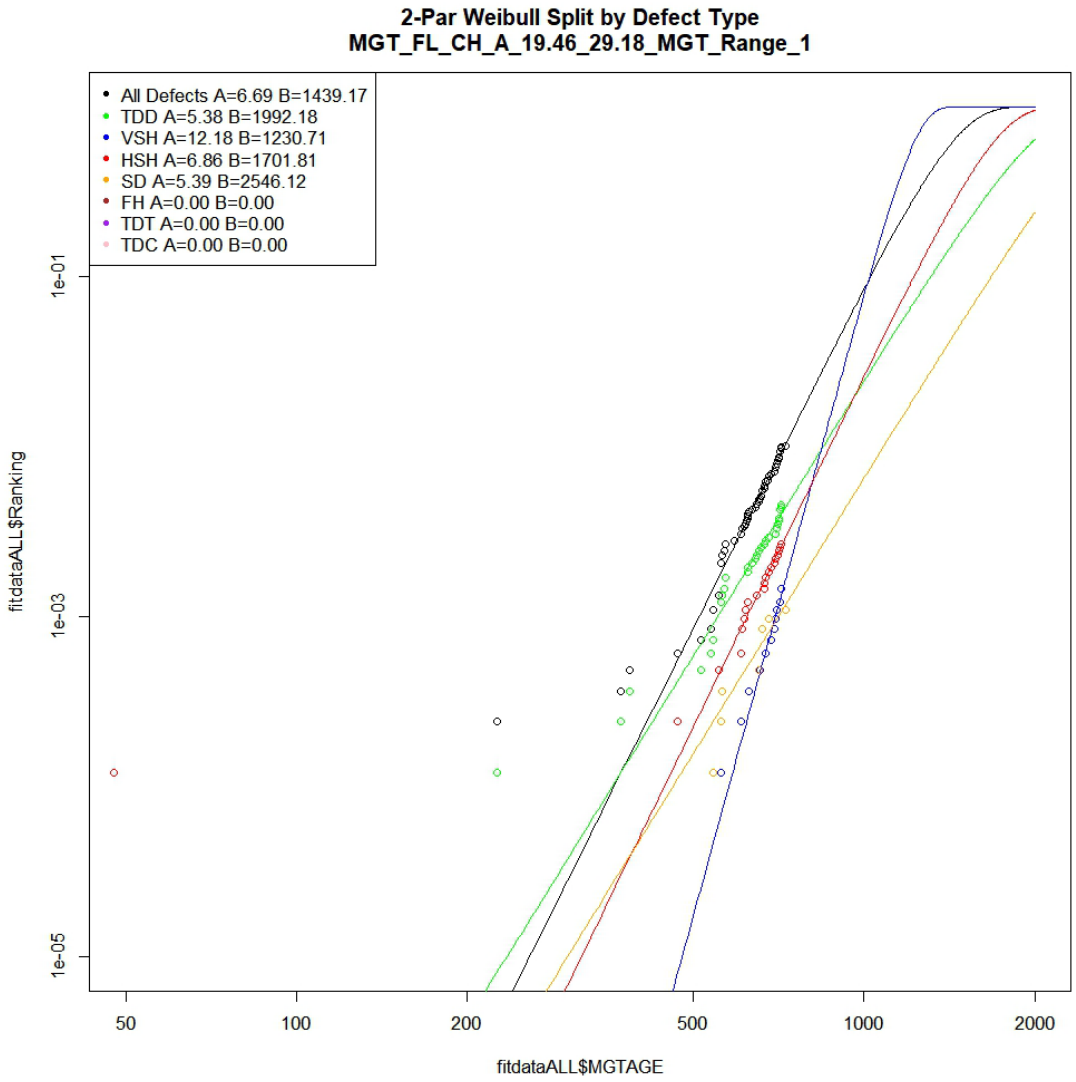


Figure 18: Overlaid 2-Parameter Weibull plots of all defect types on the same rail section

The next permutation of the analysis consisted of consolidating all track based on three broad categories: Annual MGT, Year Laid, and Curvature. As shown in Table 7 through Table 18, when the data was restricted to only "full history" track and defects, there were limited data points usable, and it was thought at the time that consolidating them along these three parameters would allow usable results.

Table 6: MGT-Year-Curve-Defect Count, Years 1999 to 2004

	1999	2000	2001	2002	2003	2004
MGT_0-5	18	6	6	9	19	2
MGT_5-10	11	16	11	9	1	0
MGT_10-15	6	17	5	21	8	19

MGT_15-20	14	17	11	27	4	3
MGT_20-25	6	5	10	43	12	25
MGT_25-30	24	18	58	59	19	17
MGT_30-35	6	13	14	46	12	23
MGT_35-40	2	3	4	24	8	10
MGT_40-45	15	5	7	18	10	6
MGT_45-50	12	10	12	23	15	15
MGT_50-55	13	1	4	14	2	3
MGT_55-60	0	0	17	5	1	16
MGT_60-65	5	3	3	3	7	2
MGT_65-70	0	3	0	4	3	1
MGT_70-75	0	0	0	0	2	0
MGT_75-80	0	0	0	0	0	0
MGT_80-85	0	0	0	0	0	0
MGT_85-90	2	8	0	0	0	4
MGT_90-95	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0

Table 7:MGT-Year-Curve-Defect Count, Years 2005 to 2010

	2005	2006	2007	2008	2009	2010
MGT_0-5	3	11	2	87	7	20
MGT_5-10	0	3	3	10	9	26
MGT_10-15	11	22	31	45	23	52
MGT_15-20	18	7	30	9	16	20
MGT_20-25	19	53	46	17	58	51
MGT_25-30	18	60	10	64	103	46
MGT_30-35	4	28	15	101	51	51
MGT_35-40	11	4	52	40	44	61
MGT_40-45	13	22	38	64	41	32
MGT_45-50	17	9	34	28	61	43
MGT_50-55	6	5	7	19	20	34
MGT_55-60	1	0	3	8	3	9
MGT_60-65	0	2	4	3	1	1
MGT_65-70	0	4	0	0	0	5
MGT_70-75	0	0	0	0	0	0
MGT_75-80	0	0	0	0	0	6
MGT_80-85	0	1	0	0	0	0
MGT_85-90	0	0	0	0	0	2
MGT_90-95	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0

Table 8: MGT-Year-Curve-Defect Count, Years 2011 to 2017

	2011	2012	2013	2014	2015	2016	2017
MGT_0-5	60	3	9	5	10	1	1
MGT_5-10	13	28	52	8	13	24	0
MGT_10-15	69	19	99	9	20	3	0
MGT_15-20	18	8	20	5	14	20	3
MGT_20-25	50	17	54	34	23	38	22
MGT_25-30	49	109	37	34	88	88	83
MGT_30-35	52	25	55	31	26	44	4
MGT_35-40	63	14	22	55	30	74	0
MGT_40-45	8	15	42	40	53	41	19
MGT_45-50	20	44	40	10	63	52	0
MGT_50-55	16	8	23	23	90	104	0
MGT_55-60	18	0	1	1	40	28	0
MGT_60-65	15	13	0	0	72	7	0
MGT_65-70	3	11	0	0	9	15	0
MGT_70-75	2	0	4	0	0	0	0
MGT_75-80	0	0	0	0	0	0	0
MGT_80-85	0	0	0	0	0	0	0
MGT_85-90	0	0	0	0	0	0	0
MGT_90-95	0	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0	0

Table 9: MGT-Year-Curve Weibull Alpha Values, Years 1999 to 2004

	1999	2000	2001	2002	2003	2004
MGT_0-5	0.58	1.44	2.74	0.00	10.22	0.00
MGT_5-10	0.00	4.74	3.78	3.25	0.00	0.00
MGT_10-15	1.54	2.92	1.39	4.67	0.00	6.59
MGT_15-20	0.00	4.47	6.64	4.79	0.00	0.00
MGT_20-25	8.29	2.68	1.36	9.92	1.18	2.09
MGT_25-30	5.30	2.35	4.46	2.90	5.93	0.00
MGT_30-35	0.00	6.51	2.11	3.42	1.65	2.70
MGT_35-40	0.00	0.00	3.58	2.13	3.88	2.35
MGT_40-45	2.10	0.00	0.00	2.52	0.00	0.00
MGT_45-50	0.00	1.77	0.00	2.81	0.00	1.60
MGT_50-55	0.00	0.00	0.00	0.00	0.00	0.00
MGT_55-60	0.00	0.00	0.00	0.00	0.00	0.00
MGT_60-65	0.00	0.00	0.00	0.00	0.00	0.00
MGT_65-70	0.00	0.00	0.00	1.05	0.00	0.00

MGT_70-75	0.00	0.00	0.00	0.00	0.00	0.00
MGT_75-80	0.00	0.00	0.00	0.00	0.00	0.00
MGT_80-85	0.00	0.00	0.00	0.00	0.00	0.00
MGT_85-90	0.00	1.99	0.00	0.00	0.00	0.00
MGT_90-95	0.00	0.00	0.00	0.00	0.00	0.00
MGT_95-100	0.00	0.00	0.00	0.00	0.00	0.00

Table 10: MGT-Year-Curve Weibull Alpha Values, Years 2005 to 2010

	2005	2006	2007	2008	2009	2010
MGT_0-5	0.00	0.00	0.00	4.57	0.00	0.86
MGT_5-10	0.00	0.00	1.70	8.68	6.83	4.52
MGT_10-15	0.00	5.95	5.44	6.15	5.87	3.41
MGT_15-20	1.58	0.00	2.89	0.00	3.45	3.15
MGT_20-25	3.91	4.74	3.31	10.89	3.09	5.00
MGT_25-30	4.01	3.91	2.84	5.00	2.84	9.26
MGT_30-35	0.00	4.78	2.47	2.66	4.50	4.47
MGT_35-40	3.08	0.00	3.06	0.00	3.55	0.00
MGT_40-45	11.56	3.26	3.54	0.00	2.51	2.60
MGT_45-50	4.48	2.83	0.00	3.51	3.38	0.00
MGT_50-55	0.00	0.00	2.08	0.00	3.41	0.00
MGT_55-60	0.00	0.00	5.11	1.16	10.03	0.92
MGT_60-65	0.00	0.00	0.00	0.00	0.00	0.00
MGT_65-70	0.00	1.37	0.00	0.00	0.00	0.91
MGT_70-75	0.00	0.00	0.00	0.00	0.00	0.00
MGT_75-80	0.00	0.00	0.00	0.00	0.00	0.00
MGT_80-85	0.00	0.00	0.00	0.00	0.00	0.00
MGT_85-90	0.00	0.00	0.00	0.00	0.00	0.00
MGT_90-95	0.00	0.00	0.00	0.00	0.00	0.00
MGT_95-100	0.00	0.00	0.00	0.00	0.00	0.00

Table 11: MGT-Year-Curve Weibull Alpha Values, Years 2011 to 2017

	2011	2012	2013	2014	2015	2016	2017
MGT_0-5	0.42	0.00	0.00	0.00	0.90	0.00	0.00
MGT_5-10	0.00	2.33	2.47	7.64	2.16	4.45	0.00
MGT_10-15	6.02	0.00	4.24	0.00	3.93	8.51	0.00
MGT_15-20	2.45	0.00	4.73	0.00	4.29	6.49	5.24
MGT_20-25	6.72	8.86	9.82	5.48	3.76	5.43	26.88
MGT_25-30	4.83	0.00	3.76	3.23	10.60	4.02	5.50
MGT_30-35	3.35	0.00	5.32	4.45	3.24	4.62	5.71
MGT_35-40	5.30	0.00	7.15	3.73	7.94	6.50	0.00

MGT_40-45	6.78	6.17	4.52	4.58	2.77	0.00	8.17
MGT_45-50	3.22	2.49	2.63	3.53	0.00	2.93	0.00
MGT_50-55	8.54	0.00	0.00	5.57	0.00	0.00	0.00
MGT_55-60	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGT_60-65	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGT_65-70	0.00	3.42	0.00	0.00	1.03	2.53	0.00
MGT_70-75	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGT_75-80	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGT_80-85	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGT_85-90	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGT_90-95	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGT_95-100	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 12: MGT-Year-Curve Weibull Beta Values, Years 1999 to 2004

	1999	2000	2001	2002	2003	2004
MGT_0-5	1861178	4471	385	0	59	0
MGT_5-10	0	738	945	1188	0	0
MGT_10-15	10416	2883	24108	1133	0	885
MGT_15-20	0	1898	1166	1588	0	0
MGT_20-25	1178	6939	55959	979	104909	7334
MGT_25-30	2258	8428	2634	3973	1890	0
MGT_30-35	0	1708	10965	2964	36035	5549
MGT_35-40	0	0	4635	9874	2204	10095
MGT_40-45	10045	0	0	8819	0	0
MGT_45-50	0	55136	0	8702	0	54722
MGT_50-55	0	0	0	0	0	0
MGT_55-60	0	0	0	0	0	0
MGT_60-65	0	0	0	0	0	0
MGT_65-70	0	0	0	125325	0	0
MGT_70-75	0	0	0	0	0	0
MGT_75-80	0	0	0	0	0	0
MGT_80-85	0	0	0	0	0	0
MGT_85-90	0	19182	0	0	0	0
MGT_90-95	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0

Table 13: MGT-Year-Curve Weibull Beta Values, Years 2005 to 2010

	2005	2006	2007	2008	2009	2010
MGT_0-5	0	0	0	301	0	21857
MGT_5-10	0	0	8506	414	470	685

MGT_10-15	0	839	852	835	838	1389
MGT_15-20	23826	0	3892	0	3075	3498
MGT_20-25	2044	1891	2962	986	3276	1583
MGT_25-30	2197	2463	5514	1882	3297	1177
MGT_30-35	0	2399	6872	4741	2538	2632
MGT_35-40	5224	0	4247	0	3309	0
MGT_40-45	1595	3733	4777	0	7668	8904
MGT_45-50	2818	7491	0	6237	4450	0
MGT_50-55	0	0	30024	0	5812	0
MGT_55-60	0	0	3417	163088	2200	1782368
MGT_60-65	0	0	0	0	0	0
MGT_65-70	0	22765	0	0	0	514156
MGT_70-75	0	0	0	0	0	0
MGT_75-80	0	0	0	0	0	0
MGT_80-85	0	0	0	0	0	0
MGT_85-90	0	0	0	0	0	0
MGT_90-95	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0

Table 14: MGT-Year-Curve Weibull Beta Values, Years 2011 to 2017

	2011	2012	2013	2014	2015	2016	2017
MGT_0-5	1056827	0	0	0	22754	0	0
MGT_5-10	0	2714	1593	530	3412	609	0
MGT_10-15	785	0	977	0	1206	679	0
MGT_15-20	7279	0	1631	0	1767	1035	979
MGT_20-25	1268	1199	1008	1689	2593	1503	606
MGT_25-30	2021	0	2636	3674	1125	2215	1456
MGT_30-35	3153	0	1883	2263	3889	2565	1733
MGT_35-40	2284	0	1958	3253	1814	2097	0
MGT_40-45	2075	2283	2344	2863	6262	0	1717
MGT_45-50	7077	7014	6937	7042	0	7734	0
MGT_50-55	2274	0	0	3122	0	0	0
MGT_55-60	0	0	0	0	0	0	0
MGT_60-65	0	0	0	0	0	0	0
MGT_65-70	0	6200	0	0	389911	12279	0
MGT_70-75	0	0	0	0	0	0	0
MGT_75-80	0	0	0	0	0	0	0
MGT_80-85	0	0	0	0	0	0	0
MGT_85-90	0	0	0	0	0	0	0
MGT_90-95	0	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0	0

Table 15: MGT-Year-Curve Rail Lengths, Years 1999 to 2004

	1999	2000	2001	2002	2003	2004
MGT_0-5	21.74	11.79	6.52	4.38	19.24	2.09
MGT_5-10	4.77	13.83	7.1	7.06	2.02	4.76
MGT_10-15	4.79	27.33	7.73	22.15	8.25	31.42
MGT_15-20	32.02	29.54	12.32	27.11	25.54	15.48
MGT_20-25	15.14	15.2	18.78	25	18.7	13.84
MGT_25-30	43.14	19.73	64.42	30.26	20.38	13.72
MGT_30-35	17.36	5.45	11.09	13.35	21.71	14.03
MGT_35-40	6.57	1.41	6.69	19.25	2.87	11.66
MGT_40-45	7.14	3.64	6.55	13.48	11.58	12.71
MGT_45-50	17.48	29.9	10.06	25.3	18.07	24.48
MGT_50-55	3.06	2.59	3.25	5.05	4.73	4.21
MGT_55-60	0.06	0.85	5.05	5.13	3.69	11.89
MGT_60-65	12.36	8.52	2.43	14.19	13.01	6.2
MGT_65-70	2.16	4.84	0.67	2.25	1.55	0.96
MGT_70-75	0.22	0.8	0.01	0.25	1.33	0.17
MGT_75-80	0.37	0.1	0	0.38	0	0.09
MGT_80-85	0.01	2.44	0	0	0	0
MGT_85-90	2.69	7.38	0	0	0.01	0.08
MGT_90-95	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0

Table 16: MGT-Year-Curve Rail Lengths, Years 2005 to 2010

	2005	2006	2007	2008	2009	2010
MGT_0-5	4.95	9.17	6.9	33.22	19.69	22.95
MGT_5-10	6.41	6.38	4.98	6.74	9.89	20.3
MGT_10-15	12.96	16.22	17.9	33.59	21.28	21.43
MGT_15-20	15.98	19.11	40.29	39.09	33.42	34.38
MGT_20-25	6.82	30.31	24.84	26.96	35.57	21.18
MGT_25-30	10.69	34.51	12.47	31.98	33.21	24.61
MGT_30-35	3.9	17.51	9.4	33.44	29.28	29.29
MGT_35-40	6.54	2.79	22.49	22.45	15.31	22.58
MGT_40-45	6.12	8.05	23.51	39.63	23.99	29.08
MGT_45-50	3.32	7.46	26.66	37.29	22.96	43.2
MGT_50-55	2.43	7.14	19.2	7.98	12.89	19.53
MGT_55-60	0.29	1.19	5.03	6.52	5.47	24.52
MGT_60-65	1.54	7.76	7.86	11.82	10.43	4.79
MGT_65-70	0.62	0.61	0.51	0.6	0.72	3.88

MGT_70-75	0.26	0.09	0.19	4.3	1.66	1.26
MGT_75-80	0.12	0.41	0.24	0	1.67	0.15
MGT_80-85	0	0.1	0	0	0.15	1.16
MGT_85-90	0	0.16	0.19	0	0	1.18
MGT_90-95	0	0	0	0	0	0
MGT_95-100	0	0	0	0	0	0

Table 17: MGT-Year-Curve Rail Lengths, Years 2011 to 2017

	2011	2012	2013	2014	2015	2016	2017
MGT_0-5	18.57	3.8	4.35	4.05	5.78	0.92	0.96
MGT_5-10	15.58	37.05	21.86	8.85	11.5	7.62	10.13
MGT_10-15	36.11	21.57	29.74	17.85	10.63	4.26	0.71
MGT_15-20	46.59	37.88	23.04	15.25	15.57	13.22	0.65
MGT_20-25	24.56	26.85	28.24	32.4	20.06	19.24	6.36
MGT_25-30	24.72	31	21	25.97	32.87	27.46	19.61
MGT_30-35	16.39	14.12	18.29	12.11	14.93	24.14	0.67
MGT_35-40	26.23	14.83	13.66	19	20.98	47.26	5.49
MGT_40-45	7.87	13.68	6.21	14.74	25.5	27.15	6.25
MGT_45-50	21.14	15.93	14.66	36.81	23.92	41.63	0.66
MGT_50-55	18.27	8.77	20.13	10.92	50.4	48.85	0
MGT_55-60	11.98	7.23	1.18	20.97	33.18	9.45	0
MGT_60-65	15.4	26.62	3.75	3.97	59.79	24.87	0
MGT_65-70	8.64	4.63	0.96	0.46	10.92	12.49	0
MGT_70-75	4.46	1.23	0.71	12.23	6.21	3.92	0
MGT_75-80	0.18	0.19	0.09	0.18	0.5	0.05	0
MGT_80-85	0.32	0	0	0	2.27	0	0
MGT_85-90	0	0	0	0	0	0	0
MGT_90-95	0	0	0	0	0	0	0
MGT_95-100	0	0.3	0.41	0	0	0	0

While some of these analyses provided results in the expected range, such as Figure 19, there were many which kept with the extreme values, such as Figure 20. For instance, if you look at the MGT 0~5 value for 2011 in all the previous tables, they have a reasonable number of defects and track length, 60 defects over 18.6 miles, but result in an Alpha of 0.42 and a Beta of 1,056,827 Cum. MGT, highly unrealistic values. This tended to happen when rail length was short, and there were few defects, making each defect that did occur to have more of an impact on the Weibull than when many more defects were found.

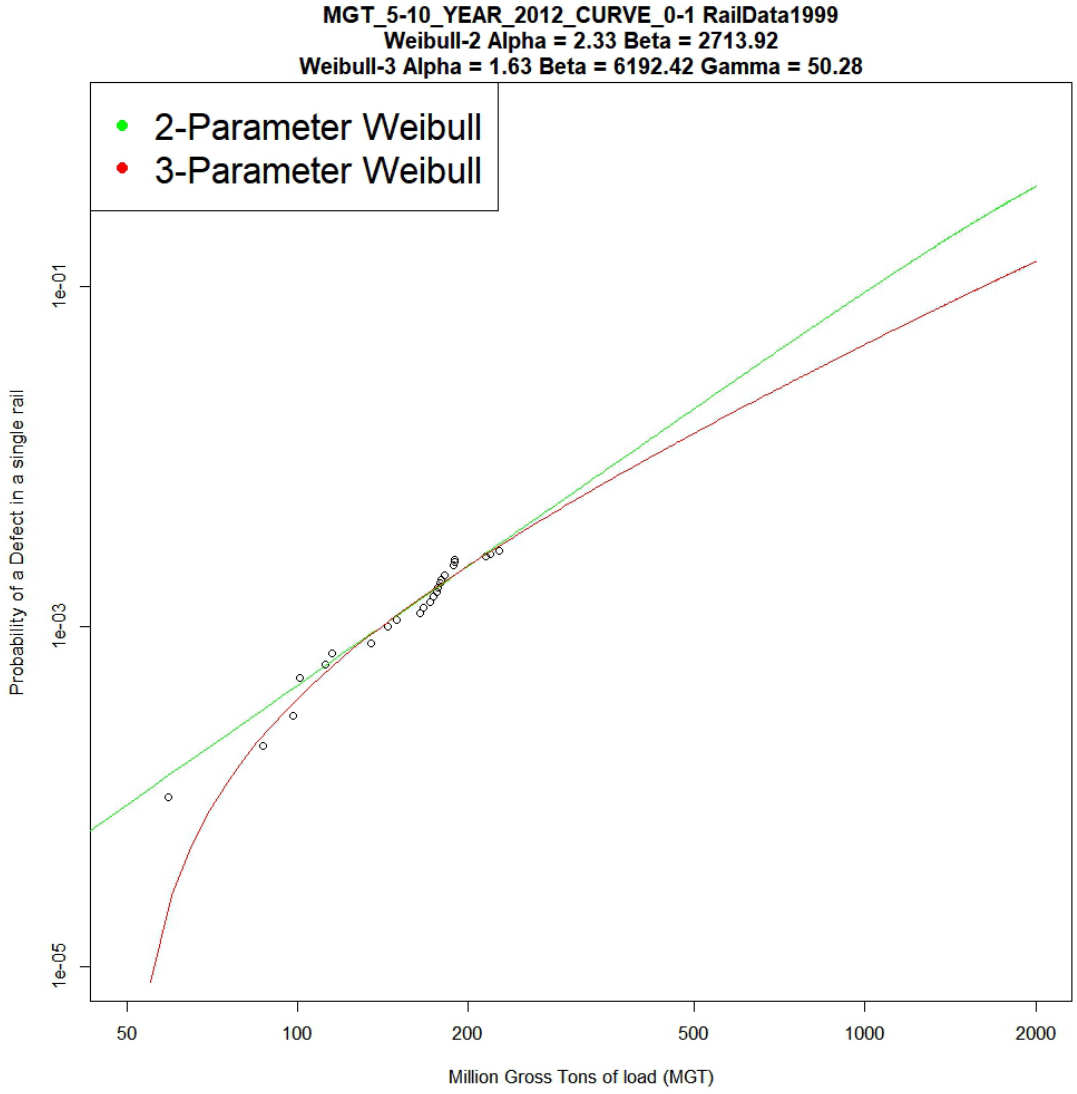


Figure 19: Weibull plot showing early-life defects

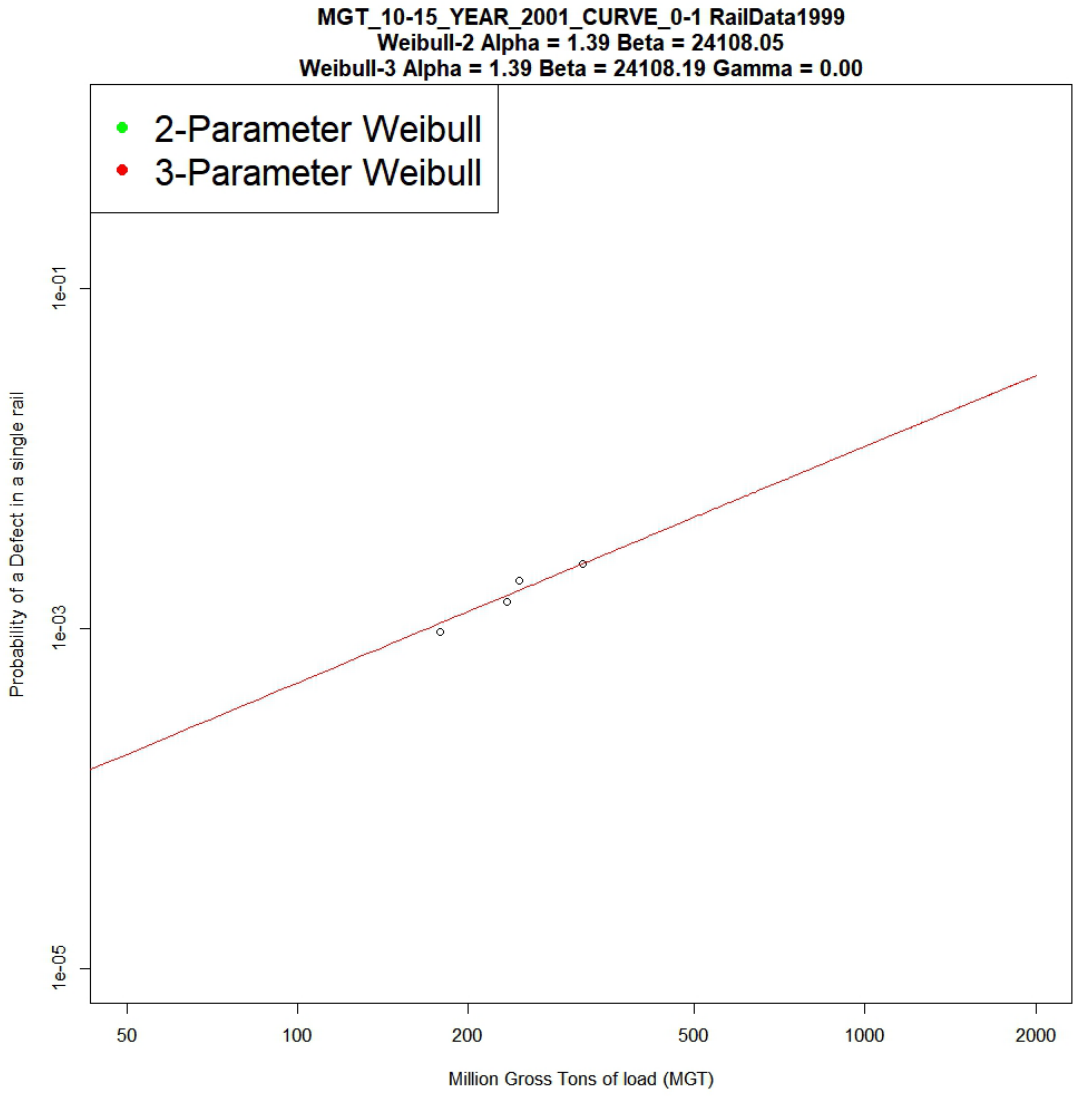


Figure 20: Weibull plot showing minimum number of defects

Following this analysis, the data itself was changed; instead of having the left and right track of a given segment be on the same line of data, they would be split into separate lines of data. Given that prior analyses did already account for the defects being on different sides of the track, this was primarily done to reduce the potential "no defect" track length, and hopefully bring the Weibull values more in line with the prior work ranges.

This separation of the rails was done by duplicating the data, and then filling in the unused-side's side-dependent data with a dummy value. Initially this value was chosen to be "9999", as it would be easy to spot errors, but later analysis would change this to "0" in order to simplify calculations which took into account this missing data. While this did change things up a bit, the results were generally in line with the previous analyses, leading to the same issues with extreme Weibull values.

3.1.3 Failings of the First Weibull Analysis

One of the biggest issues was the initial error with the Mean Ranking formula. While it was corrected before anything important had been done, it tainted the later results when they appeared in unexpected ranges. Instead of realizing that these values were actually indicative of the data, they tended to be looked at with skeptical views, and various alternative analyses were done to see if there was some missing step that was missing compared to previously done work by others. This resulted in significant delays in the research process.

Other issues were comparatively minor; determining what location identifier(s) would be used to match data, issues with the data structure in R, and determining appropriate fitting methods in R. These minor issues are common to almost all technical work, and should be expected whenever such tasks are undertaken for the first time.

3.1.4 Lessons learned from First Weibull Analysis

Most of the lessons learned from the First Weibull Analysis are related to learning and becoming proficient in the R programming language. Shifting from “for” loops to indexing for parameter replacement, thinking about how a process could be parallelized from the start, allowing multiple iterations to be run concurrently, and developing better methods of graphing the data.

3.1.5 Second Weibull Analysis

The second Weibull analysis started upon the acquisition of more data from the Class 1 Railroad. This data, as covered in Chapter 3, Section 2, doubled the number of years of defect data, as well as provided a better baseline for the track network to be used. The process used in the first Weibull analysis was used for bringing in, cleaning, and checking the data, but was expanded to account for the additional datasets. Beyond this, the analyses were effectively the same; variations of the dataset were used in developing Weibull plots and values that were within reasonable ranges. However, these results were either too low in number while in the reasonable range, or were far out of the reasonable range

Figure 21 and Figure 22 are representative of the plots developed before the analysis shifted to machine learning techniques in order to try and rectify the lack of reasonable results.

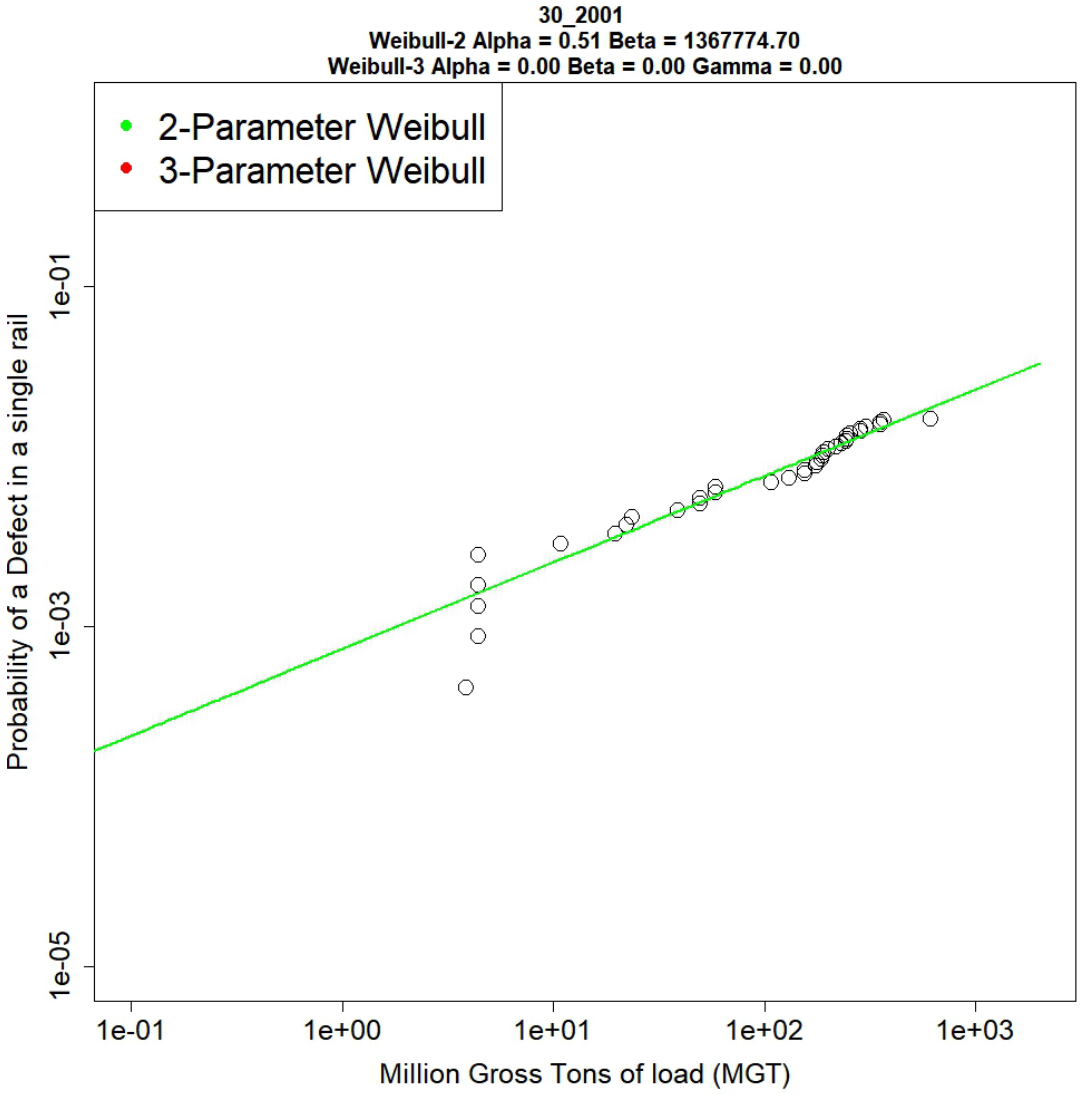


Figure 21: Weibull Analysis using Second set of data

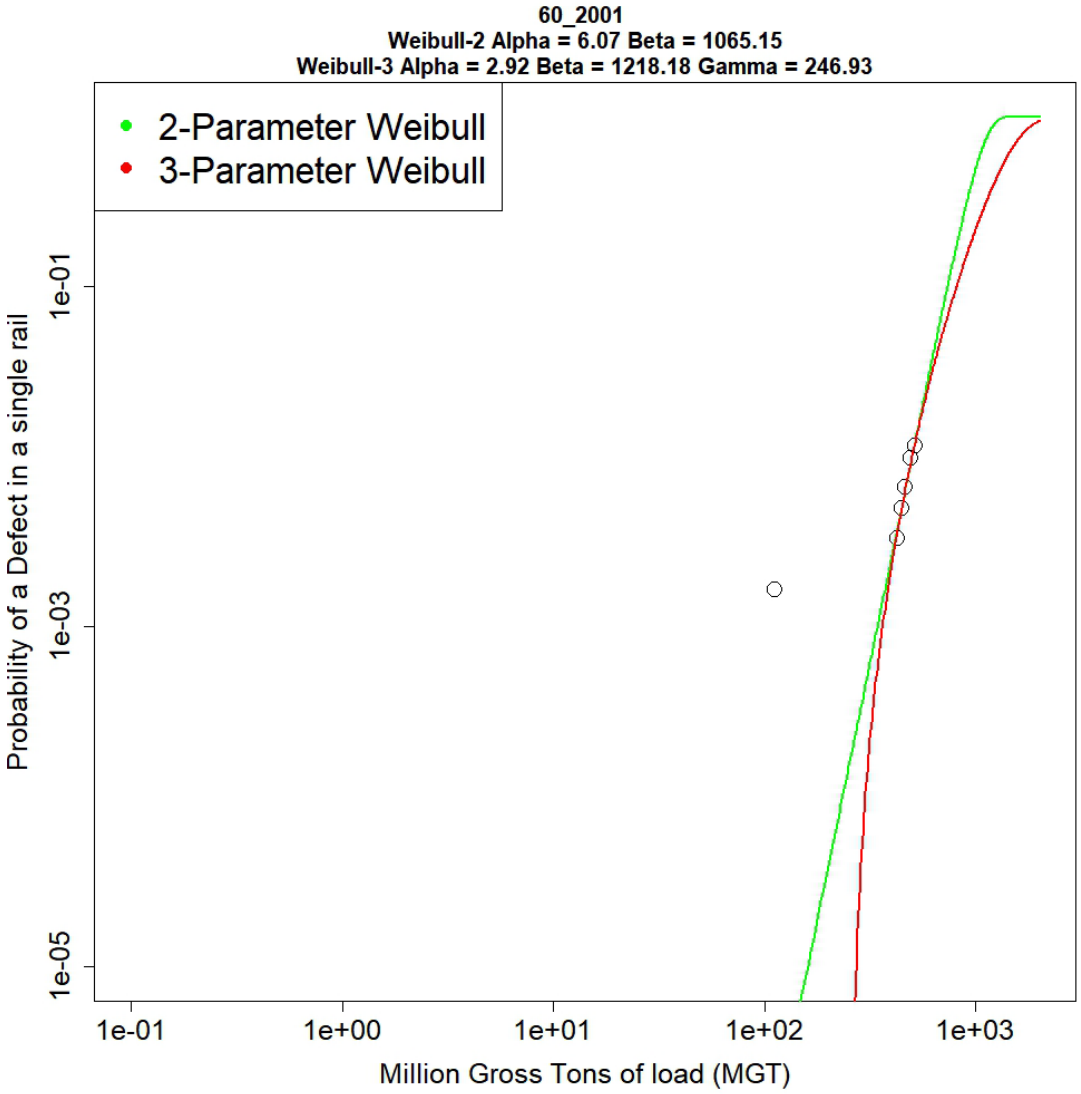


Figure 22: Weibull Plot using Second set of data

Following the acquisition of more data, it was decided to halt the previous analyses and restart from the beginning, as was done previously. This was more of a result of how the datasets were formed and put together, necessitating additions to be done at the beginning of the process. This was due to how the data was put together during the cleaning and merging process, which relied upon knowing as much about everything at that time. However, it should be possible to alter the methodology to account for new data without requiring the process to start from the beginning.

One bonus from this reset was that the iterative IF/THEN/ELSE checks for different variables was able to be changed to a method which relied upon the datasets themselves to pick out the correct indices to be changed. As shown in the following diagrams, Figure 23 and Figure 24, the process shifted from checking each individual row, into one which checked the column of interest, pulled out the indices of interest, and then used those to direct the replacement of data. This greatly increased the speed of the data cleaning, allowing the new data to be processed in less time than it

took for the first analysis' data; the tradeoff being that some debugging features, such as being able to quickly tell which row had bad data, and being able to process only a portion of the data at a time, are unable to be used.

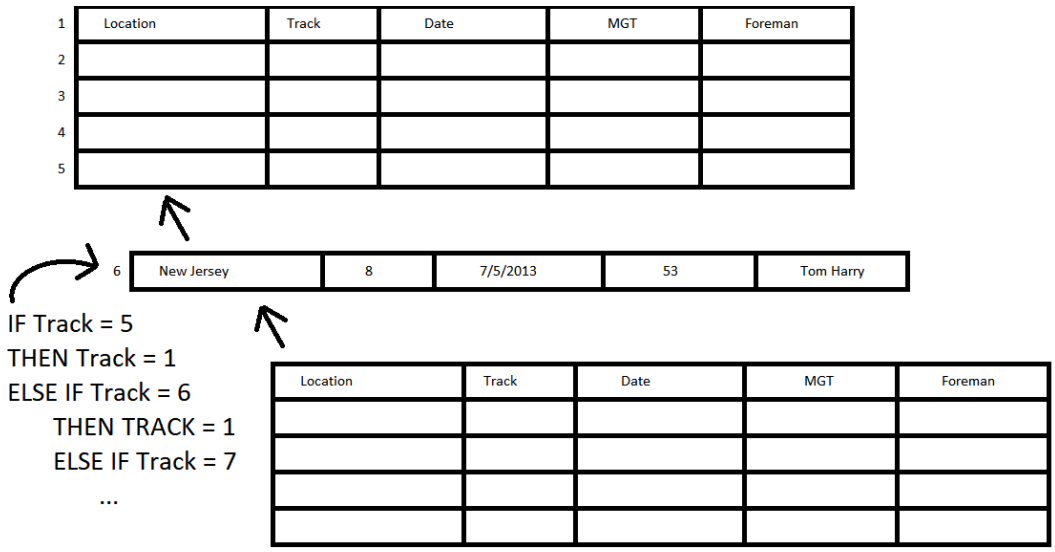


Figure 23: Representation of the Iterative Checking method; each row was analyzed individually

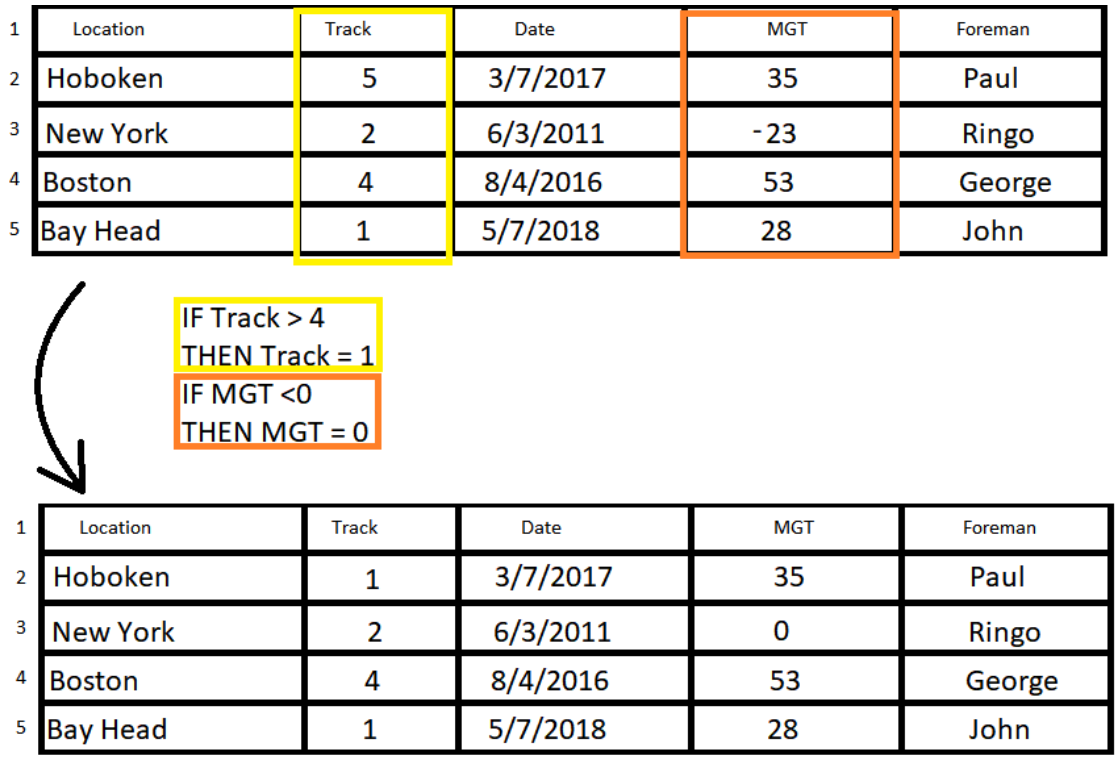


Figure 24: Representation of Index Checking methodology; each column is checked in the entirety at once.

3.1.6 Third Weibull Analysis

As before, there were attempts to perform the standard Weibull Analysis in order to develop a baseline for which future analyses could refer back to in order to determine accuracy. These graphs showed that, with the new data, the Weibull values fell more in line with the expected values, as well as occurring much more often; 1474 total graphs were produced with Weibull values within the expected range.

Part of this difference compared to previous analyses is that the Cumulative MGT was calculated differently. Instead of relying upon a formula in order to estimate the value, the 15 years of know Annual MGT was used as a reference point in order to scale the reported average Annual MGT as published by AREMA (55). This was done by taking the known historical data and interpolating it for the years not reported, as well as estimating the value based on the location (east or west), as there was a noticeable difference between the two pairs of major railways. The MGT values were then calculated based on the miles of track, and total ton-miles of transportation, allowing a derivation of the Annual MGT to be determined (see Figures 25, 26, 27).

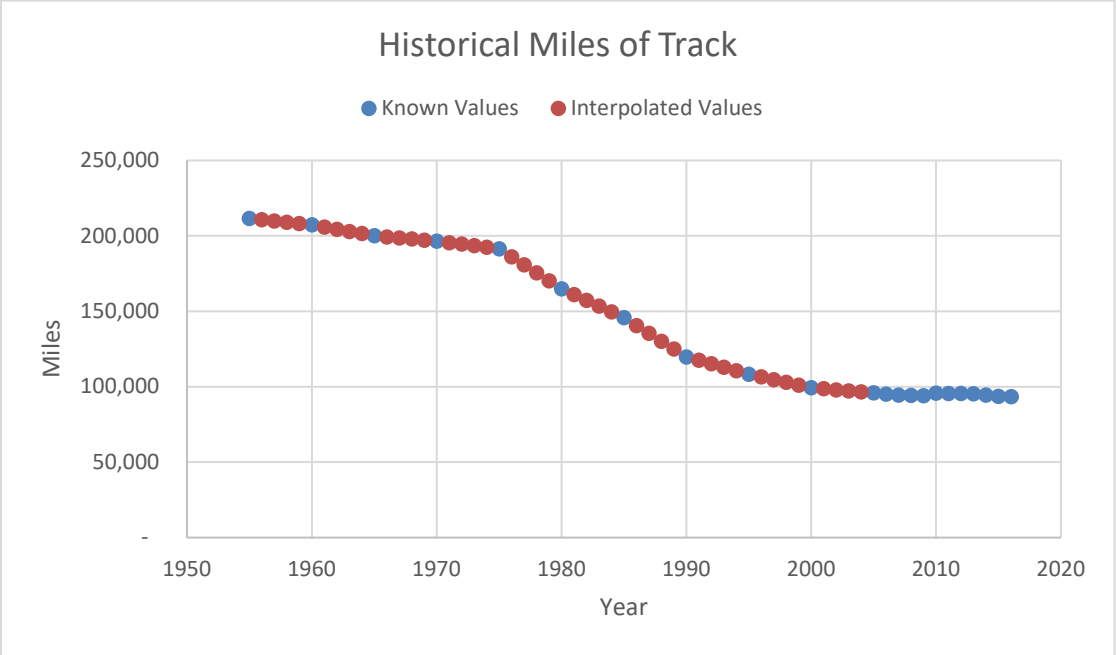


Figure 25: Known and Interpolated Historical Miles of Track

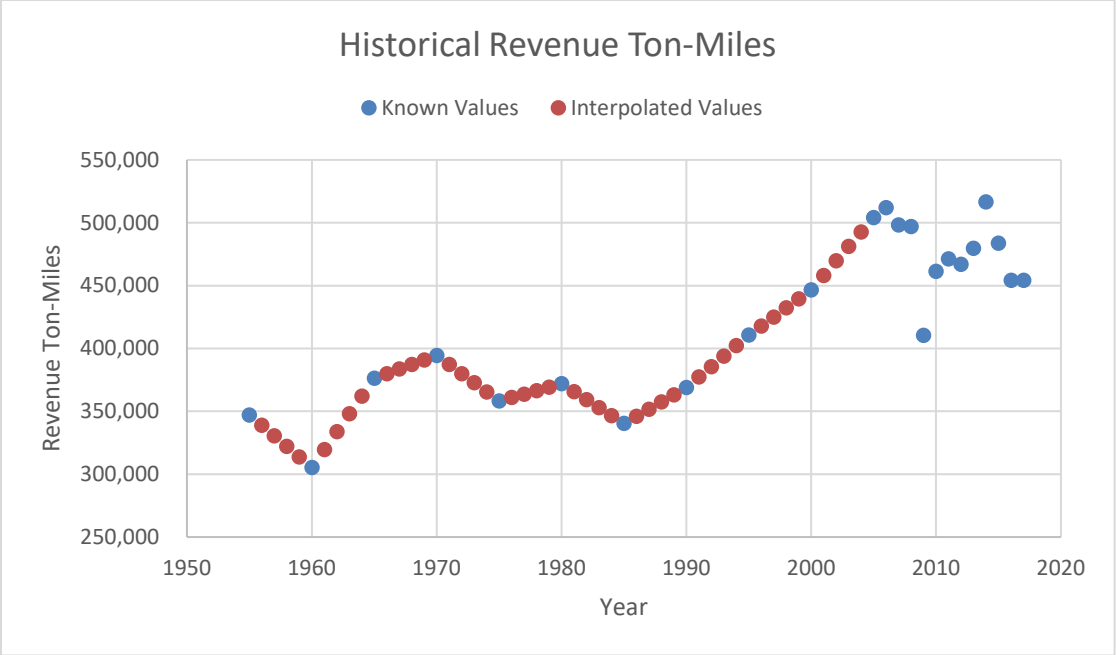


Figure 26: Known and Interpolated Historical Revenue Ton-Miles

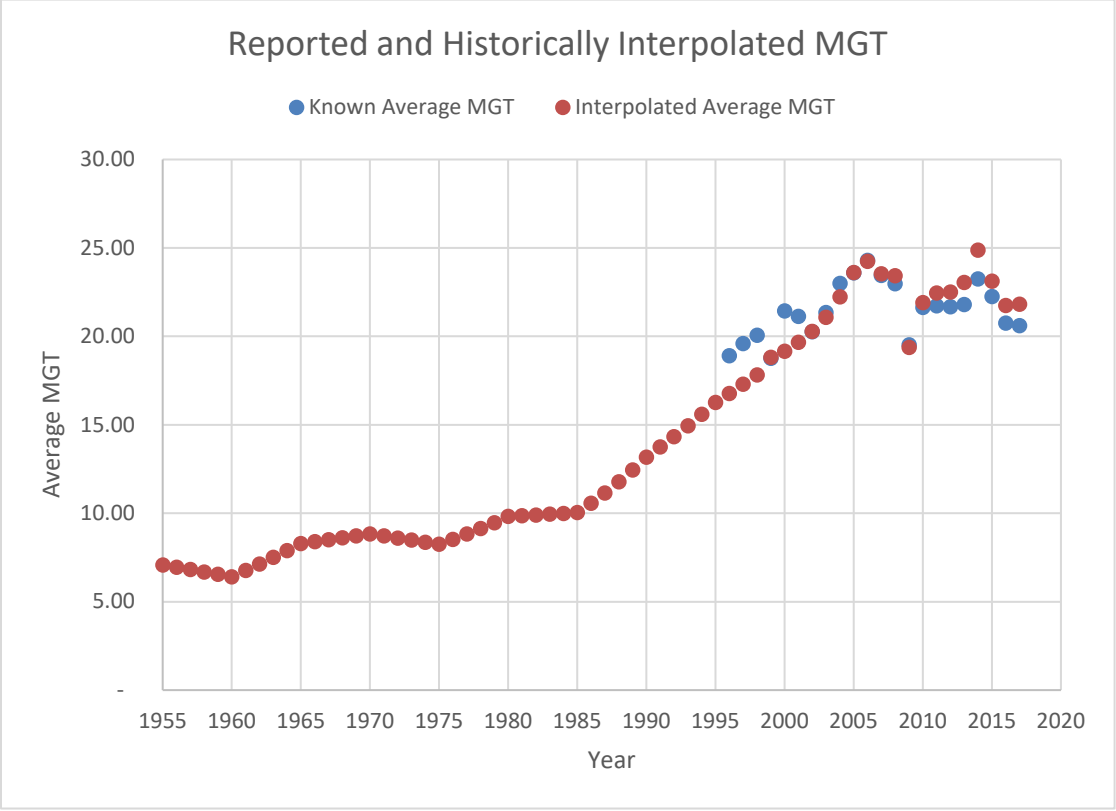


Figure 27: Known and Interpolated Historical Annual MGT

With these new Annual MGT values going back to 1955, the Cumulative MGT of track could be better estimated. Of course, there are several caveats to this method, such as assuming that the proportionality of traffic stays the same across the network. While this does not hold true in real life, due to changes in demands for certain products, the gain in reasonable values compared to the previous method of estimating Cumulative MGT deemed it to be of acceptable use.

3.2 Bootstrapping Weibull Analysis

As part of the previous KNN and Logistic Regression analyses, datapoints were over-represented by repeat picking of them in certain tests. This led to some of the issues observed in the results of the previous sections. A similar process, which is designed to be used, where there are limited data points is the overall process of Bootstrapping which develops a representation of the overall population of data, by randomly re-picking points.

However, normal Bootstrapping uses the exact values picked, meaning that if "1.34" and "2.45" were the only two values, then only combinations of those two would be results. This would severe limit the analyses used here, given that many track segments have few defect data points. Given the nature of the data, continuous positive values, alternatives were sought in an effort to better represent the population data. This led to the use of Parametric Bootstrapping for the Weibull Analyses and prediction efforts.

3.2.1 Parametric Bootstrapping Weibull Analysis

Parametric Bootstrapping is different from normal bootstrapping in that the known values are used to develop distributions from which new values are picked from (56). This greatly increases the population from which "re-picked" data is chosen. Furthermore, this lets the values chosen exceed the known values by treating the value as a "mean" and allowing a distribution of points on either side of the mean. This in turn, can help give insight into what the population distribution is.

To start, the analysis used the basic Weibull equation to come up with known Alpha and Beta values for each segment in the population segments selected. These segments were chosen at set intervals of 50 to 500 rows from the last chosen in order to ensure that limited duplication of analyses would result. In addition, the reduction from the initial population of approximately 200,000 segments to a few thousand actually being calculated provided far more reasonable computation times. These Alpha and Beta values formed the core of the Bootstrapping parametric bootstrapping models and were examined at this point to see how they were distributed in order to get a better understanding of how the population of values may be distributed. Figure 28 shows the distribution of Alpha values, and Figure 29 shows the distribution of Beta values that were found.

Initially, a Gamma distribution was used to fit a density function to the data, because the Gamma distribution had similarities with the overall distribution of the Alpha and Beta parameters. However, while working with the Gamma distribution, errors occurred due to the scaling of the Gamma's scale parameter, which was very sensitive to the magnitude of the Weibull Beta value. Based on the types of errors, it was determined that since the Beta values had a far greater range, any distribution chosen would have to work with this range. After several tries with various

distributions, including Chi-Squared, Weibull, and Exponential distributions, it was determined that the Log-Normal distribution would work best overall. The resulting Log-Normal distributions are presented, superimposed on the Alpha and Beta distributions in Figure 28 and Figure 29.

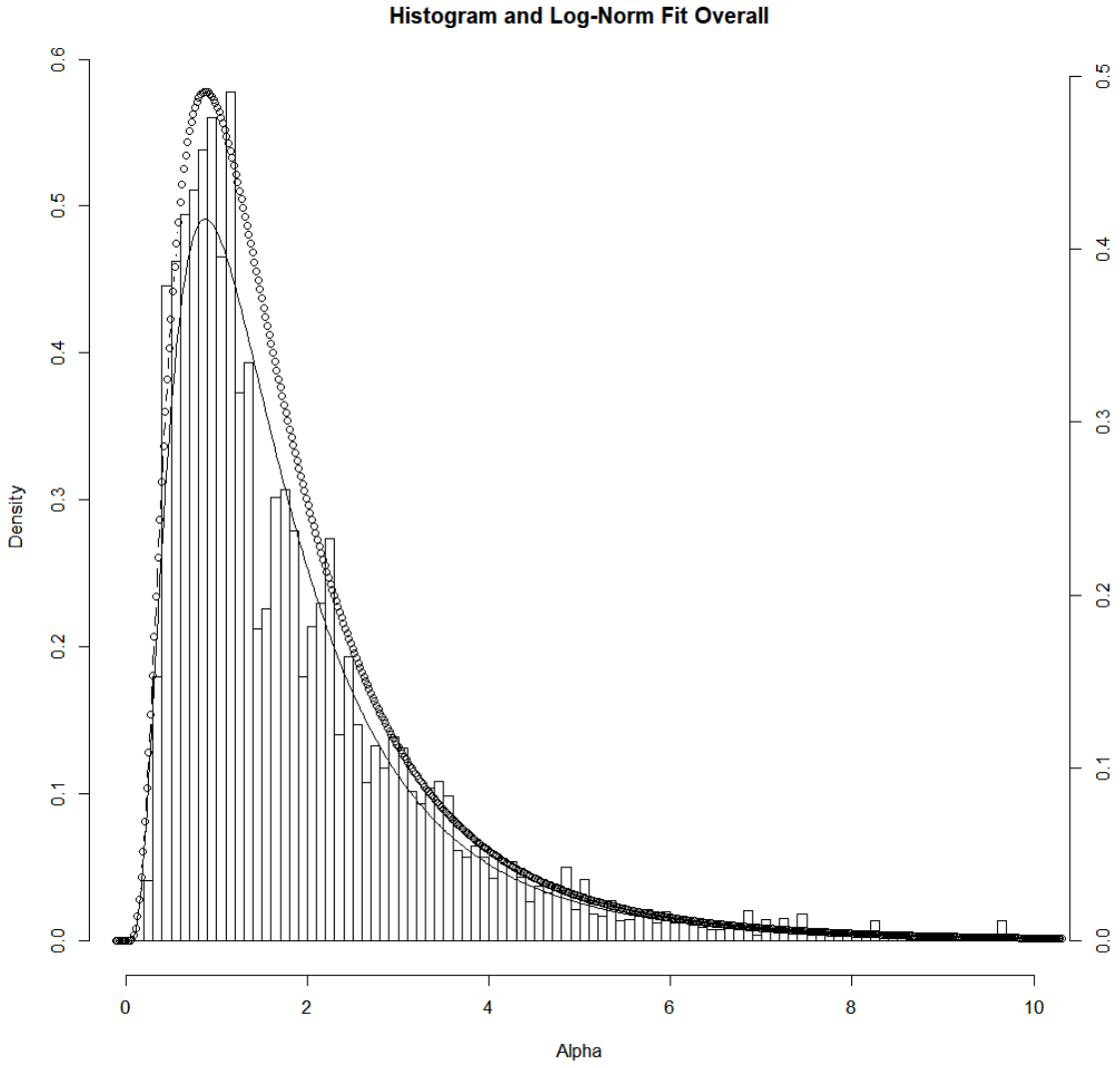


Figure 28: Overall Distribution of the Alpha Values

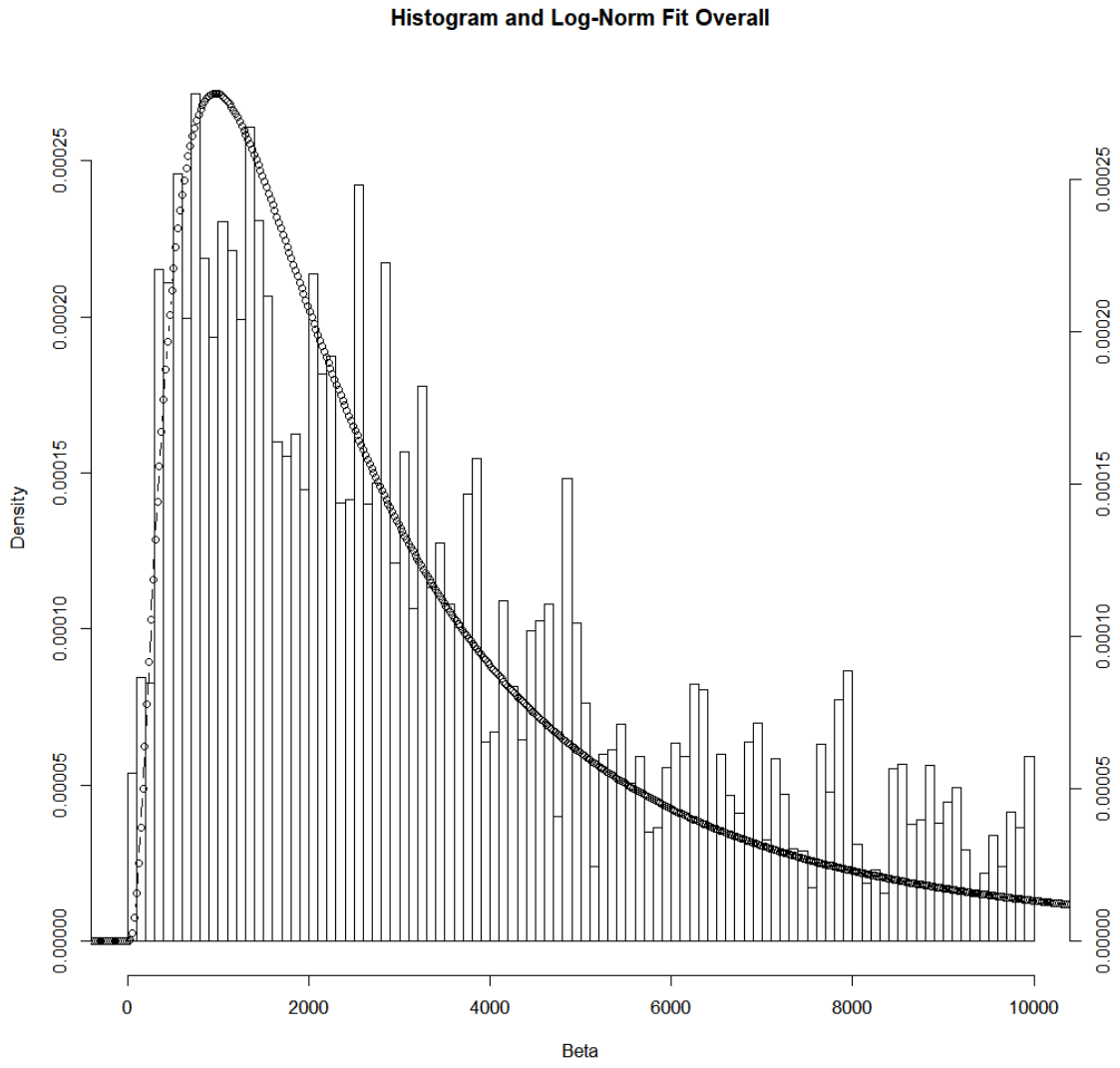


Figure 29: Overall Distribution of Beta Values

The next process was to select the meaningful source data for each given track segment. In this manner, the population of segments is analyzed to identify those that can generate Weibull Alpha and Beta values Bootstrapping relies upon some form of prior known data in order to extrapolate the distribution for the object being looked at, which in this case required the more limited number of track segments that were able to compute a Weibull Alpha and Beta to be used as the source data. As using the entirety of the data would render any results meaningless, since it shows the entire population acting around the 'average" the source data was pruned based on "similarity" to the track segment being looked at. Thus, the segments were grouped into "similar" segment groupings. In order to determine what track segments are considered "similar" enough to the segment in question, several parameters were used in pruning the dataset into the "similar" dataset. In determining what constituted a "similar" track segment, the decision was made to use user-determined values, based on the sensitivity of the Weibull results to these values. This is instead of using other algorithms to select the "similar" segments; such algorithms including K-Nearest

Neighbors or Clustering. This allows for the bypass of the issues that cropped up in previous machine learning attempts, and also allowed for ease of change in variables and their distribution. The selected parameters are Annual MGT, within +/-10%, Track Length, within +/- 10%, and Laid Year, within +/-5 years. For the very first few attempts, this pruning was not done in order to see how the data acted without any bounds, but the results were not useful since the results tended to group around the "mean" or "average" values. Pruning was then applied for all following analyses, with more effective results.

Once this source data was obtained, it was used to develop Alpha and Beta distributions, as shown in Figure 30 and Figure 31. These graphs formed the basis of the Parametric Bootstrapping, as they provided the distributions from which values could be obtained for the bootstrapping analysis.

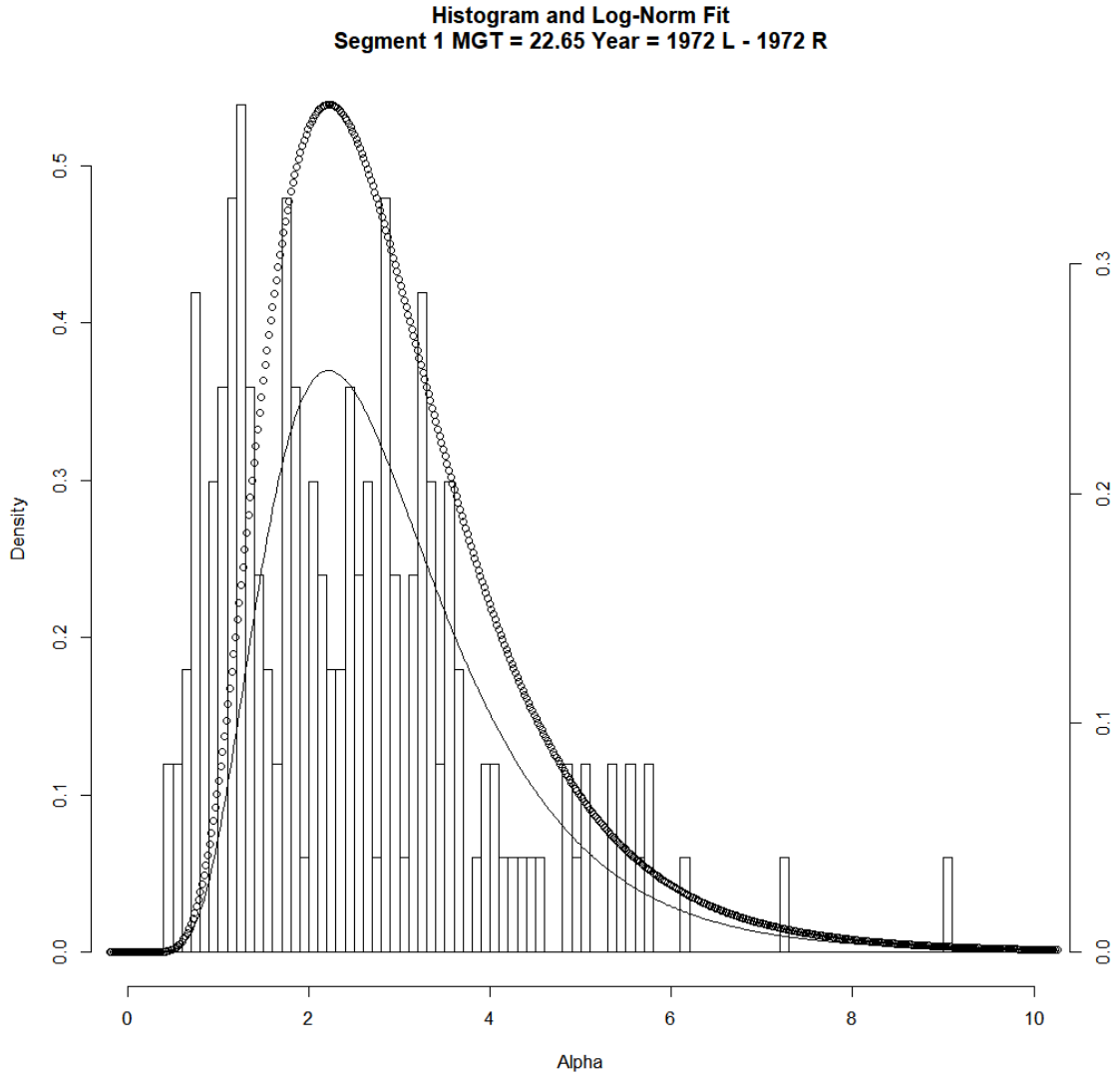


Figure 30: Histogram and Log-Normal fit of Alpha Values for Segment 1 The dotted line is the density fit scaled on the left axis, and the solid line is the Log-Normal distribution, scaled on the right axis

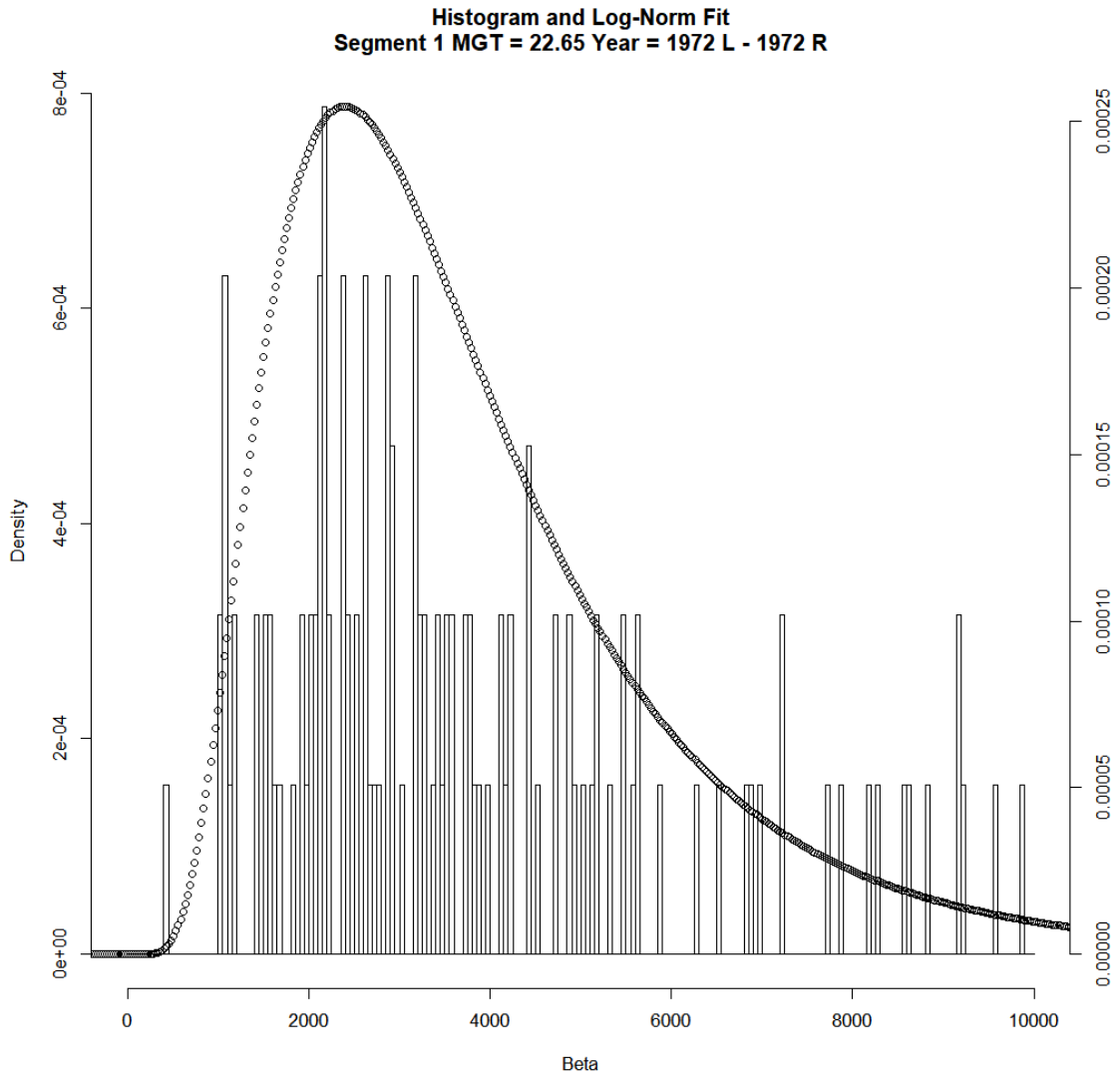


Figure 31: Histogram and Log-Normal fit of Beta values for Segment 1

Given these two distributions, a list of pairs of Alpha and Beta values were created, as shown in a small sample in Table 21. These pairs were then used to develop a Weibull curve for every entry, which would then be used in the calculations to determine probability densities. Initially, these pairs were not restricted in value in any way, which led to issues where extreme values would occur, throwing off the Weibull plotting, as shown in Figure 32. This figure shows the combined output of the Parametric Bootstrapping Analysis, as plotted on a Weibull graph. The Blue vertical line is the Cumulative MGT of the last known defect, the Green line is the best-fit Weibull values for the known defects. The Dark Red line is the best fit median Weibull equation, the dashed red lines are +/- 40% of the median value, and the Purple lines are the minimum/maximum values. While not as clearly shown as in later figures, the gray lines represent each unique Weibull pair that was used in the bootstrapping analysis.

In order to rectify the issues that resulted from the use of the extreme values, the Weibull parameters were limited to 0 to 10 for Alpha, and 500 to 10,000 for Beta, which represents a range

of realistic Weibull parameters based on literature review and analyses preformed previously. This restriction was done in two areas; first when the source data was being chosen, and again after the Alpha and Beta distributions were developed, and pairs were chosen. The pairs were checked for fitting within the boundaries, and if they were found to be outside, the pair was discarded and a new pair picked. In Table 21, Pairs #2 and #4 would be removed from the table, and then new values calculated as replacements.

Table 18:Example of Weibull Alpha and Beta pairs

ID	Alpha	Beta
1	2.42	1486
2	11.57	1843
3	4.22	5047
4	3.61	18762
5	2.59	3425

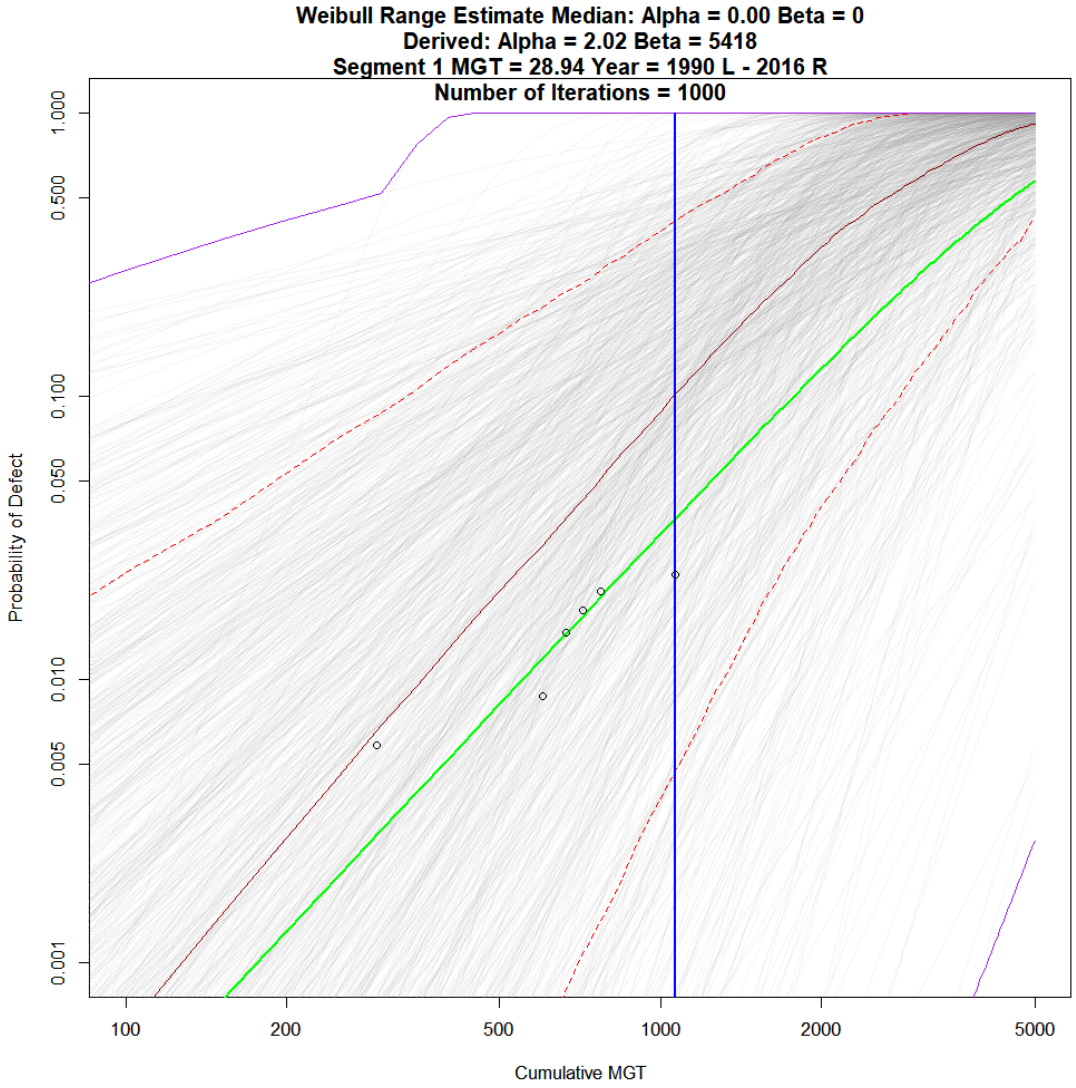


Figure 32: Unrestricted/unpruned initial parameter Weibull Bootstrapping results

Figure 32 shows this first Weibull Bootstrapping result. Since this segment had defects known prior to the bootstrapping attempt, they were included in the plotting of the graph to give a representation of what the estimated vs actual looked like. They are the small circles around the green line, which as noted is the best-fit Weibull values for the known defects.

Figure 33 is a similar plot, where the extreme values are not used in the boot strapping and the range of alpha and beta values are confined to the realistic range noted above. Thus, Figure 33 shows the distribution around the "median" Weibull values for a range of cumulative life (cumulative MGT) of 100 to 5000. Compared to Figure 32, the biggest difference is the loss of the Weibull plots in the lower right triangle area, due to Beta being restricted. There are also a few low Alpha Weibulls that were removed, as evident in the upper extreme boundary being lower than before.

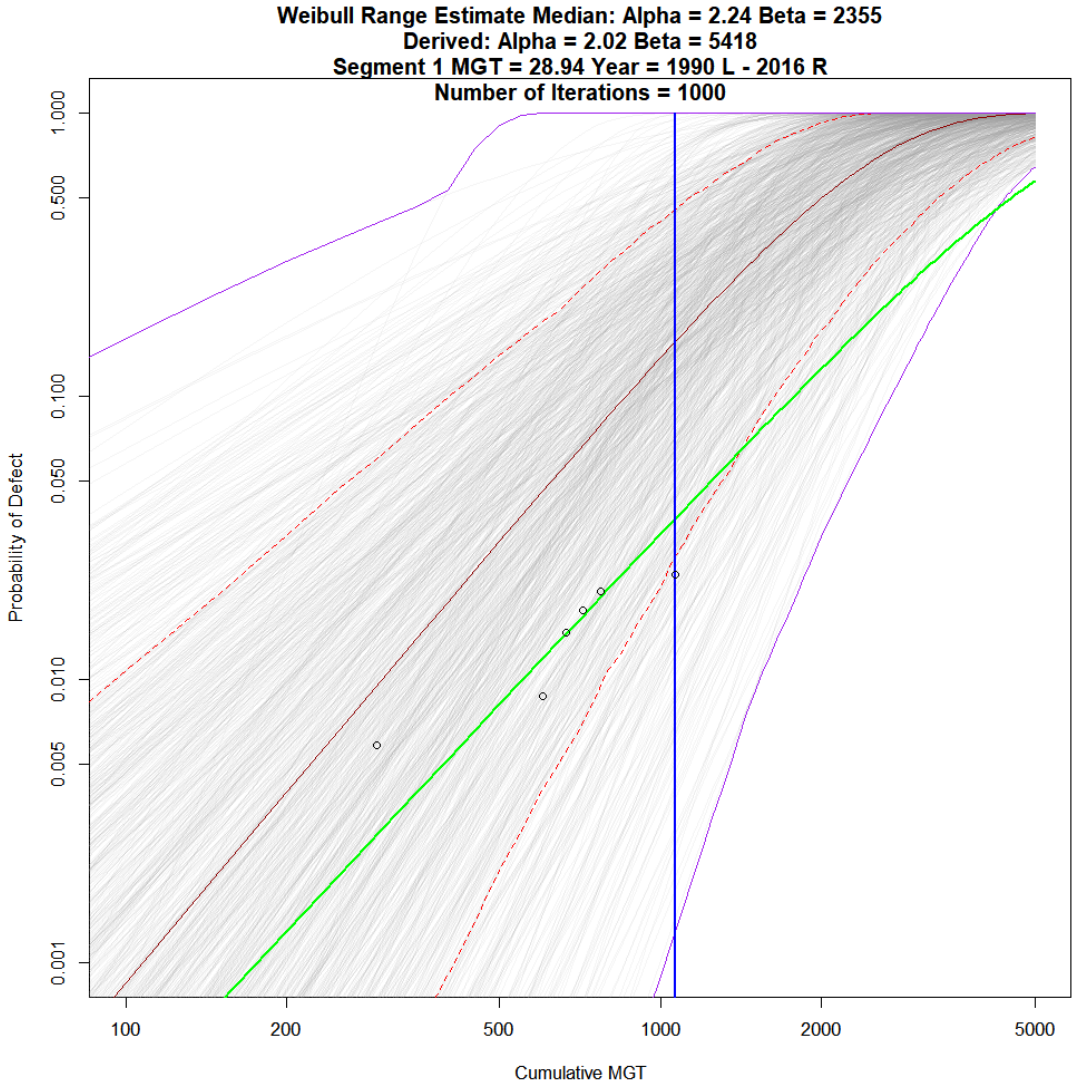


Figure 33: Restricted Initial Parameter Bootstrapped Weibull

Figure 34 shows a Bootstrapped Weibull with the source data plotted outside of the expected area. As can be seen, there is a large discrepancy between the median and best fit values, due to the existing data being discarded due to being outside the expected range of values. This is one consequence of the use of pruning measures to remove valid datapoints which are outside the expected limits of the data. In this case, the datapoint has a Beta value of over 16,000, 60% over the upper limit put in place for Beta values.

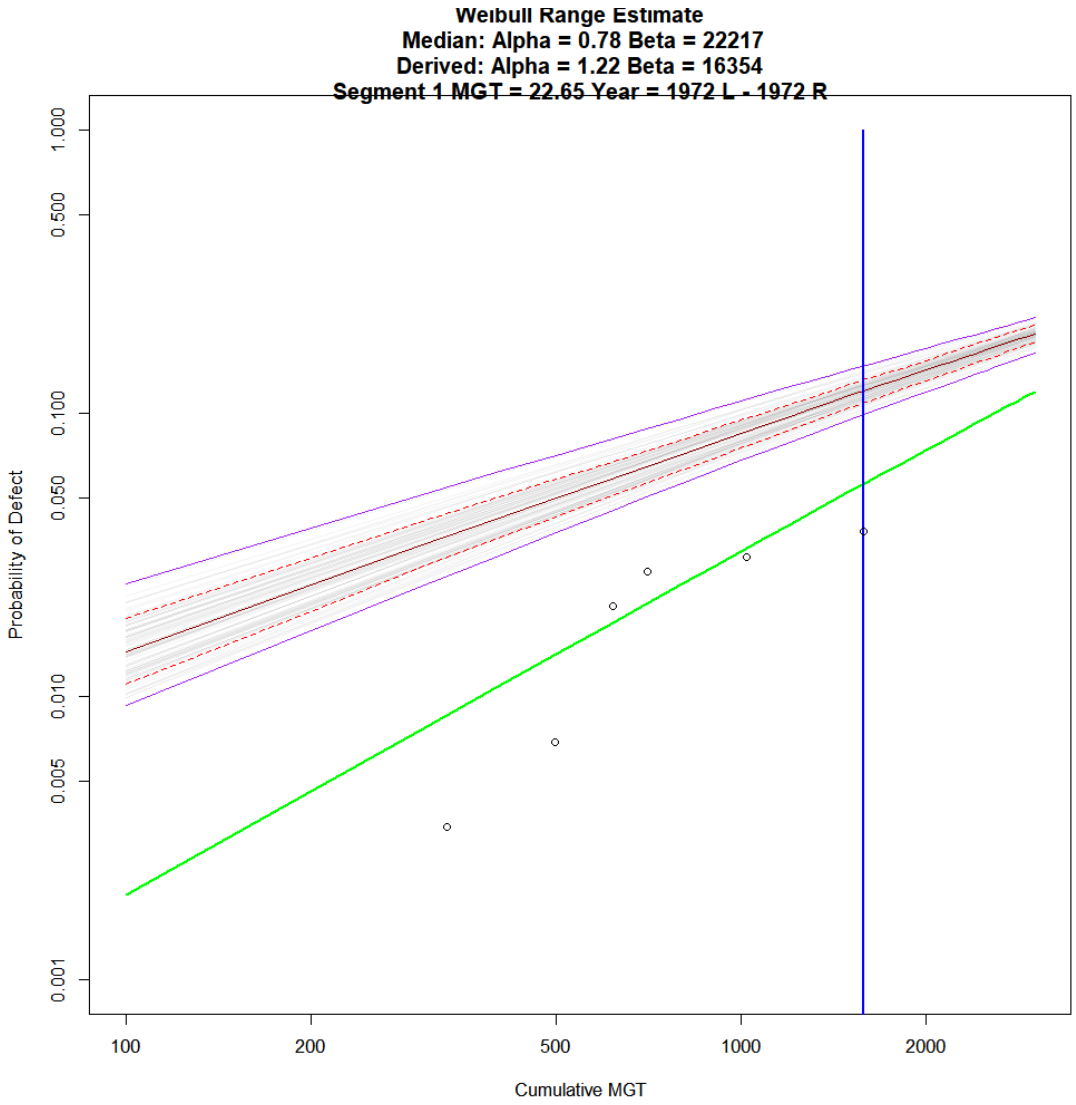


Figure 34: Weibull Bootstrapping results showing higher expected parameters vs known current

While this showed that the process of bootstrapping was working with the code as written, there was still more to do in order to improve the results. First to be done was expanding the number of times values were chosen ("re-picked") from the Alpha and Beta distributions, thus changing the number of Weibull lines plotted and used for calculations. Figure 35 (100 iterations) , Figure 36 (200 iterations), Figure 37 (300 iterations), Figure 38 (400 iterations), Figure 39 (500 iterations), Figure 40 (600 iterations), Figure 41 (700 iterations), Figure 42 (800 iterations), Figure 43 (900 iterations), and Figure 44 (1000 iterations), show the differences as the number of used Weibull Pairs increases. Of note is that there is a slight change in the expected median Weibull values, but a major difference in the smoothness of the resulting graphs. This smoothness helps interpretation of the results of the following applications to the data, while not impacting the values found. Examining the effect of the number of iterations to determine the relationship between smoothness and calculation time, the value of 1000 iterations was chosen as a value for all future Parametric Bootstrapped Weibull calculations.

Note, these figures represent a final iteration of the analysis process, where the process was applied correctly and reasonably. Prior attempts had various issues that were worked through and addressed, such as code errors that led to variables being the same for multiple runs, leading to the same segments being reused regardless of their applicability). Another such issue related to use of data without any boundary on the upper or lower values taken from the Alpha and Beta distributions, allowing a very wide range of predictions, as shown in Figure 35.

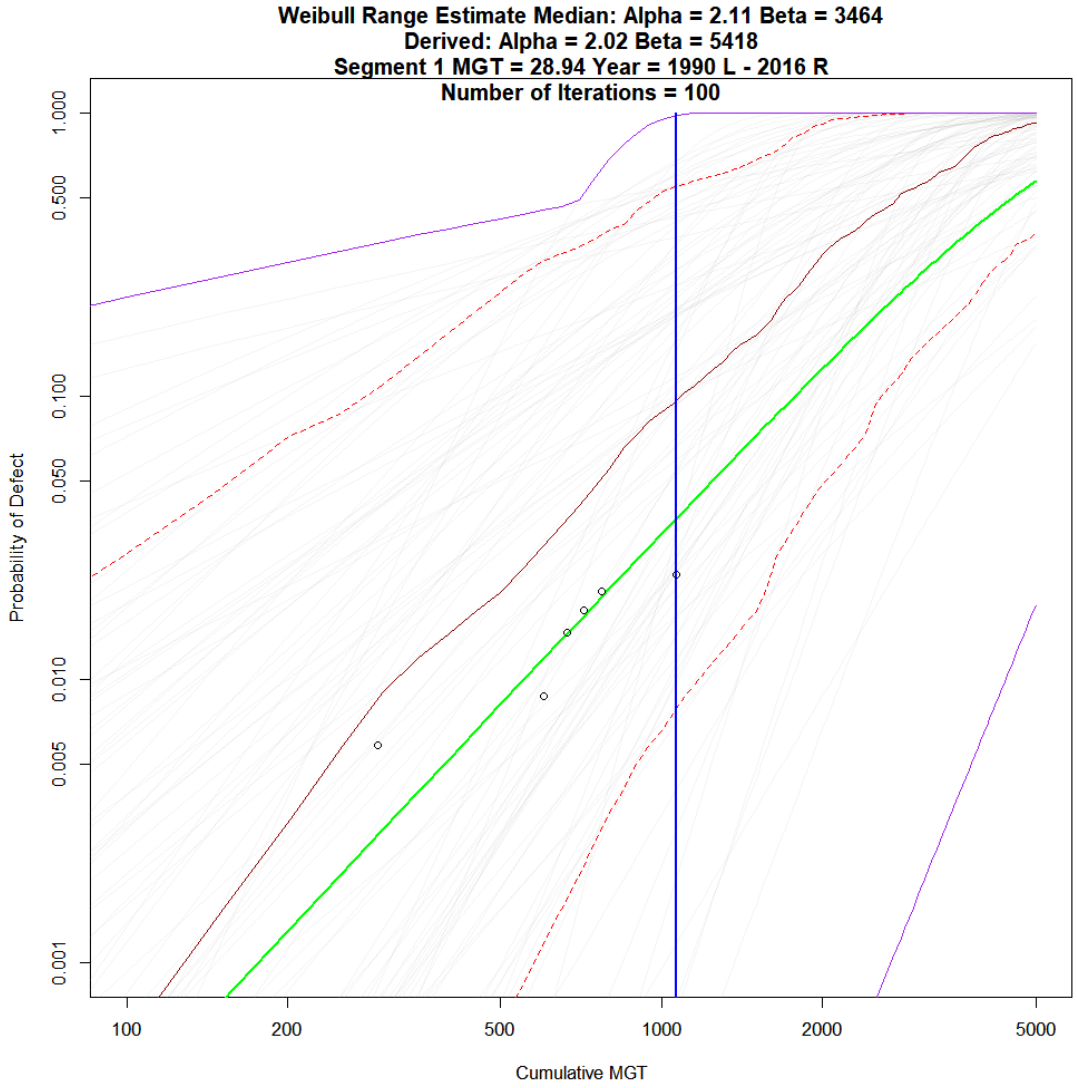


Figure 35: 100 Bootstrapping Iterations Weibull Output

Figure 35 presents the results after 100 iterations. Note, the results are rather jagged. This is due to both the limited number of iterations, as well as the wider spread of the lines; a narrower spread will tend to have smoother lines, as any differences in the lines are reduced as the spread decreases.

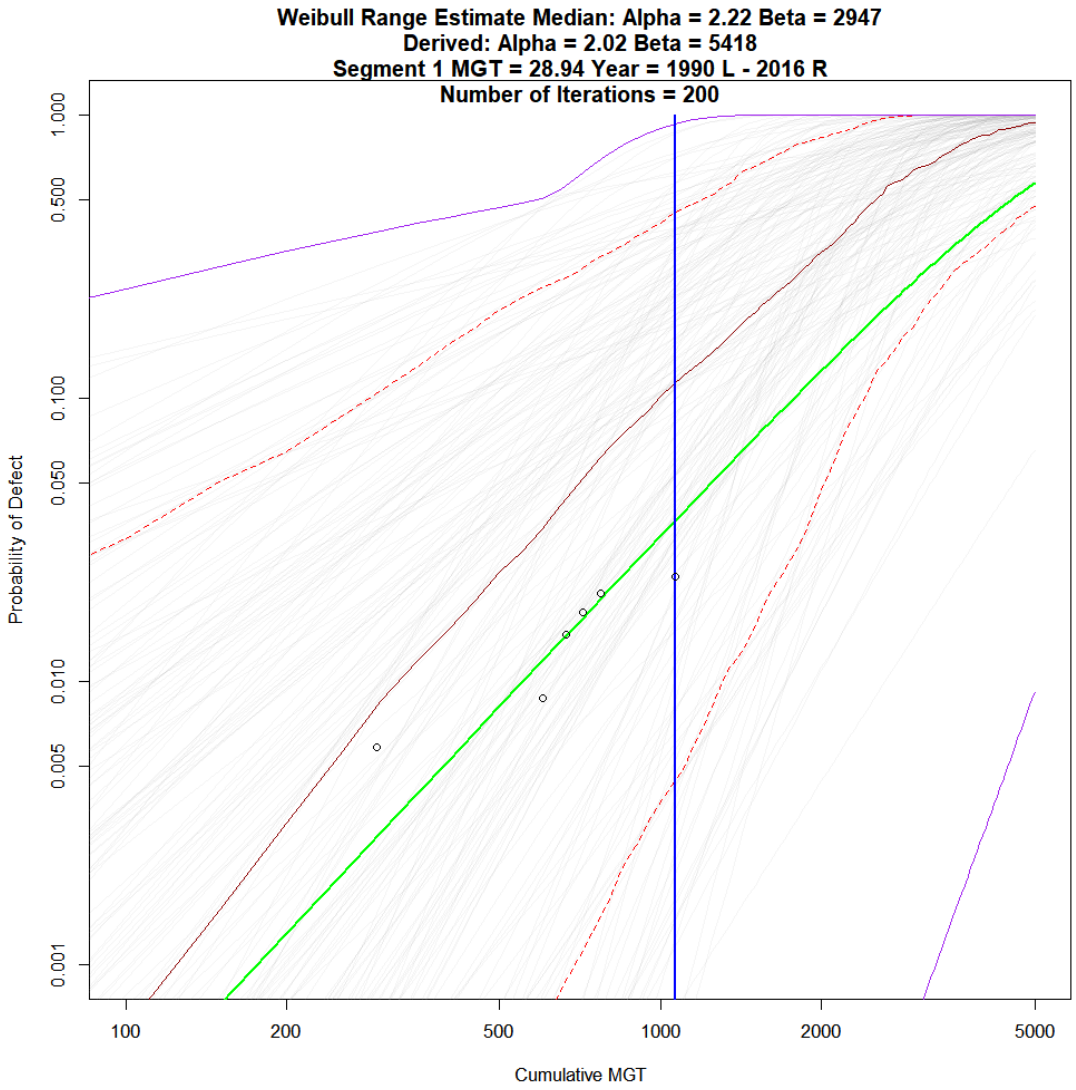


Figure 36: 200 Bootstrapping Iterations Weibull Output

Figure 36 presents the results after 2200 iterations; note; the median line starts to smooth a bit. This results in a shift of +0.11 to Alpha, and -517 to Beta.

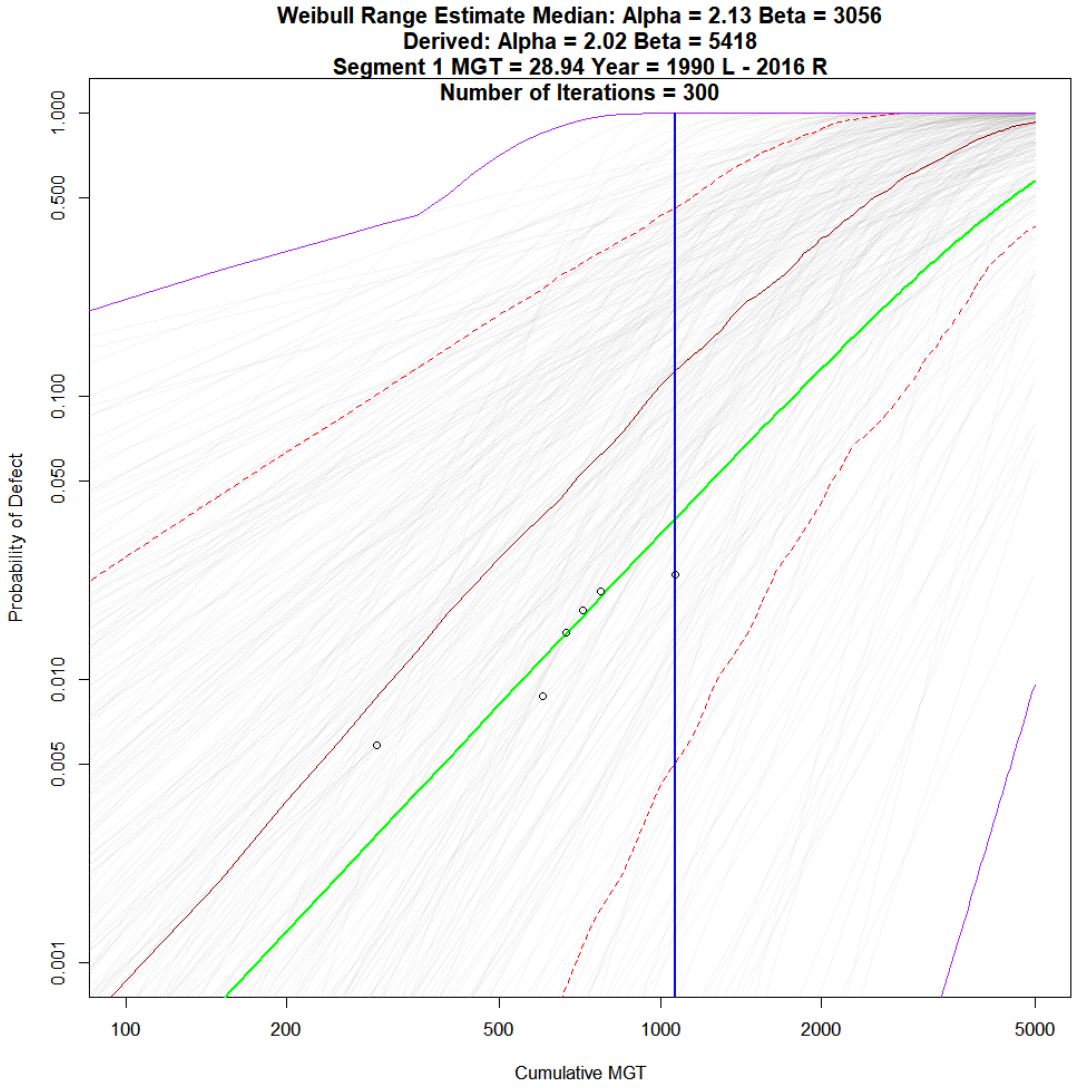


Figure 37: 300 Bootstrapping Iterations

Figure 37 presents the results after3after 300 iterations, note the Alpha value has moved closer to the initial 100 iteration's 2.11, while the Beta value stays near the 200 Iteration's 2947.

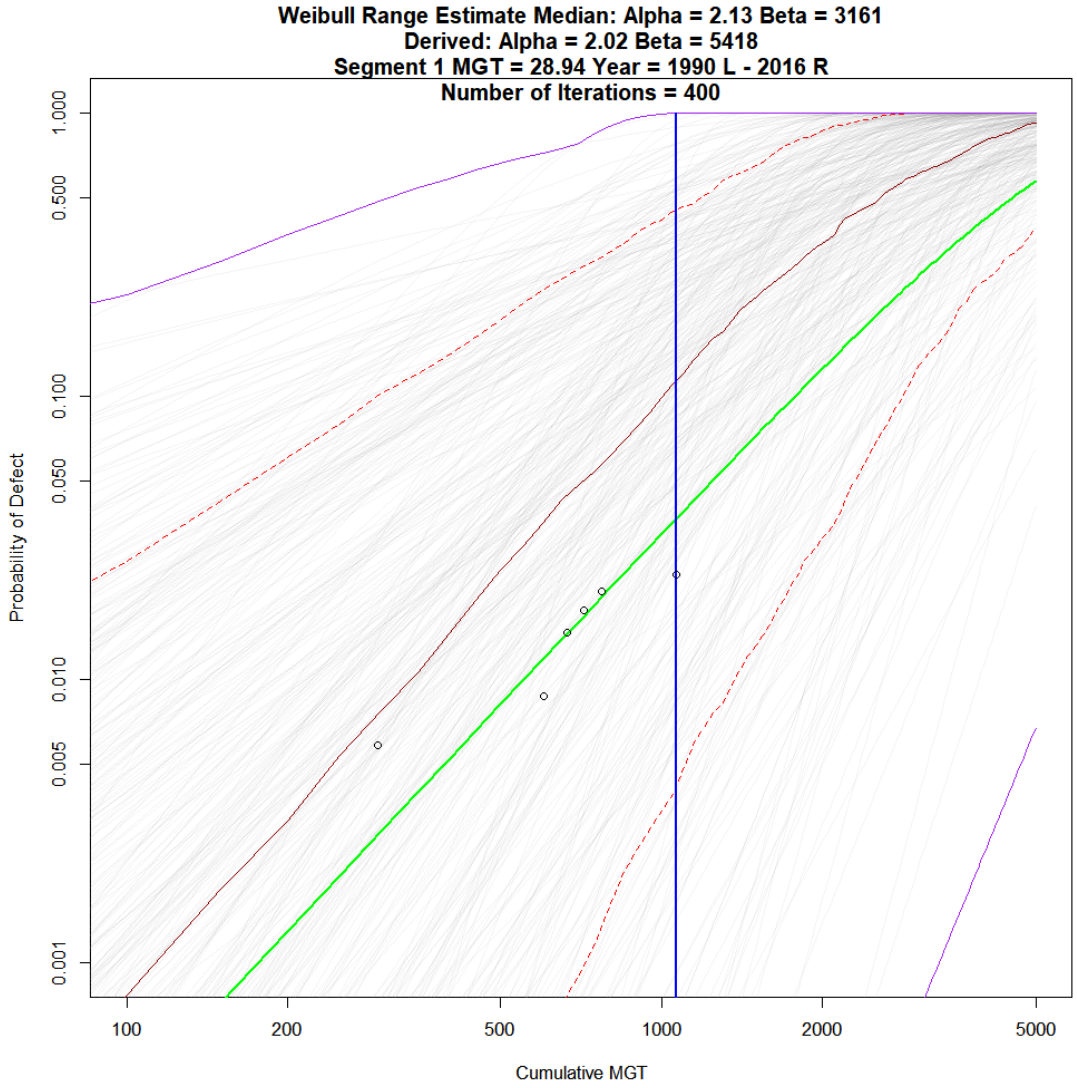


Figure 38: 400 Bootstrapping Iterations

Figure 38 presents the results after 400 iterations. At this point, it seems that the Alpha value is stabilizing around 2.13, most likely minor changes in the 3rd and beyond decimal places which are causing the Beta shifts. There is also the addition of a new higher extreme value compared to the previous iterations, removing the noticeable "bump" that was present. This "bump" was the result of two different "extreme" Weibull curves, one which dominated due to a moderate to low Alpha value and low Beta value, and then one with a high Alpha value and low Beta value, which allowed the second Weibull to exceed the first. In general, this is happening within the main body of the Weibull overlays, with slight differences in Alpha and Beta causing Weibull plots to cross each other at various points as well, but because it is happening with numerous plots at once, it is harder to see.

Figure 39 through Figure 44 show increasing iterations, from 500 to 1000. Note, the increasing smoothness of the curves. This increase in smoothness is a result of having more sources from

which to develop the values for the curves, as well as having the same general trend for the values. In a few cases, you can see how the extreme value curves have large changes in direction as one Weibull overtakes another; since these outliers are fewer in number compared to the center range, they have less of an effect on changing the inner boundary curves.

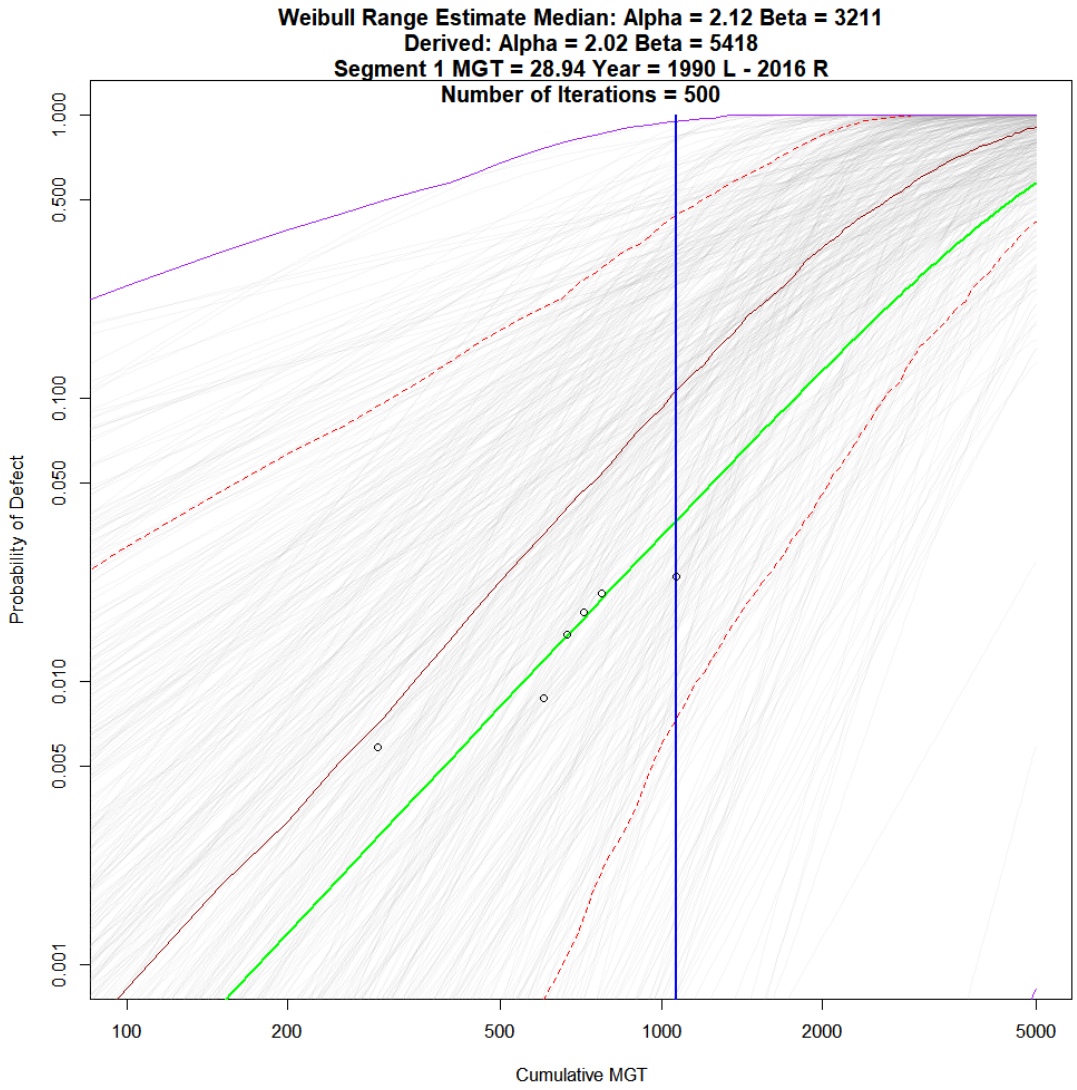


Figure 39: 500 Bootstrapping Iterations

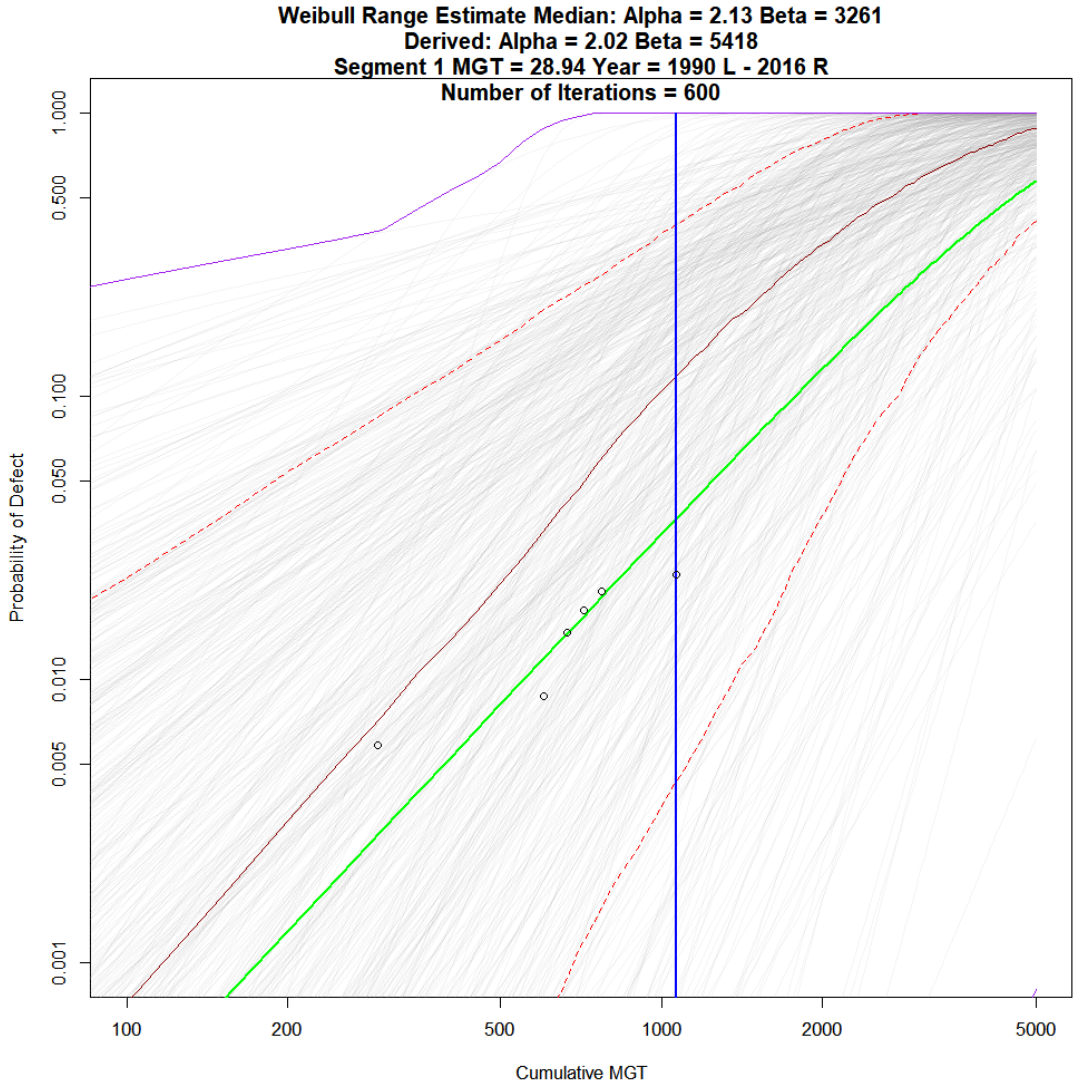


Figure 40: 600 Bootstrapping Iterations

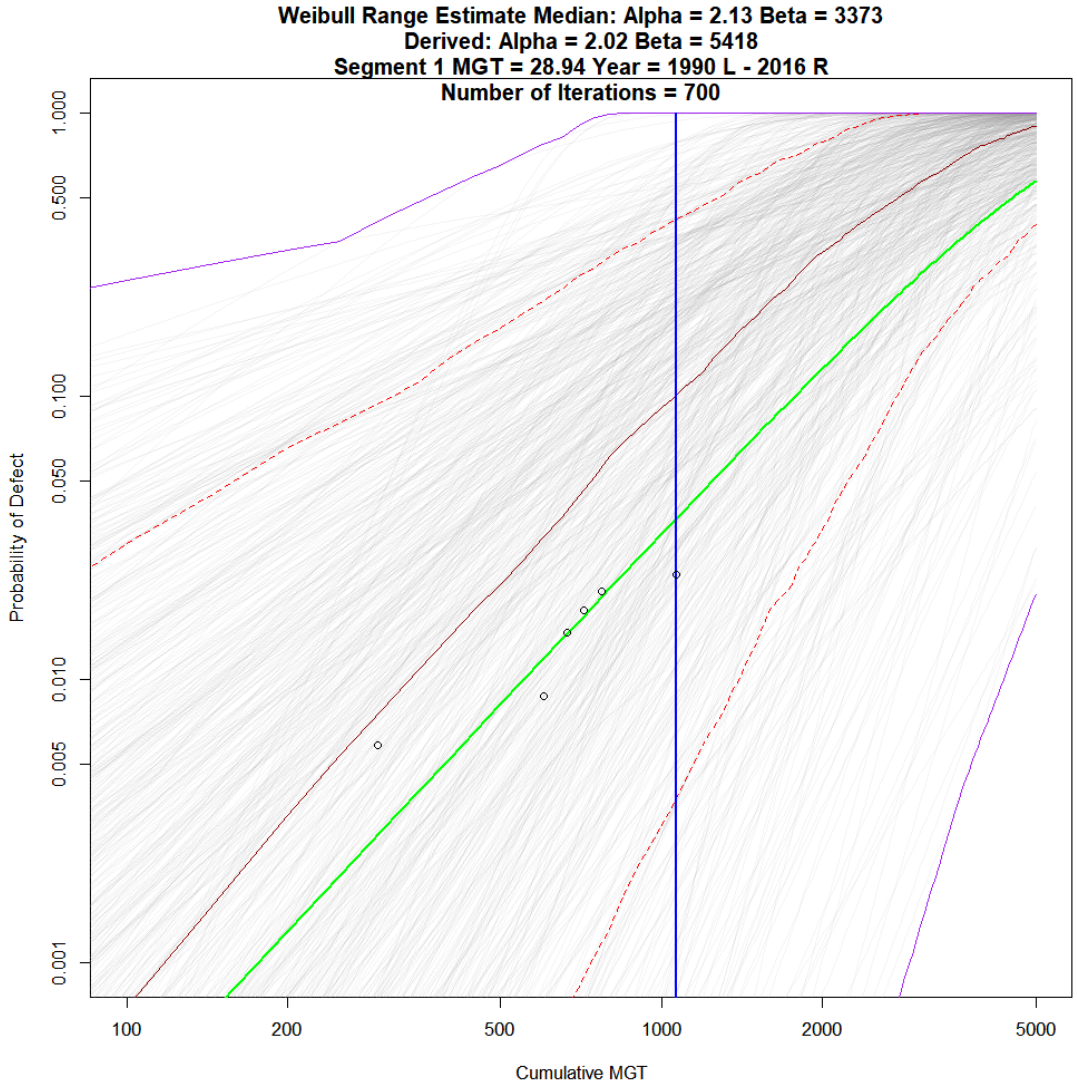


Figure 41: 700 Bootstrapping Iterations

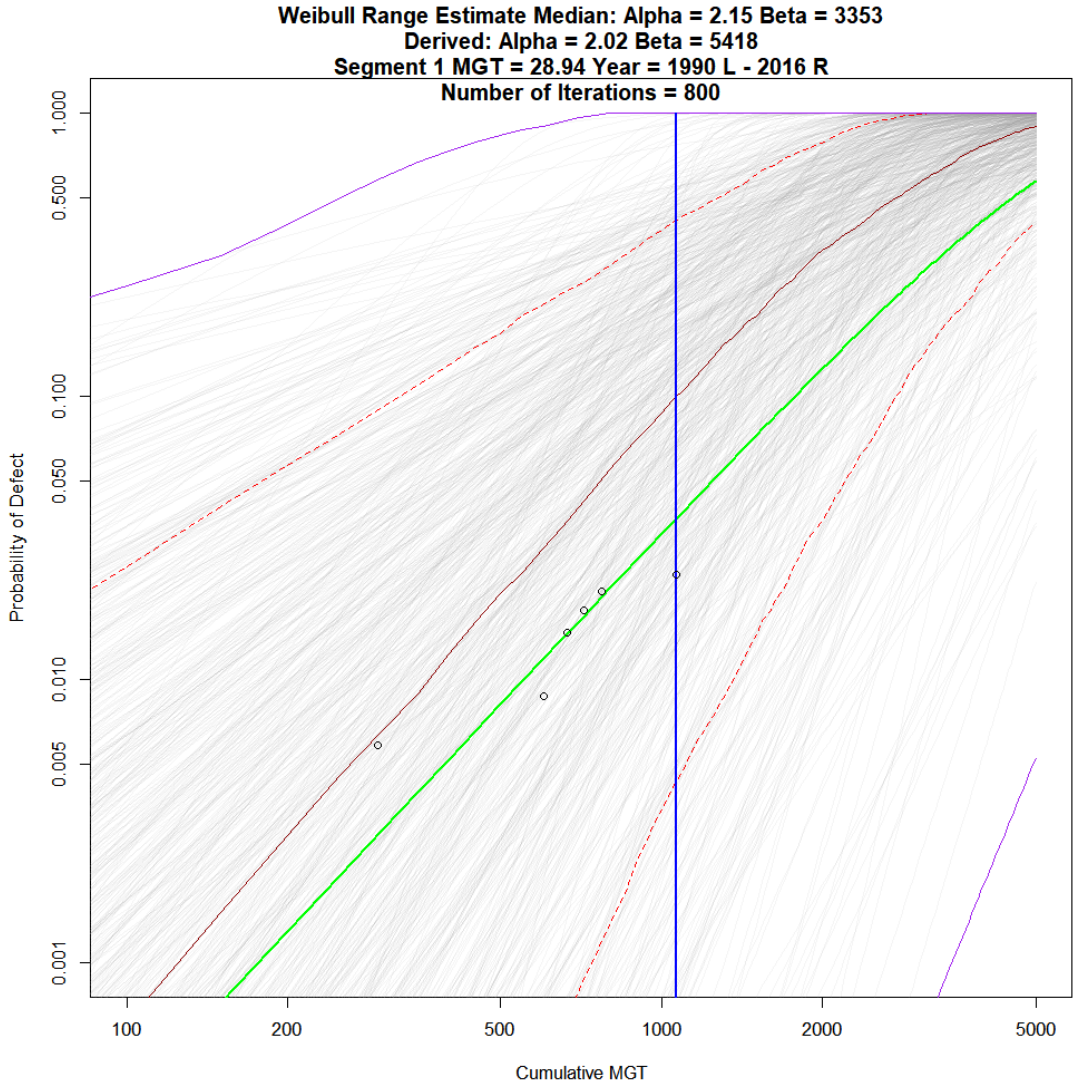


Figure 42: 800 Bootstrapping Iterations

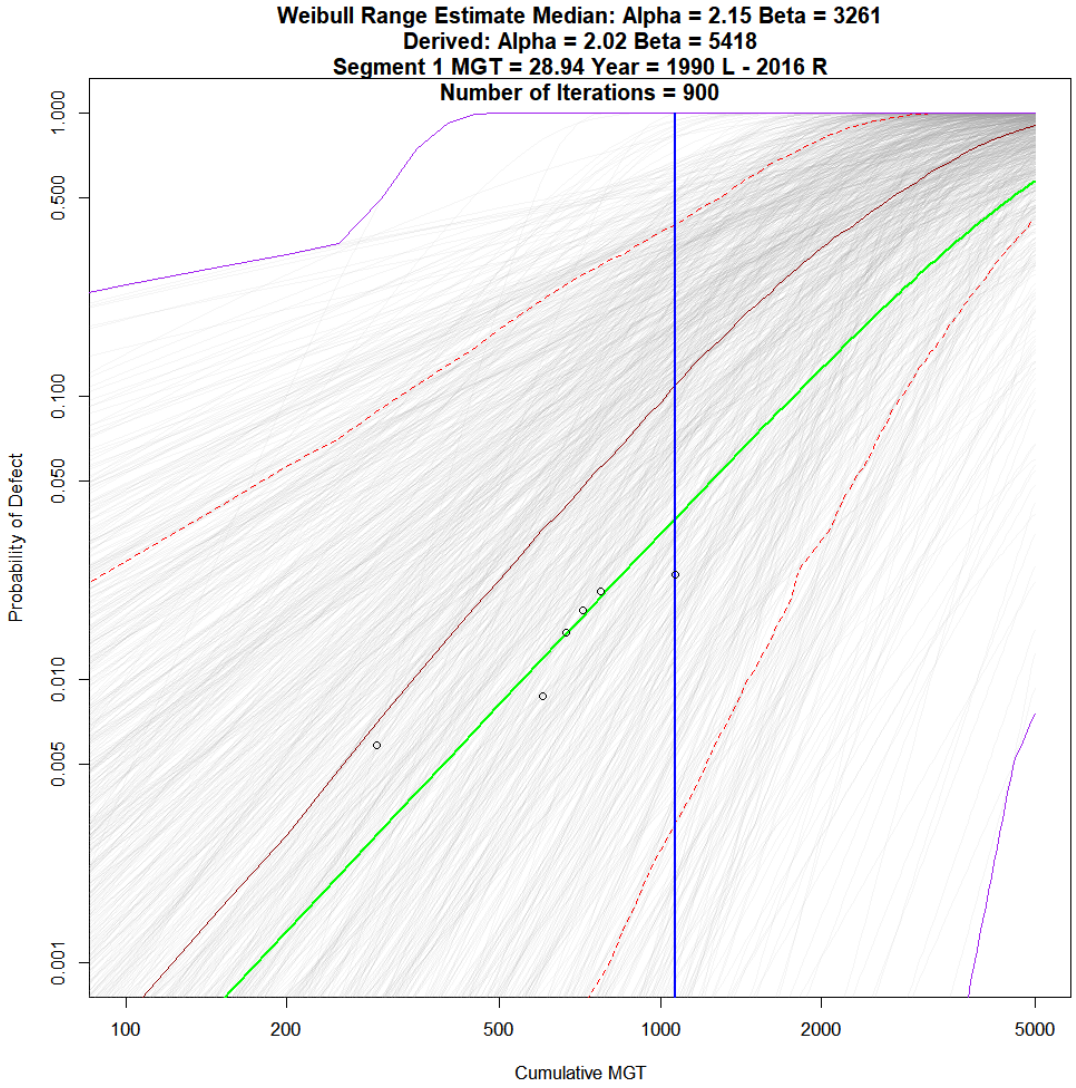


Figure 43: 900 Bootstrapping Iterations

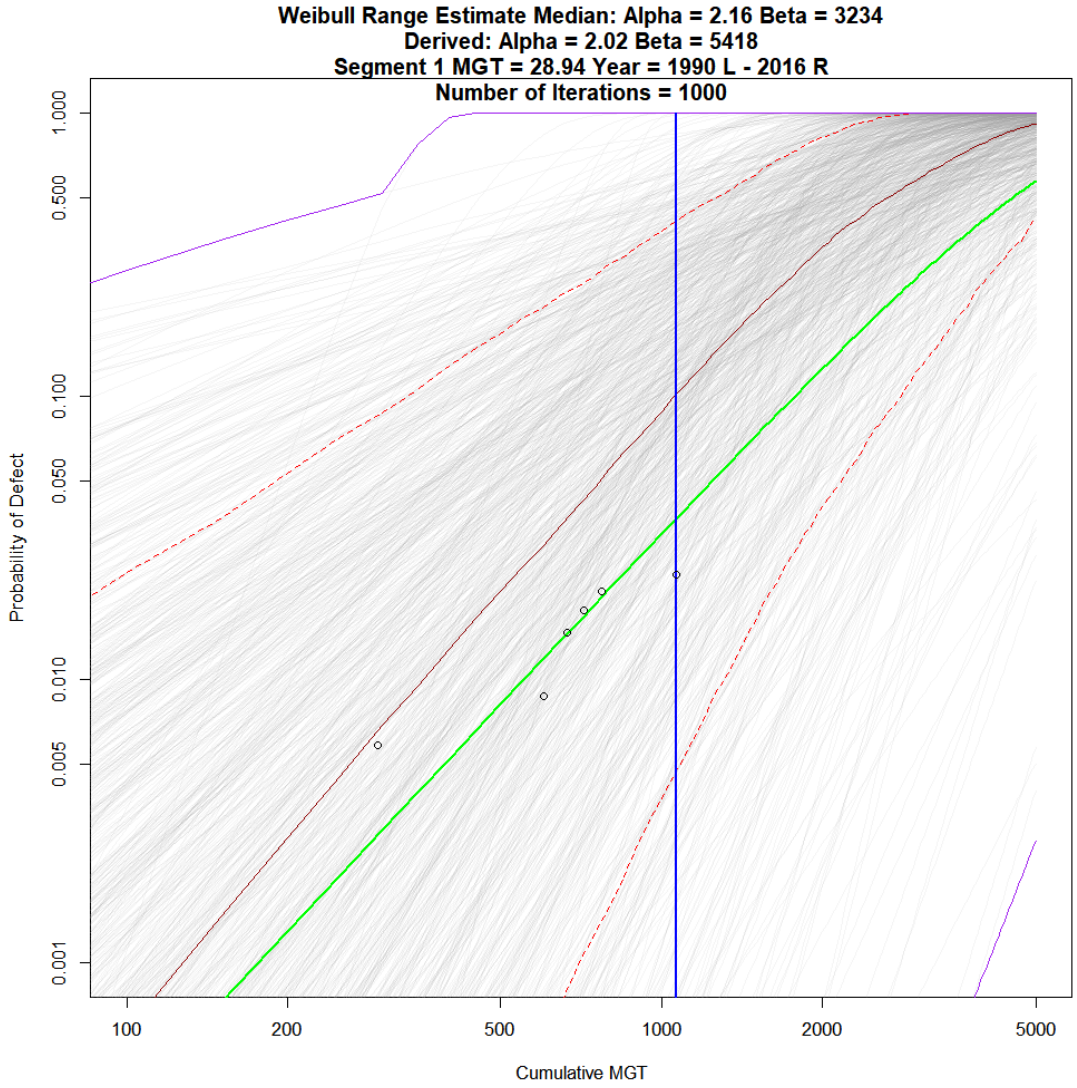


Figure 44: 1000 Bootstrapping Iterations

Accordingly, with the boundaries in place while selecting Weibull Alpha and Beta values, the distribution of values tended to shrink and close in when plotted, as shown in the comparison between Figure 43 and Figure 44, whereas the extreme values from the Iterations grow as rarer extreme values occur than the prior (Figure 35 through Figure 42) shown results. While known data was plotted when available[8], it sometimes did not match up with the bootstrapped data, an indication that the track segment itself was acting differently than what similar segments would be doing. Figure 45 shows how the initial 100 iterations do not include the known data's curve, yet the 1000 iteration, Figure 46 , version does. In Figure 45, the green line represents the current "Best fit" 2-parameter Weibull applied to the defect data; note that between approximately 600 to

[8] Known data is plotted as "circles" in the Figures

1000 MGT the best-fit line crosses the “extreme value” line. Whereas in Figure 46, the green “Best Fit” line stays within the “extreme value” boundary the entire time.

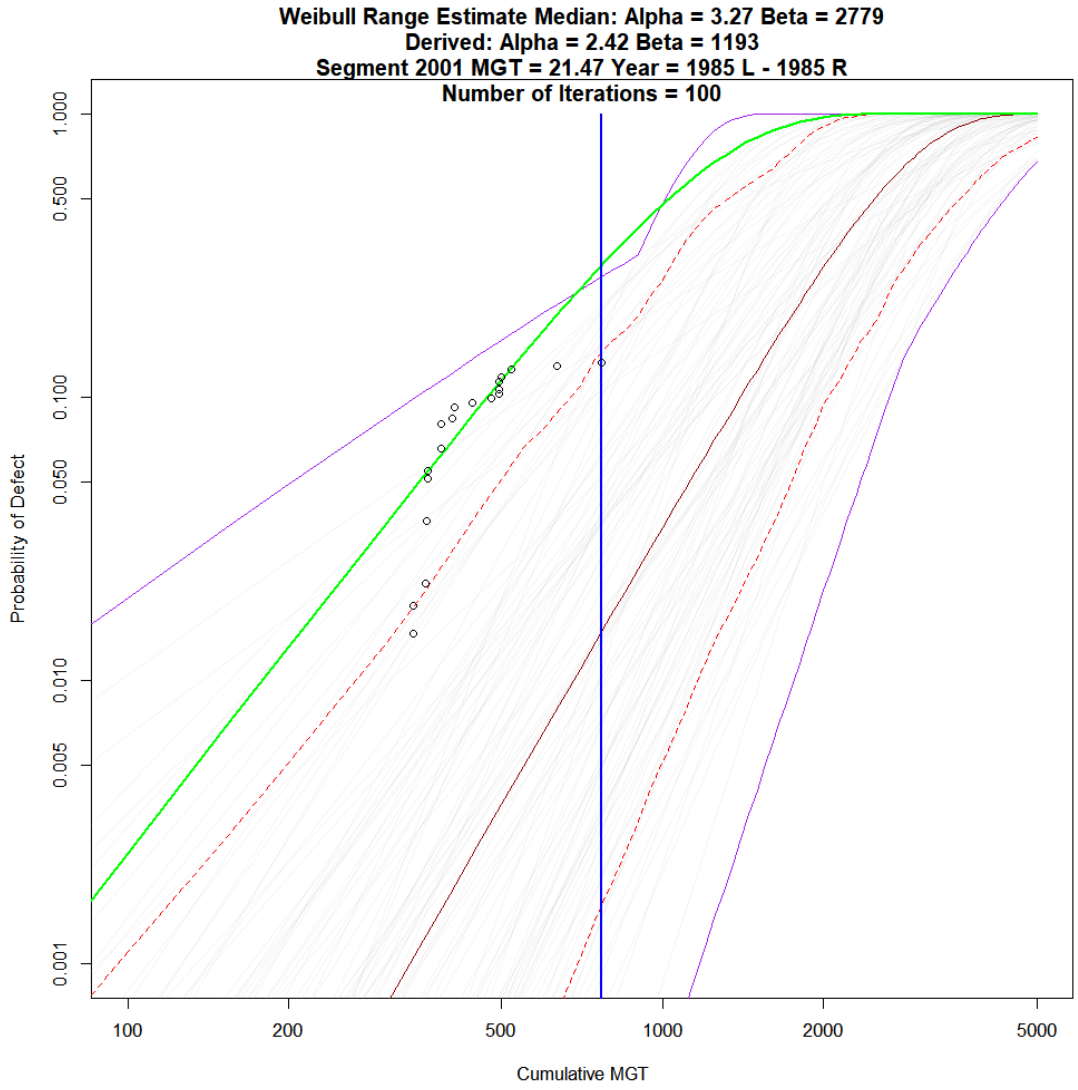


Figure 45: Initial 100 Iterations of Bootstrapping showing Known Data outside of Prediction space

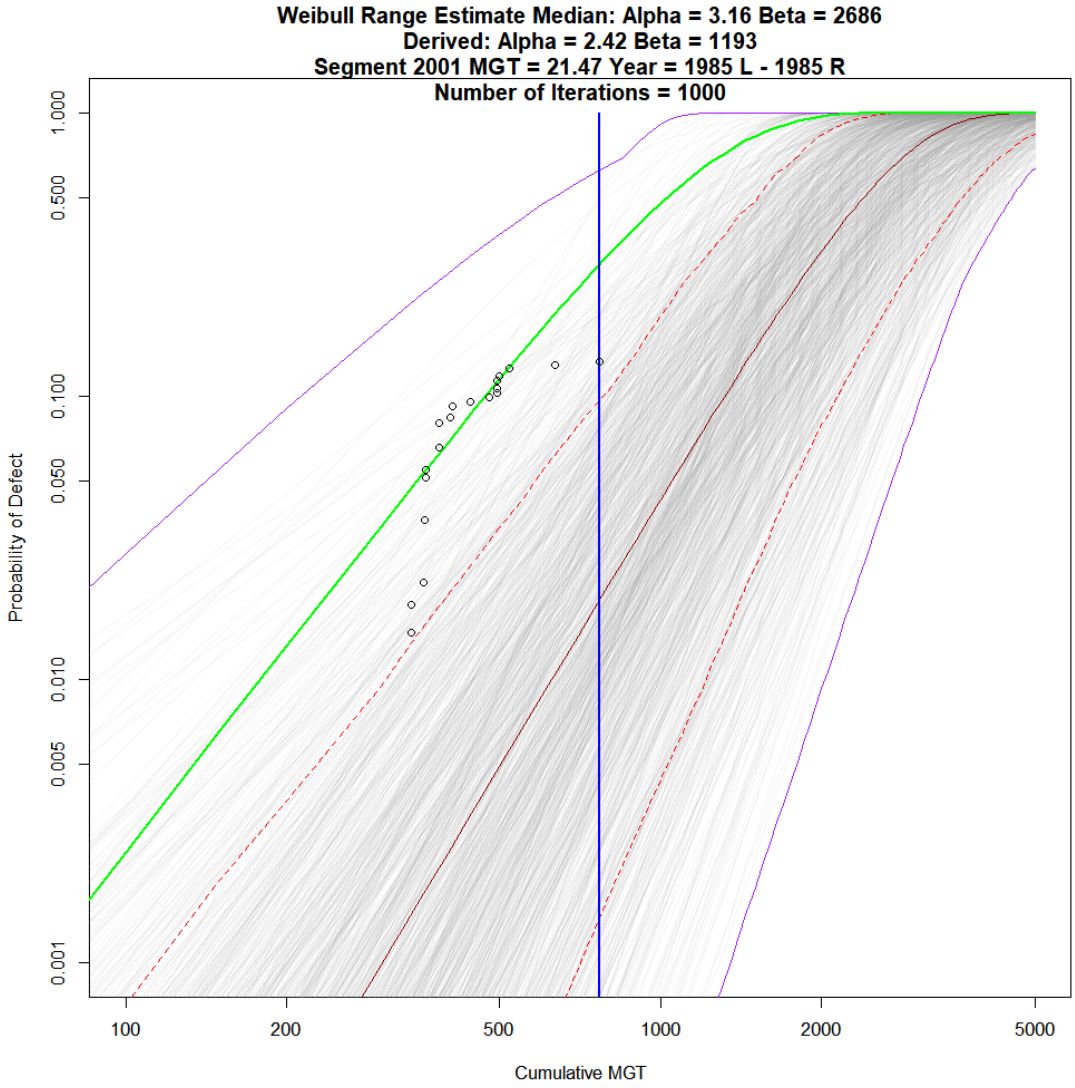


Figure 46: 1000 Iterations of Bootstrapping showing Known Data within the Prediction space

So far, all the Weibull bootstrapped plots shown have had the same general shape; this is not the case for all of the data. Several different types of shapes became apparent as the data was processed and inspected. These shapes depended upon the initial Alpha and Beta distributions; a narrow Beta distribution would result in a tighter clustering of the lines, while a tighter Alpha would result in more straight lines. Figure 47 shows a case of a narrow Beta distribution, and a wider Alpha distribution resulting in a funnel effect of the lines.

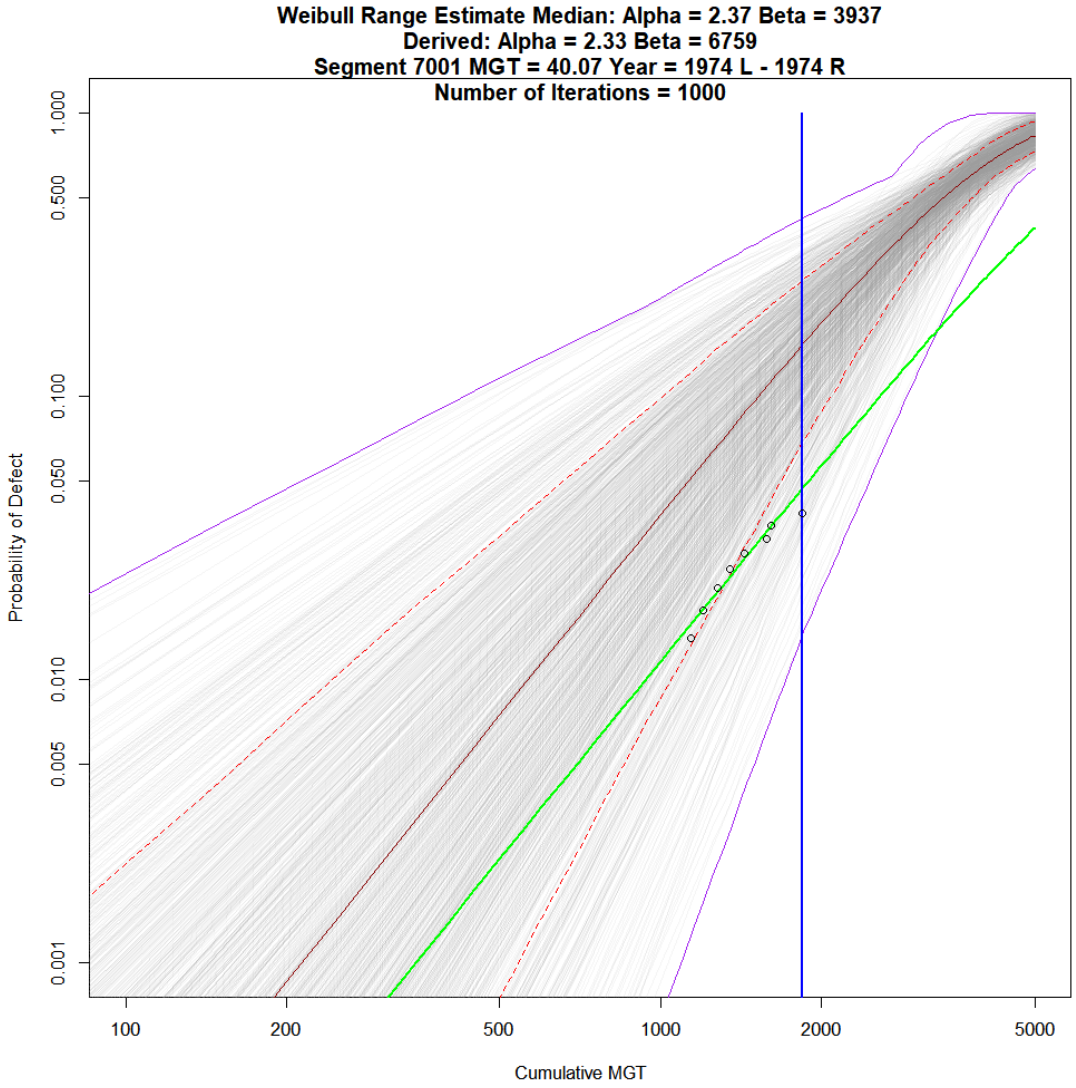


Figure 47: Bootstrapped Weibull results showing "funnel" behavior

Figure 48 shows the effect when a narrow Alpha distribution is used, showcasing the relatively straight lines going across the graph.

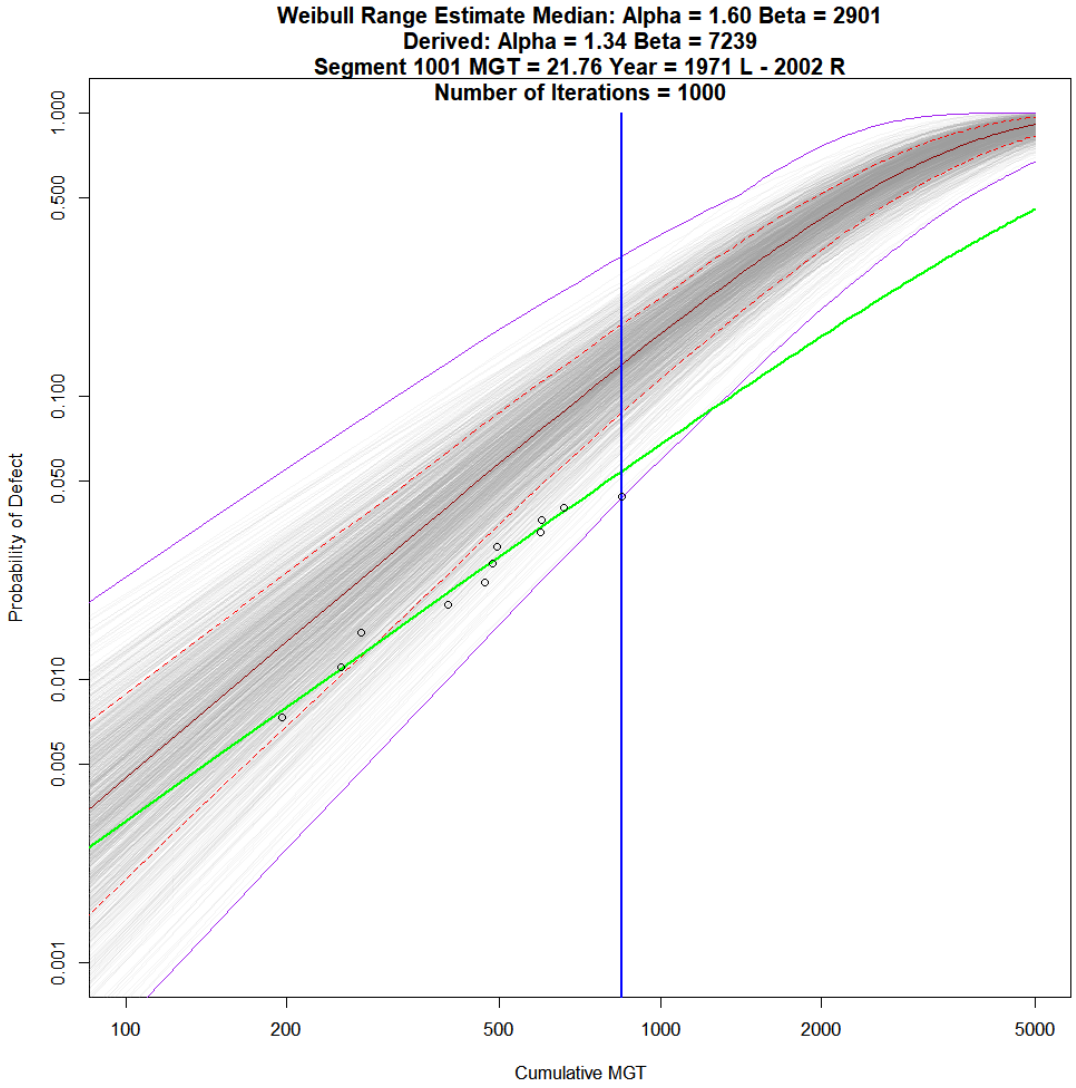


Figure 48: Bootstrapped Weibull results showing behavior of low Alpha range

These shapes come into play with the next part of the Weibull Bootstrapping analysis: prediction of future defect probabilities. From these layered Weibull graphs, it became possible to develop a distribution based on the frequency of Weibull lines crossing at certain points. In essence, the graph was "cut" vertically or horizontally, and the resulting distribution of the Weibull lines was used to develop future probability densities. Using Figure 47 as an example, a vertical cut was made at the Blue line, and at +50, +100, +150 MGT, giving four views of how increasing Cumulative MGT shifted the relative probabilities of the Probability of a Defect. Figure 49 shows these cuts using that data. Note, the shift to the right (increasing MGT) as expected,

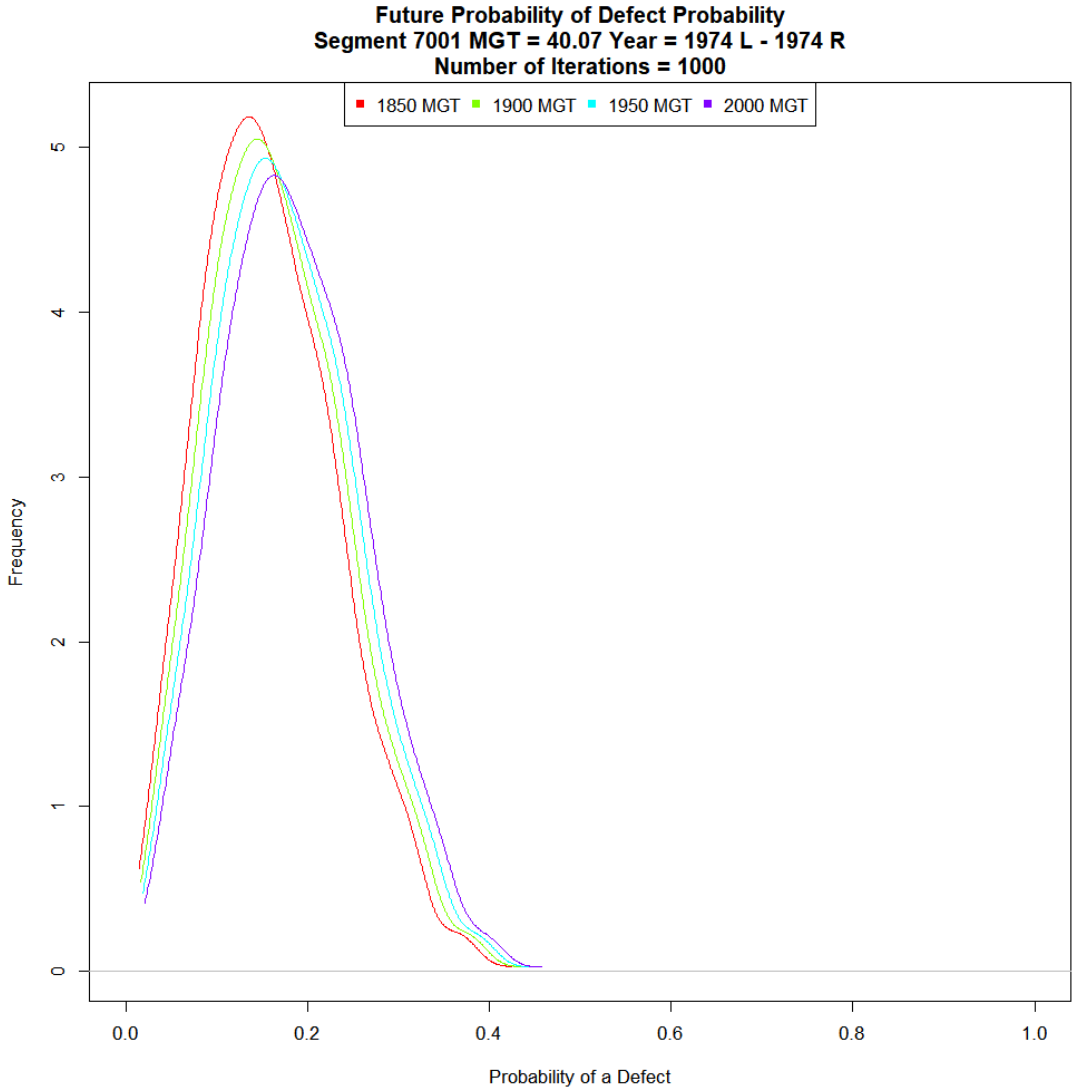


Figure 49: Early result of Frequency Density cut from Bootstrapped Weibull output

From this graph, it becomes apparent that as the Cumulative MGT increases, the frequency of having a low Probability of a Defect drops, while the frequency of having a moderate, in this case, Probability of a Defect rises. In the case of Figure 49, if we look at the relative frequencies at a Probability of a Defect value of 0.1, we can see that the 1850 MGT value (red curve) is higher than the others, indicating that it is more likely to happen. As we move to Probability of a Defect value 0.2, the 1850 MGT values peak, then drop below the other values, indicating that past approximately 0.18 to 0.2, the other MGT values are more frequent compared to the 1850 MGT slice. In addition, the peak of each successive MGT line drops compared to the previous, while the endpoint also shifts to the right, indicating that as Cumulative MGT increases, the Probability of a Defect's extreme value grows, making it more likely that previously unexpected values will become more frequent

In order to better present these results a series of new graphical presentations were developed. Figure 50 and Figure 51 show the source data, and Inverse Cumulative Density Function

respectively, for a given track segment. The Cumulative Density Function (CDF) is a way to generate probabilities from a density function, as generated from the slices in the previous Figure 50. As the area under a density curve is equal to 1, each point taken from the curve is related to the actual probability of occurrence, with the CDF being the cumulative sum of the densities along the x-axis. Since the Cumulative Density Function is essentially the probability that the value will be equal to or less than "x", the Inverse CDF, or 1-CDF, is the probability that the value will be greater than "x". The use of the Inverse CDF is intended to showcase how the Probability of a Defect was shifting towards higher values; since the lines can be thought of as "% make it to here", having a line shift to the right indicates that a higher probability of a defect is more likely compared with a line to the left. While these functions provide additional information about the frequencies of the probabilities, they were found to be confusing and difficult to understand, especially the Inverse CDF. As a result, it was decided to use the density method to present these density "cuts" as presented in Figure 49, as they are easier to understand without requiring detailed explanation.

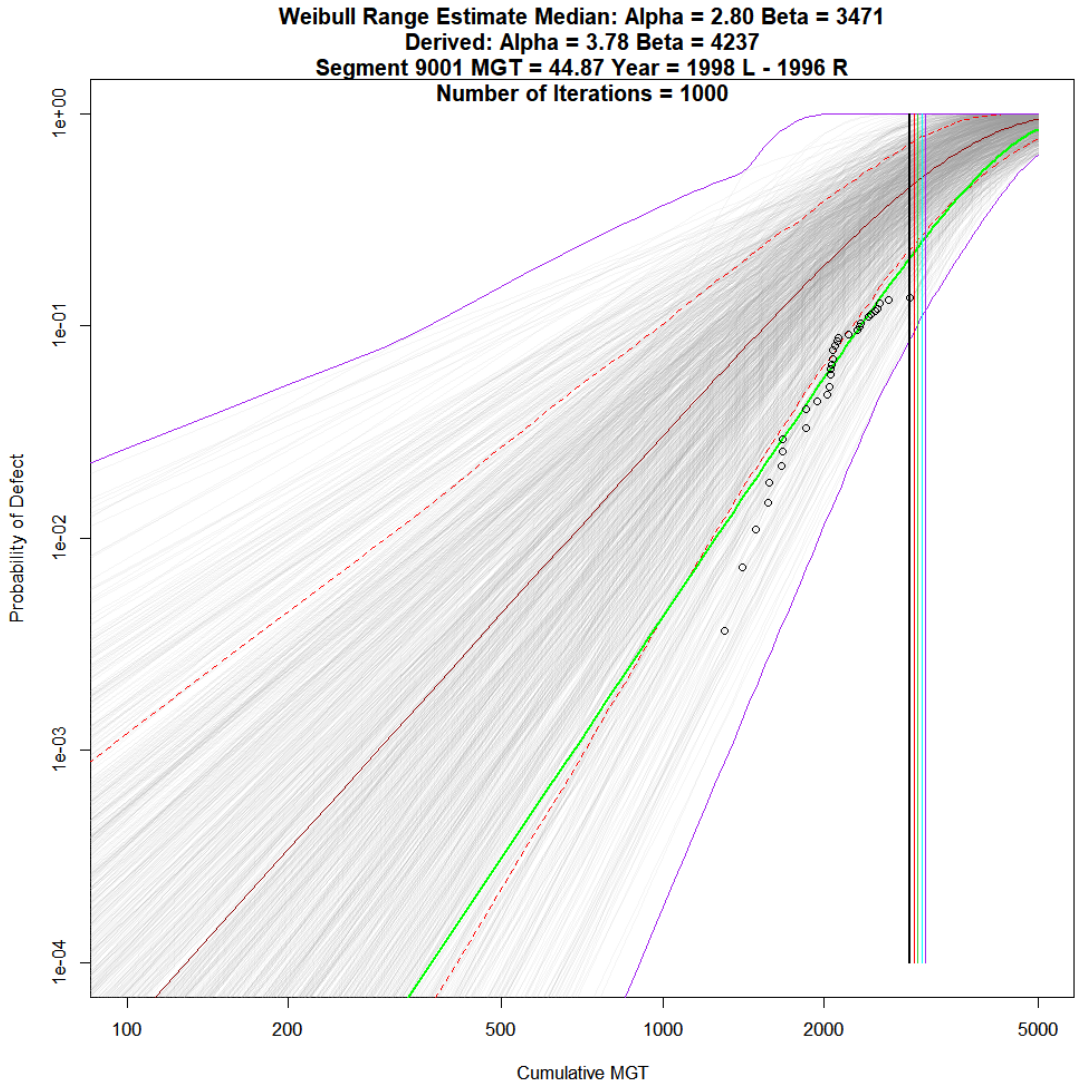


Figure 50: Bootstrapped Weibull data source for Density Cut example

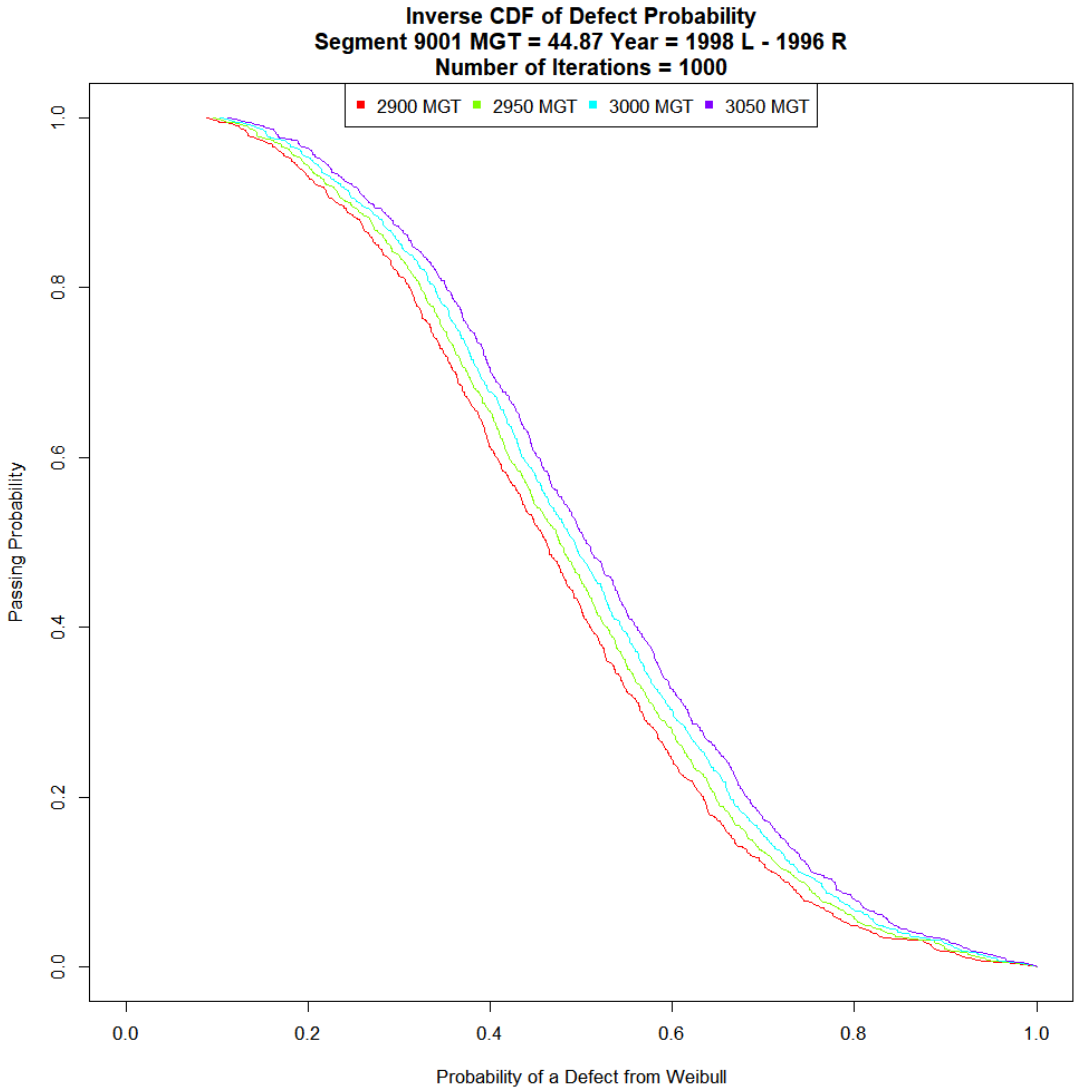


Figure 51: Inverse Cumulative Distribution of Density Cut data for Weibull Bootstrapping

These density cuts also showed some unusual behaviors due to the way the Weibulls were plotted; due to the asymptote at 1/100%. This manifest itself as Weibull lines becoming horizontal, as shown at the top of Figure 52, as they reached this asymptote. This also resulted in producing a double-peak effect in the density cuts, as seen in Figure 53. This behavior in the density graph can be seen as a representation of the track segment in question reaching "failure" or time for replacement due to excessive number of defects, rather than strictly meaning that there is a 100% probability of a defect.

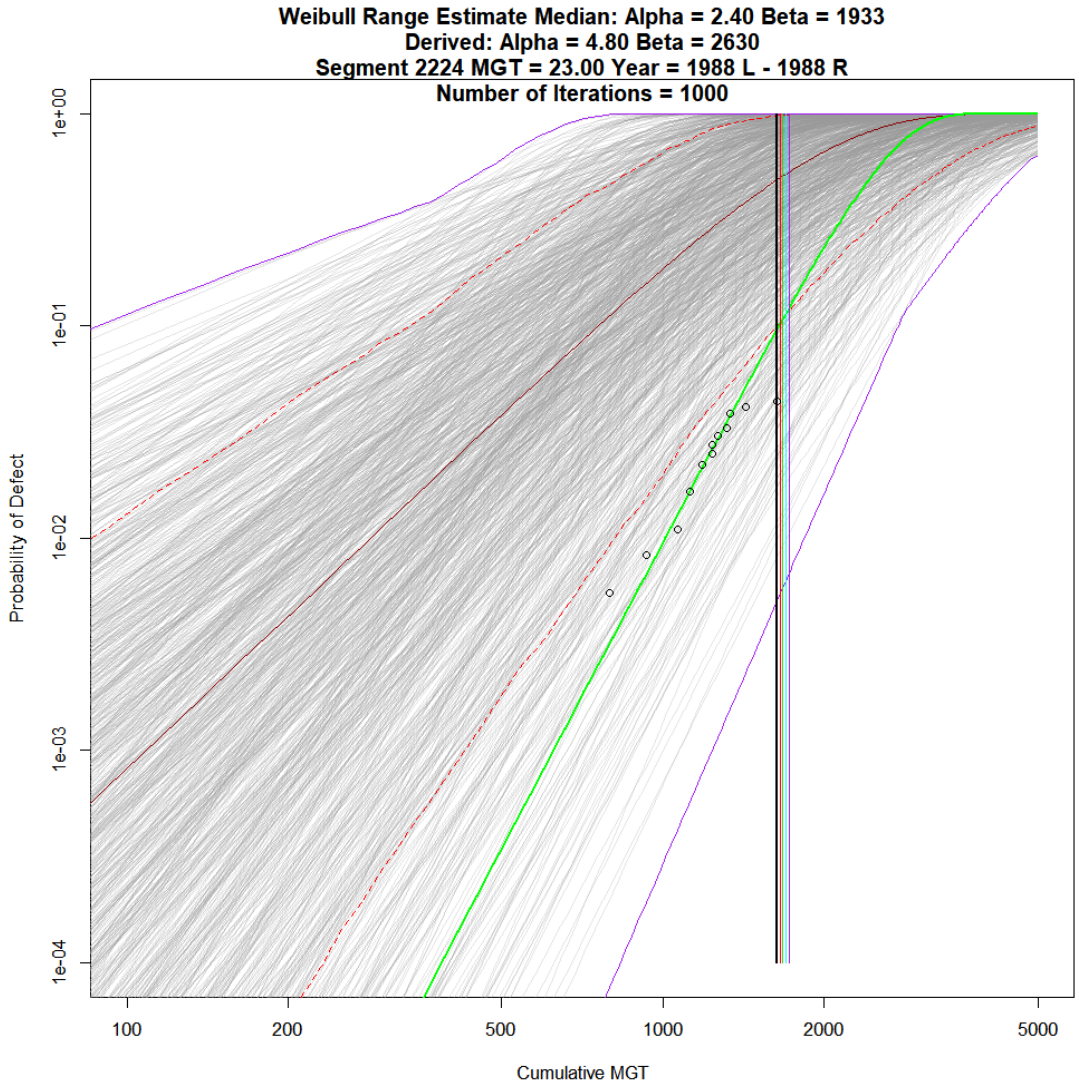


Figure 52: Source Weibull Bootstrap for Density Example; two-peak distribution

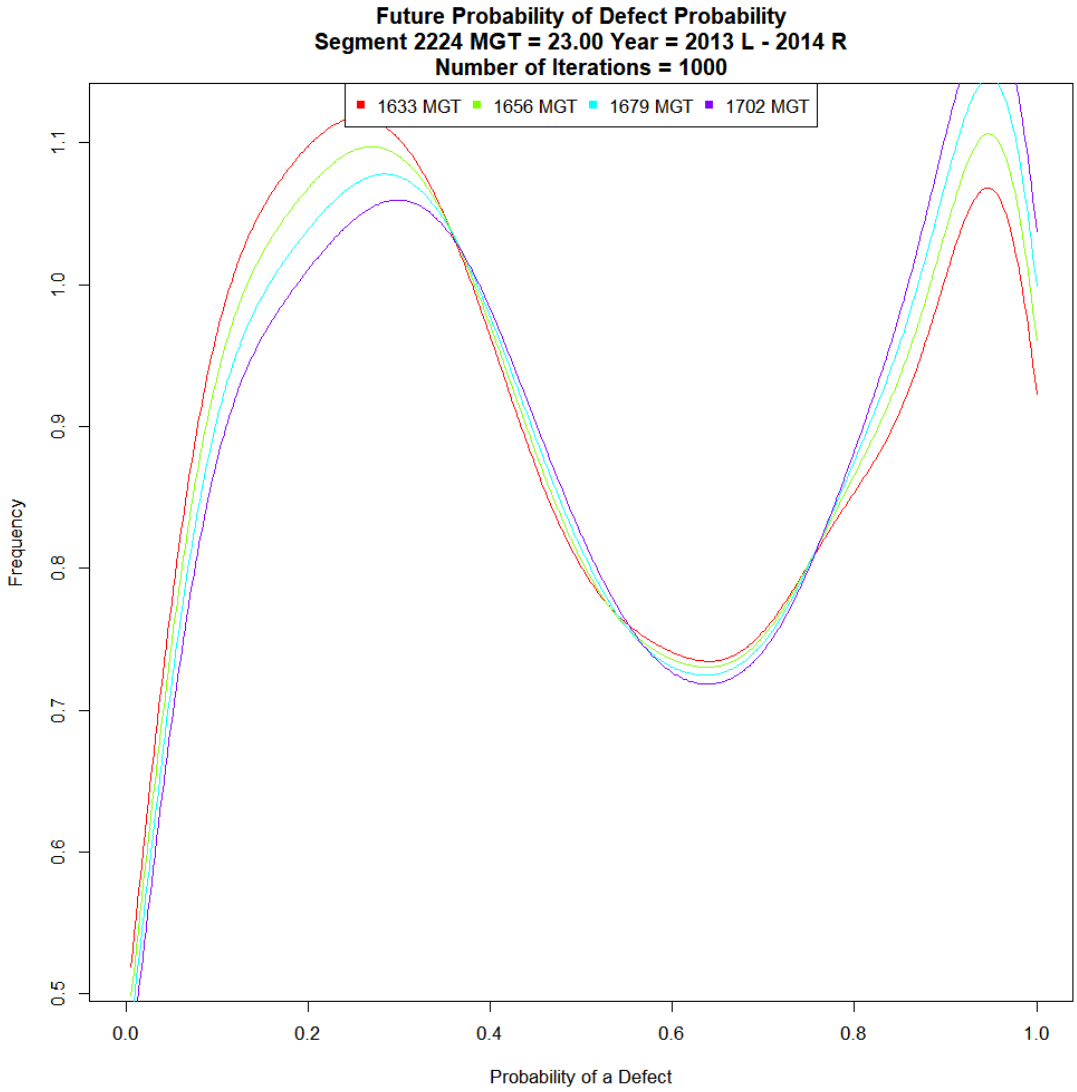


Figure 53: Weibull Density Cut showing Double Peak effect

The Density Cuts were originally defined as cutting across a set Cumulative MGT interval on the Weibull plot, e.g. the red density cut in Figure 53 corresponds to the 1633 MGT vertical line in Figure 52. These cuts were initially made at intervals corresponding to a year's worth of traffic; a year's worth of MGT allows for a visualization of the annual progression of broken rail risk as seen in Figure 52. Alternately a default value for change in MGT can be chosen if the Annual MGT was considered low.

Since railroads often define rail replacement criterion in in terms of such concepts as defect rate (defects/mile/year), this would change the way the Weibull results are presented. To achieve this, it is necessary to "cut" the Weibull curve at the Probability of a Defect, where the resulting density curves would be the relative frequency at which Cumulative MGT values reach that Probability of a Defect. Initially, the cuts were at set intervals as a proof-of-concept display, as shown in Figure 54 and Figure 55, the source data and corresponding density cuts, respectively. These figures present the slices as a function of the probability of a defect, i.e. the red curve in Figure 55

corresponds to a 80% probability that the next defect will occur at the cumulative defined MGT level (horizontal axis).

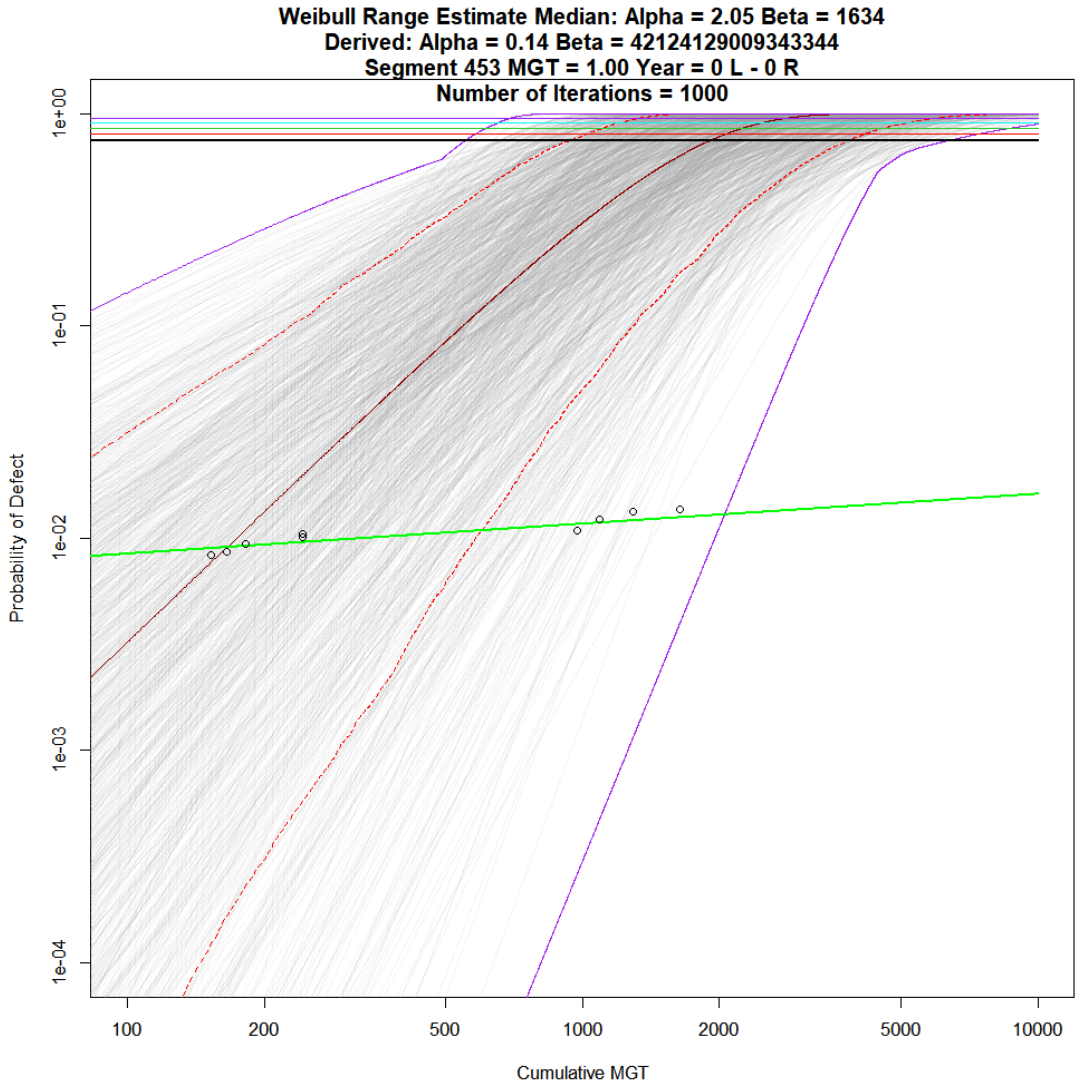


Figure 54: Source Data/Graph for Horizontal Density Cut

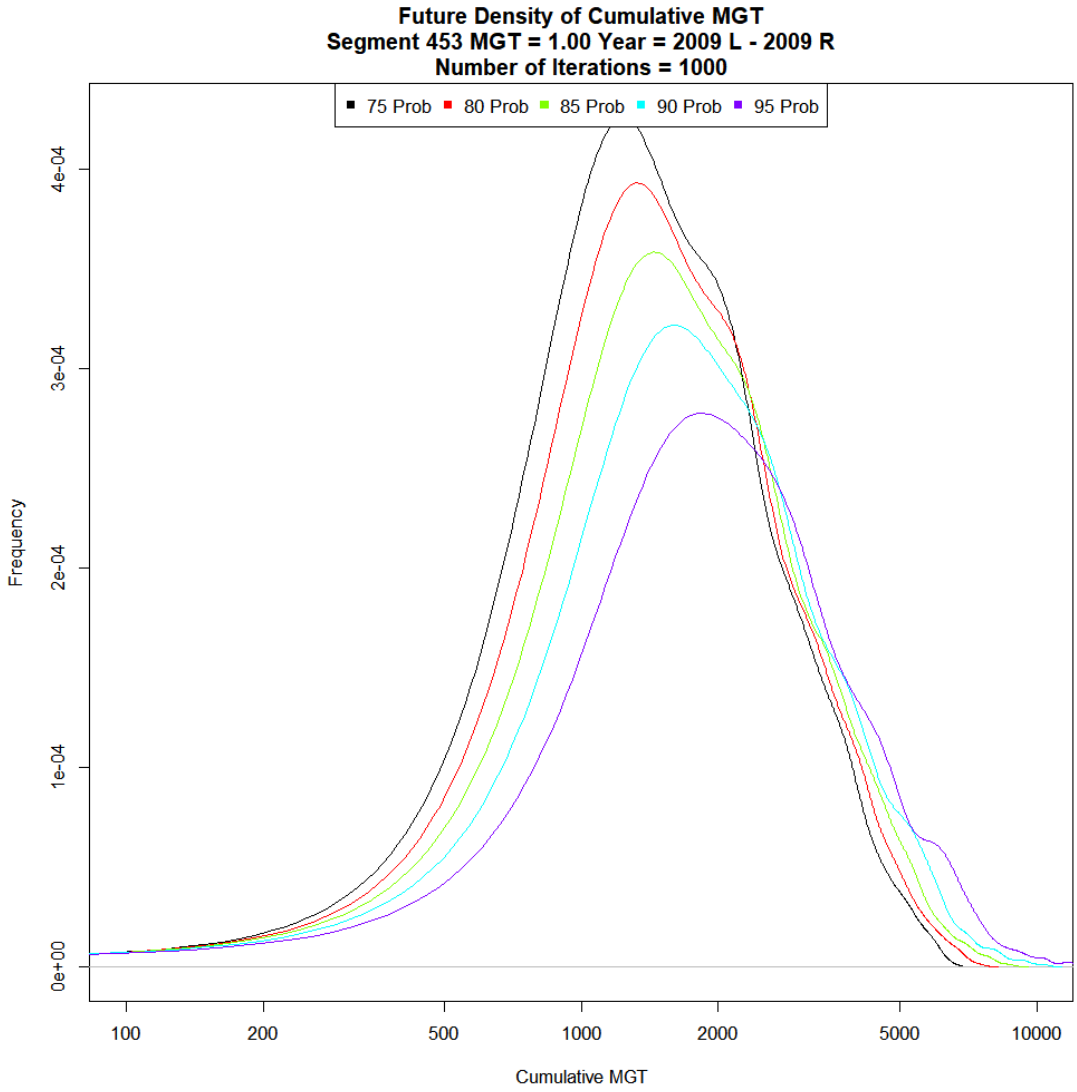


Figure 55: Density of Probability of a Defect

Interpreting this graph, Figure 55, was found to be somewhat complex, since it introduced a new concept, that of the probability of the next rail defect at a specific cumulative MGT level. Each line is the Probability of a Defect in a given rail, for the track segment in question. Given a Cumulative MGT value, it can be seen how the relative frequencies change between the Probability intervals. For example, at 500 Cumulative MGT, the 75% Probability of a defect in a rail is almost twice that of the 95% Probability of a defect in a rail, indicating that the 75% Probability of a defect is more common at the lower Cumulative MGT. Looking at 5,000 Cumulative MGT, the frequencies are reversed, with 95% probability of a defect in a rail being more frequent.

Proceeding from this, the next set of graphs focused on "slicing" the Weibull plot based on the Defects per Mile per Year values, the criterion commonly used by railroads to replace rail[9]. One major issue came about during this work; because the Weibull curve has an asymptote at 1/100%, any value which would correspond to a Cumulative MGT in that asymptote range would return a value of 1/100%. Figure 56 and Figure 57 display the normal behavior, while Figure 58 and Figure 59 show the odd behavior when the asymptote comes into play.

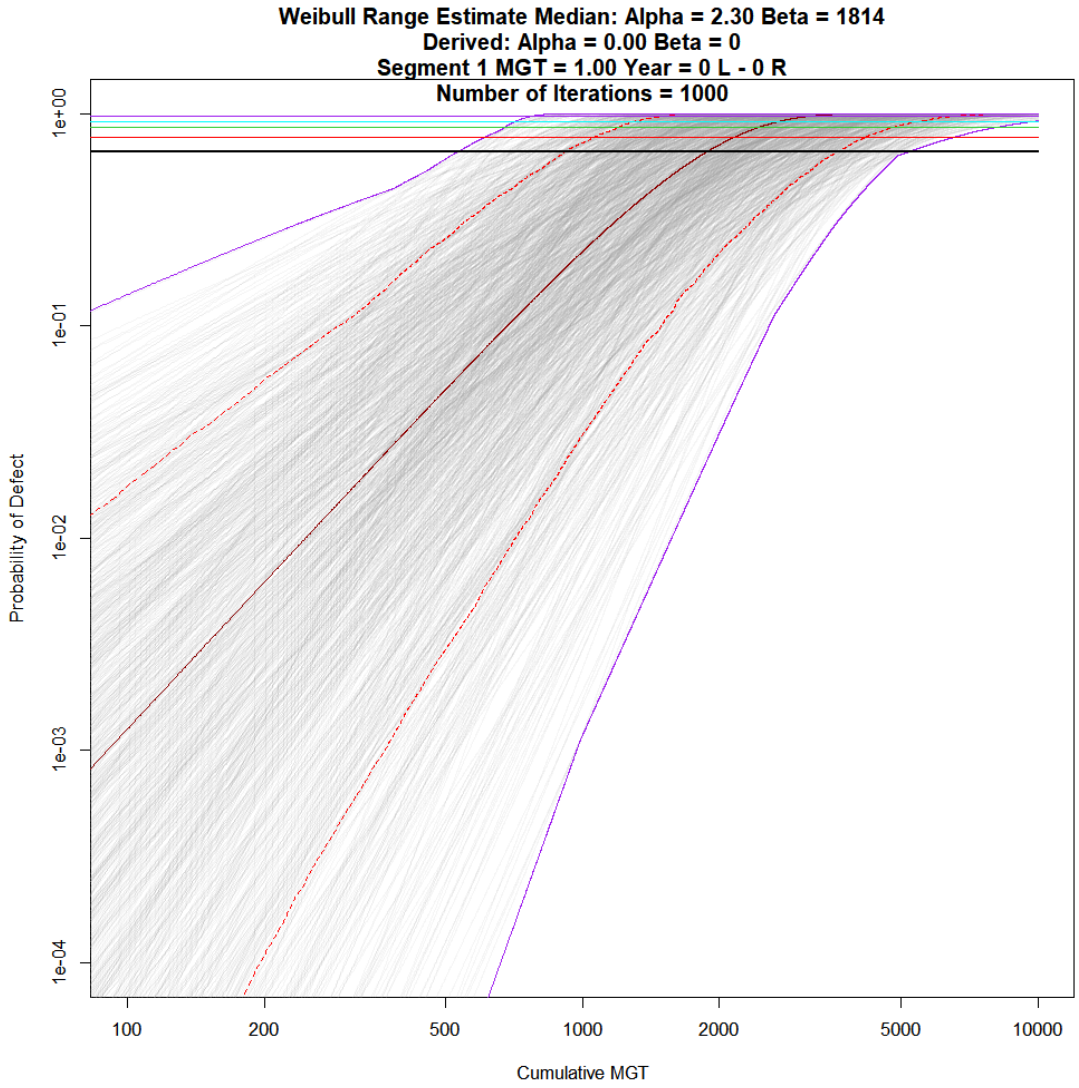


Figure 56: Source Data/Graph for Horizontal Density Cut Defects/Mile/Year

[9] Thus 5 defects/mile/year is a common rail replacement threshold for defect based replacement.

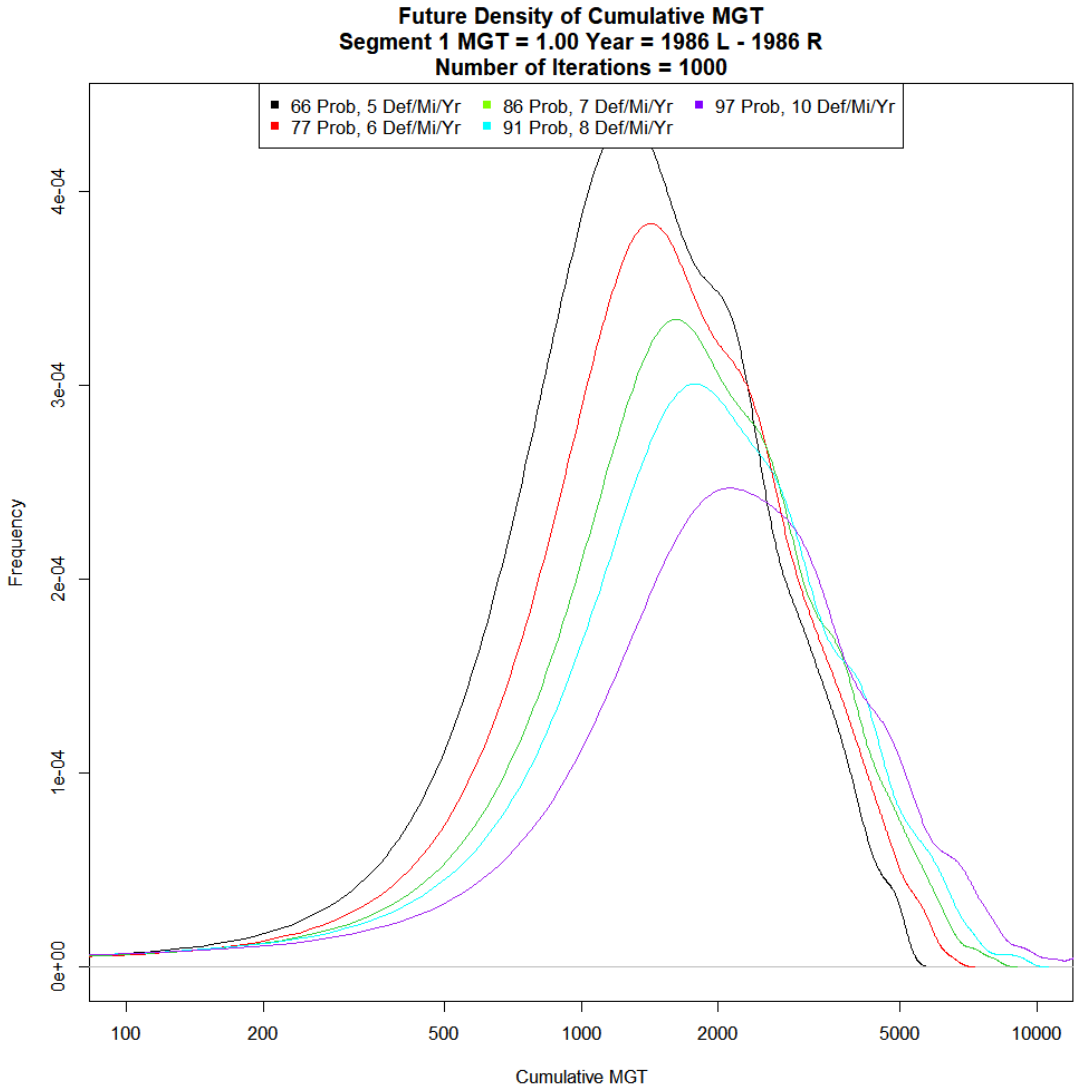


Figure 57: Density Graph of Defects/Miles/Year

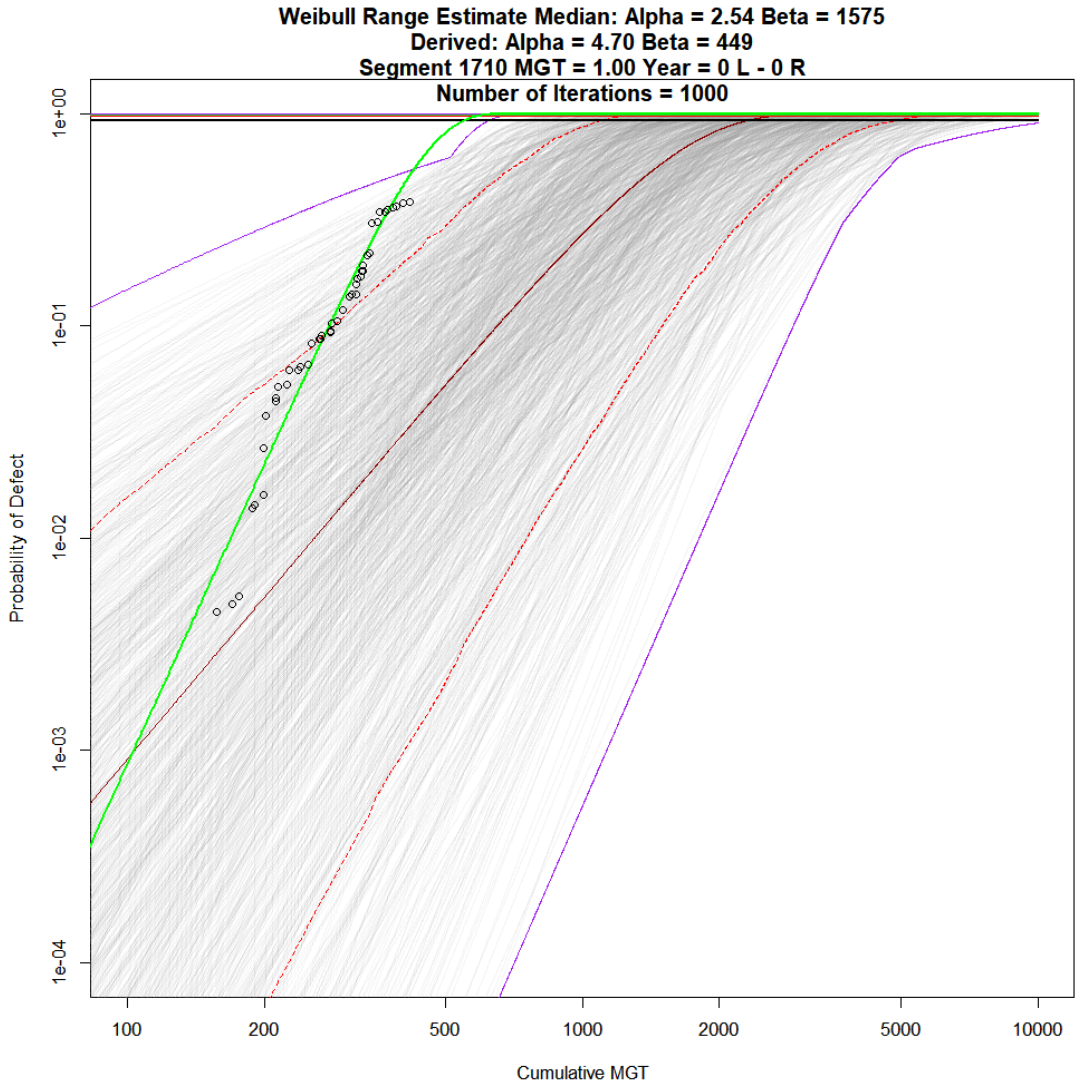


Figure 58: Source Data/Graph for Figure 59

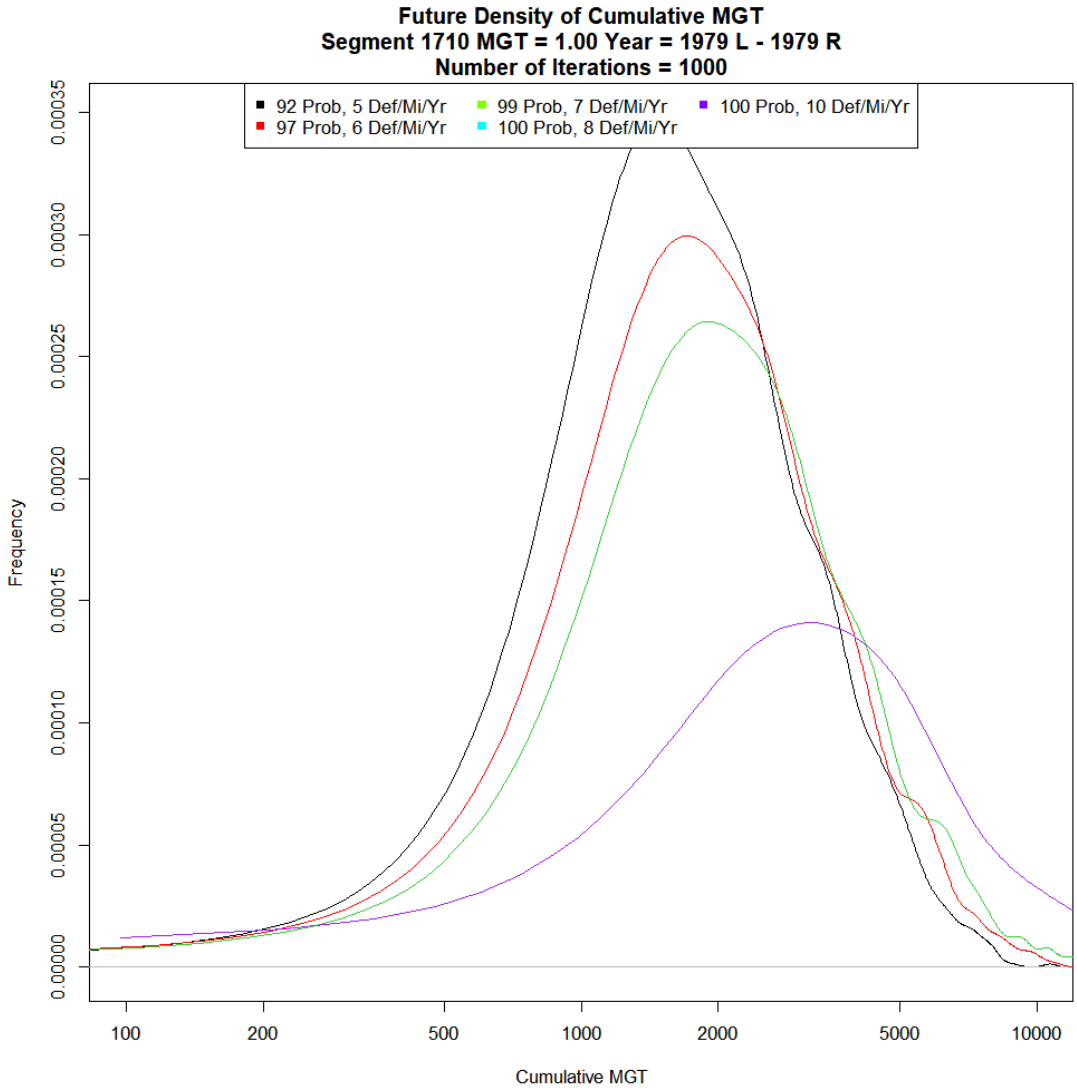


Figure 59: Density Graph showing combined densities due to Def/Mi/Year to Probability conversion issues

The next step of the process focused on combining individual segments together to develop a single probability density. The general idea is to provide a frequency prediction of the probability of a defect, or defect/mile/year, for a combined group of track segments, such a replacing a continuous 3 to 5 mile stretch of track[10]. Since each segment is slightly different than others, and thus pull different source data for the bootstrapping analysis, different Weibull sets are developed. Combining these Weibull plots was done by shifting the actual point values along the X axis such that all of the plots had their Current Cumulative MGT point at the same Cumulative MGT value on the X axis. This was done by determining the largest Current Cumulative MGT of the set of plots, and then shifting the plots with smaller Current Cumulative MGT to the right until they lined

[10] Which would correspond to a single rail train of rail.

up. Figure 60 shows the stacking and lining up of several Weibull plots, while Figure 61 shows the resulting density of the relative frequencies.

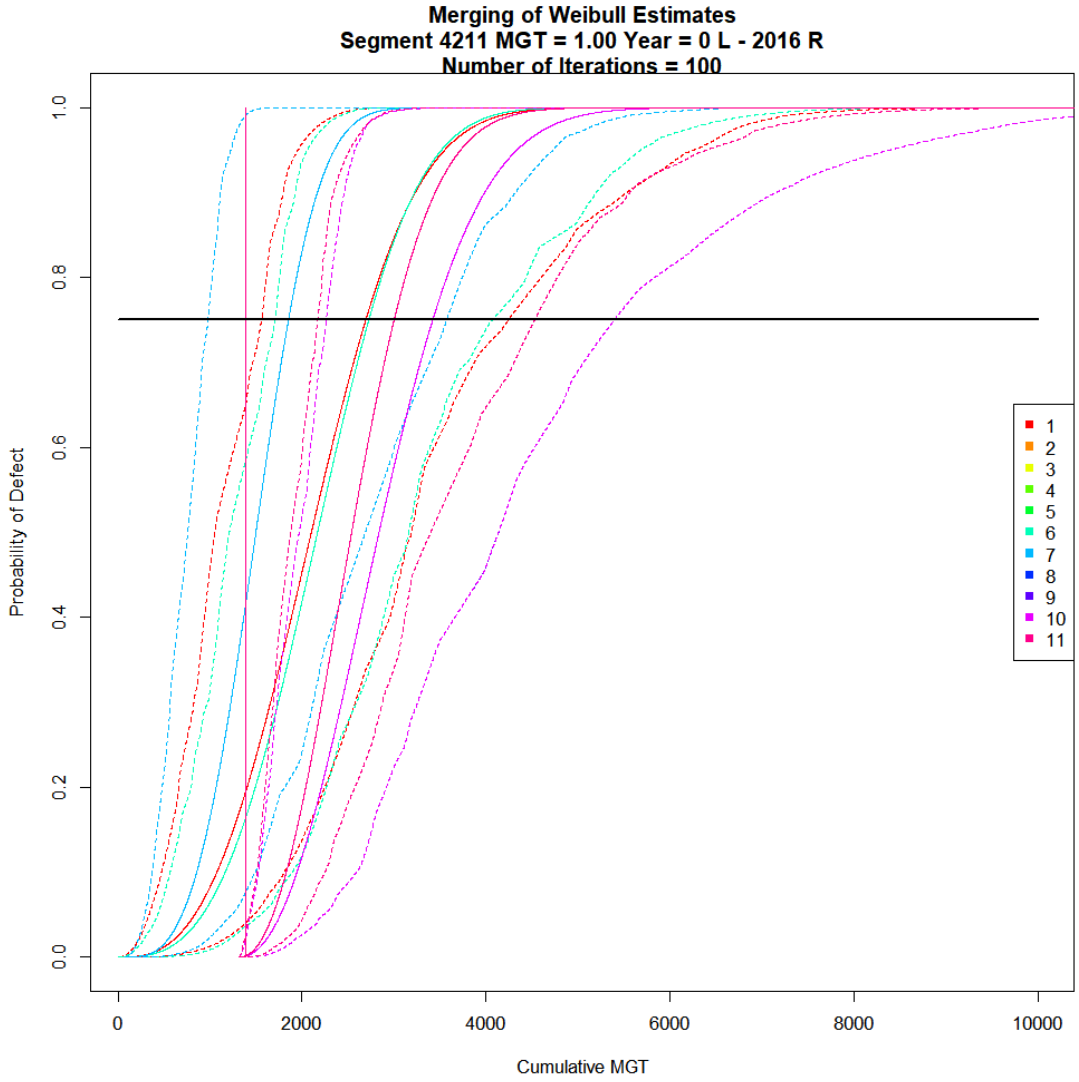


Figure 60: Graph showing combination of Weibulls and offsets by Cumulative MGT

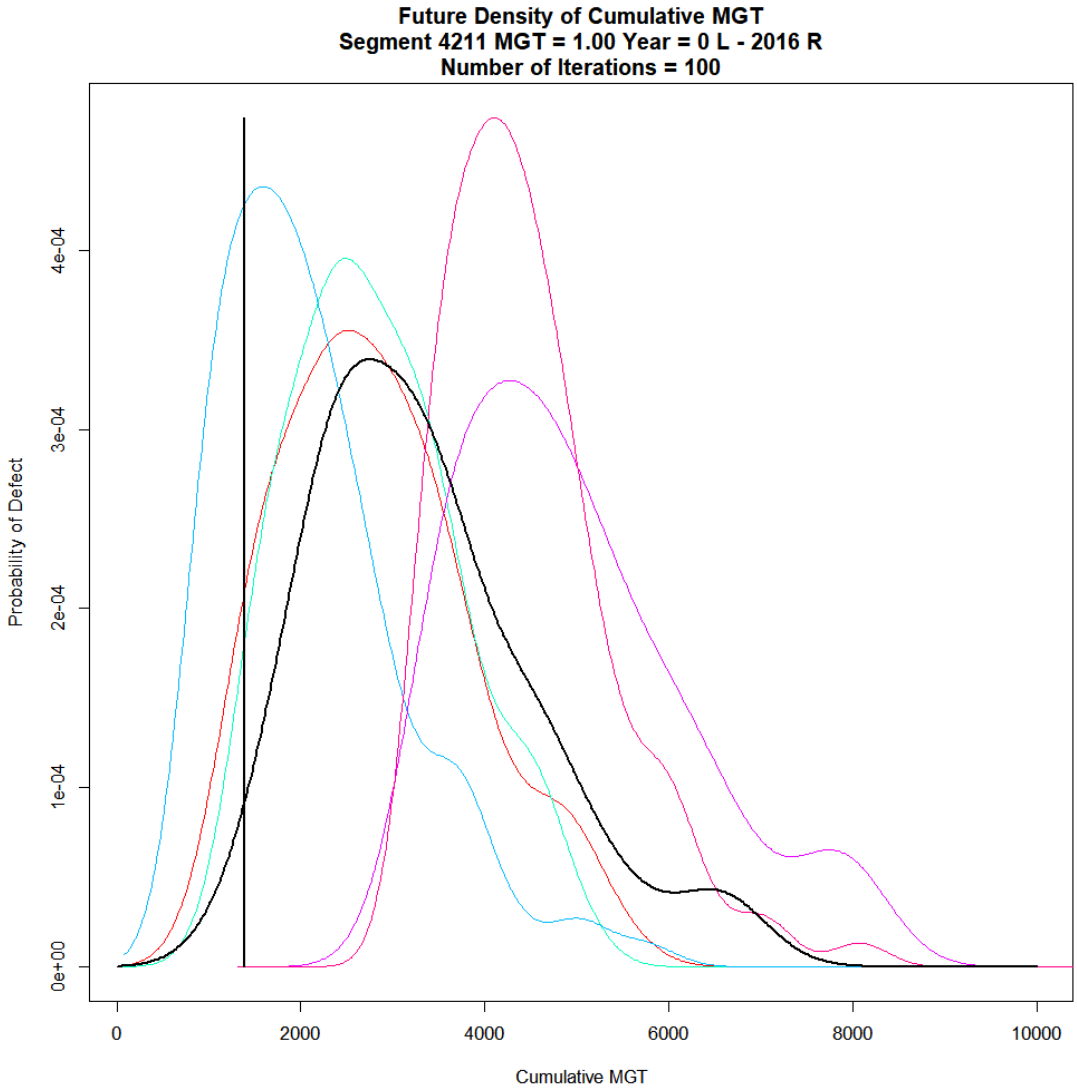


Figure 61: Weibull Combined Density with Weighted Expected Density

With the combined density graph, it becomes possible to schedule maintenance efforts based on the likelihood of defect occurrence, along a contiguous segment of track. This aids maintenance planning efforts by allowing scheduling to be done based on the structure of the maintenance itself, which tends to focus on replacing miles of track as a single job, instead of multiple sub-mile lengths in different maintenance events for preventative work.

3.3 Overall Results

This research has developed a way to help forecast the probability of a defect in a single rail, to include rails with a well-defined defect history and adjacent rails with a less well-defined defect history by using parametric bootstrapping in conjunction with available defect data. This now allows for the determination of a probabilistic distribution of forecasts which in turn allows for the development of reasonable approximations as to the failure rate of these less well-defined rails. By extending the parametric bootstrapping Weibull analysis approach to those rail segments with

limited defect history, this approach opens up significantly increased amounts of data to be used in maintenance planning efforts.

While the other machine learning methods discussed did not result in useful output, some information has been gleaned from them; the dominance of "no-defect" data, lack of unique identifiers on a rail-by-rail basis aside from milepost, and extensive time requirements for calculations.

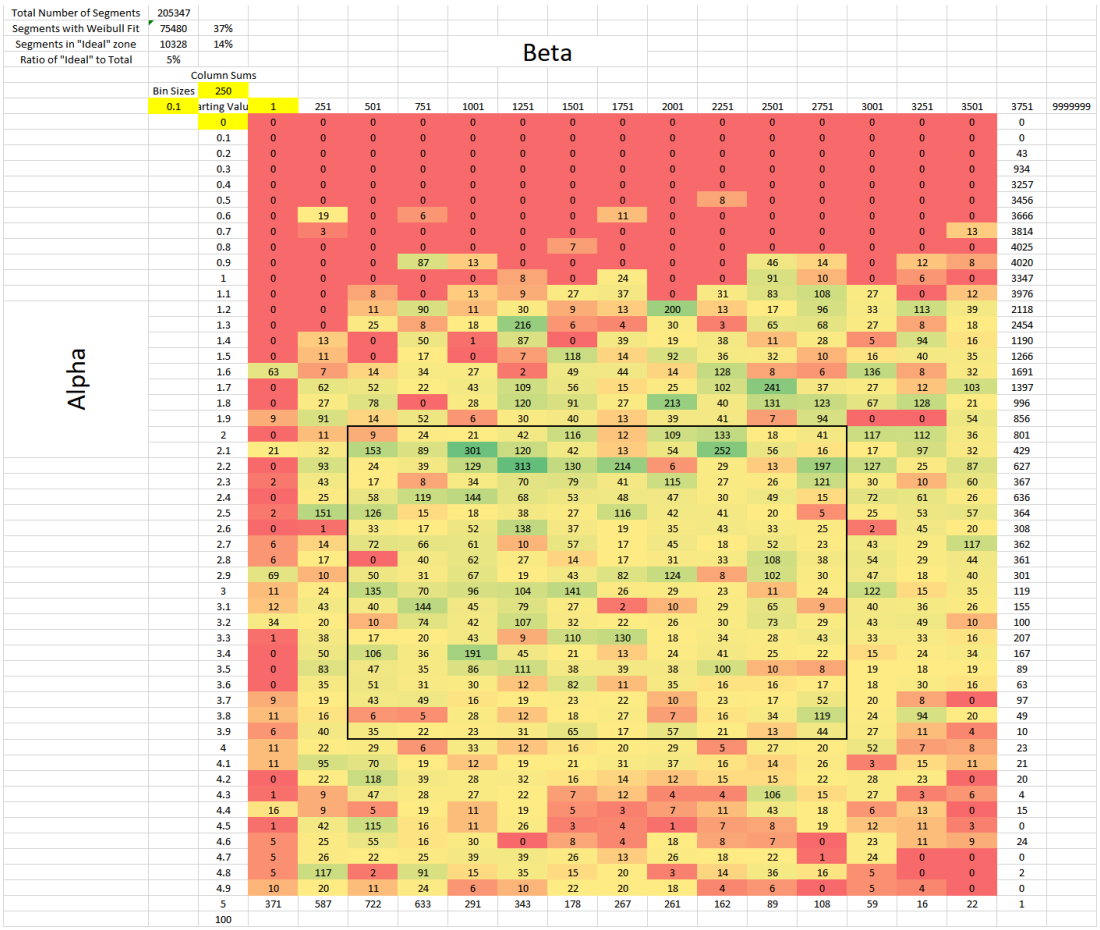
The "traditional" Weibull analysis has also been shown to be weak when applied to the data used, particularly due to the fact that often there was limited rail defect history data available for individual segments. Frequently the Weibull analysis would not work, due to a lack of defect data, but the reporting and structure of that data also caused issues in providing reasonable Weibull values.

However, the introduction of the bootstrapping approach to the Weibull analysis allowed for the extension of the rail life forecasting model to virtually all segments on a railroad with a broad distribution of rail condition and defect history.

As for concrete results, comparisons were done comparing the initial Weibull Analysis that resulted in some 10,000 usable track segments, and the results of the Parametric Bootstrapped Weibull analyses. Given that the original Weibull analysis resulted in Figure 62's distribution across Alpha and Beta values, it was determined that the best way of comparing the effectiveness of the Parametric Bootstrapped Method would be to measure the distance, from the "ideal" Weibull parameters, specifically Alpha = 3, and Beta = 2000 MGT (Equation 15). This formula was chosen as taking the Logarithm of Beta would bring down the range of values for Beta into the same general range as Alpha, allowing the distance formula to be more sensitive to changes in both values.

$$\sqrt{\left((\Delta Alpha^2) + \mathrm{Log}(\Delta Beta^2)\right)} \qquad \text{Equation 12}$$

In addition, the boundary of the "reasonable range" of Alpha and Beta parameters was considered, with a range of Alpha 2 to 4 and Beta = 500 to 3000. A second boundary encompassed the Alpha = 0 to 6, Beta = 0 to 4000 condition. As shown in Figure 63 through Figure 68, these two boundaries are marked on the graphs by two vertical lines.



Total Number of Segments	205347																	
Segments with Weibull Fit	75480	37%																
Segments in "Ideal" zone	10328	14%																
Ratio of "Ideal" to Total	5%																	
	Column Sums																	
	Bin Sizes	250																
	0.1	arting Valu	1	251	501	751	1001	1251	1501	1751	2001	2251	2501	2751	3001	3251	3501	3751	9999999
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43	
		0.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	934	
		0.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3257	
		0.5	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	3456	
		0.6	0	19	0	6	0	0	0	11	0	0	0	0	0	0	0	3666	
		0.7	0	3	0	0	0	0	0	0	0	0	0	0	0	0	13	3814	
		0.8	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	4025	
		0.9	0	0	0	87	13	0	0	0	0	0	46	14	0	12	8	4020	
		1	0	0	0	0	0	8	0	24	0	0	91	10	0	6	0	3347	
		1.1	0	0	8	0	13	9	27	37	0	31	83	108	27	0	12	3976	
		1.2	0	0	11	90	11	30	9	13	200	13	17	96	33	113	39	2118	
		1.3	0	0	25	8	18	216	6	4	30	3	65	68	27	8	18	2454	
		1.4	0	13	0	50	1	87	0	39	19	38	11	28	5	94	16	1190	
		1.5	0	11	0	17	0	7	118	14	92	36	32	10	16	40	35	1266	
		1.6	63	7	14	34	27	2	49	44	14	128	8	6	136	8	32	1691	
		1.7	0	62	52	22	43	109	56	15	25	102	241	37	27	12	103	1397	
		1.8	0	27	78	0	28	120	91	27	213	40	131	123	67	128	21	996	
		1.9	9	91	14	52	6	30	40	13	39	41	7	94	0	0	54	856	
		2	0	11	9	24	21	42	116	12	109	133	18	41	117	112	36	801	
		2.1	21	32	153	89	301	120	42	13	54	252	56	16	17	97	32	429	
		2.2	0	93	24	39	129	313	130	214	6	29	13	197	127	25	87	627	
		2.3	2	43	17	8	34	70	79	41	115	27	26	121	30	10	60	367	
		2.4	0	25	58	119	144	68	53	48	47	30	49	15	72	61	26	636	
		2.5	2	151	126	15	18	38	27	116	42	41	20	5	25	53	57	364	
		2.6	0	1	33	17	52	138	37	19	35	43	33	25	2	45	20	308	
		2.7	6	14	72	66	61	10	57	17	45	18	52	23	43	29	117	362	
		2.8	6	17	0	40	62	27	14	17	31	33	108	38	54	29	44	361	
		2.9	69	10	50	31	67	19	43	82	124	8	102	30	47	18	40	301	
		3	11	24	135	70	96	104	141	26	29	23	11	24	122	15	35	119	
		3.1	12	43	40	144	45	79	27	2	10	29	65	9	40	36	26	155	
		3.2	34	20	10	74	42	107	32	22	26	30	73	29	43	49	10	100	
		3.3	1	38	17	20	43	9	110	130	18	34	28	43	33	33	16	207	
		3.4	0	50	106	36	191	45	21	13	24	41	25	22	15	24	34	167	
		3.5	0	83	47	35	86	111	38	39	38	100	10	8	19	18	19	89	
		3.6	0	35	51	31	30	12	82	11	35	16	16	17	18	30	16	63	
		3.7	9	19	43	49	16	19	23	22	10	23	17	52	20	8	0	97	
		3.8	11	16	6	5	28	12	18	27	7	16	34	119	24	94	20	49	
		3.9	6	40	35	22	23	31	65	17	57	21	13	44	27	11	4	10	
		4	11	22	29	6	33	12	16	20	29	5	27	20	52	7	8	23	
		4.1	11	95	70	19	12	19	21	31	37	16	14	26	3	15	11	21	
		4.2	0	22	118	39	28	32	16	14	12	15	15	22	28	23	0	20	
		4.3	1	9	47	28	27	22	7	12	4	4	106	15	27	3	6	4	
		4.4	16	9	5	19	11	19	5	3	7	11	43	18	6	13	0	15	
		4.5	1	42	115	16	11	26	3	4	1	7	8	19	12	11	3	0	
		4.6	5	25	55	16	30	0	8	4	18	8	7	0	23	11	9	24	
		4.7	5	26	22	25	39	39	26	13	26	18	22	1	24	0	0	0	
		4.8	5	117	2	91	15	35	15	20	3	14	36	16	5	0	0	2	
		4.9	10	20	11	24	6	10	22	20	18	4	6	0	5	4	0	0	
		5	371	587	722	633	291	343	178	267	261	162	89	108	59	16	22	1	
		100																	

Figure 62: Distribution of Weibull parameters from Traditional analysis[11]

Figure 62 shows the results of the traditional Weibull analysis, where each entry is the count of Analyses which had parameters in that interval. The black bounding box covers the "realistic" Weibull parameter values, Alpha between 2 and 4, and Beta between 500 and 3000. One of the points taken from this visualization of the data is that the actual railway supplied data did not conform to the prior expectation the research group had; Weibull parameters were spread out and not concentrated around the "realistic" values that were expected. In addition, the ratios of Weibulls computed and Total Track Segments showed that for a common methodology, the Weibull could only be applied approximately 37% of the time, as well as only giving reasonable results about 10% of the time. This had the effect of changing the focus of the research from improving upon the Weibull analysis, to developing a method which could be applied to all track.

[11] Alpha increases going down, Beta increases going right, last values are catchalls for results exceeding 5 Alpha and 3750 Beta

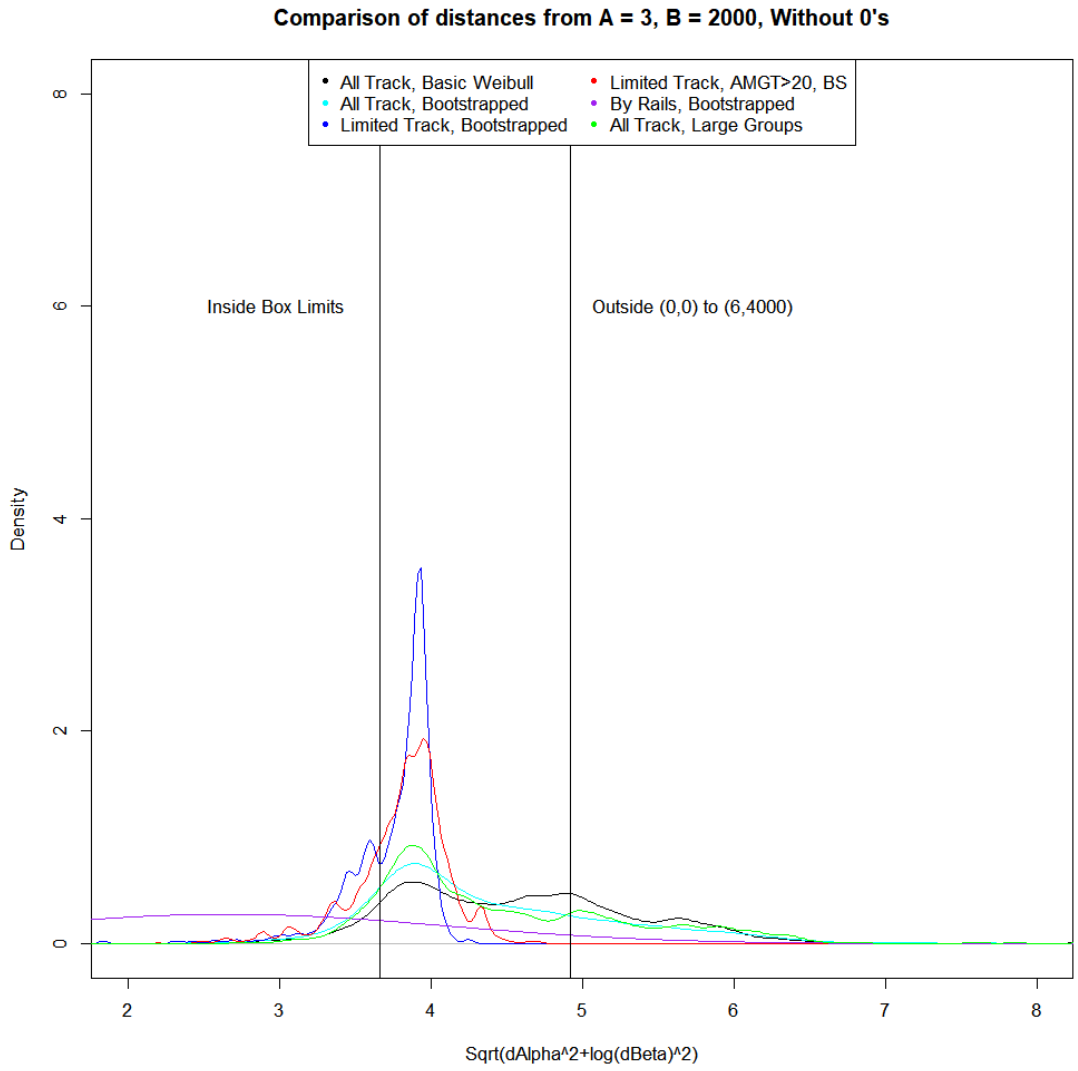


Figure 63: Density graph of Basic Weibull and Bootstrapped Weibull parameter results, Log Beta distance formula

One of the first things that is brought up by this graph is the spike in frequencies right outside the idealized Weibull parameters. The two biggest spikes are from the Bootstrapped data, working from the data that was limited to be within 1 to 10 for Alpha, and 500 to 10,000 for Beta. By Rails indicates source data that was split into individual rails, essentially doubling the number of track entries, while keeping the same number of defects. The Large Group indicator stands for the use of 100-mile long segments of track instead of the individual segments.

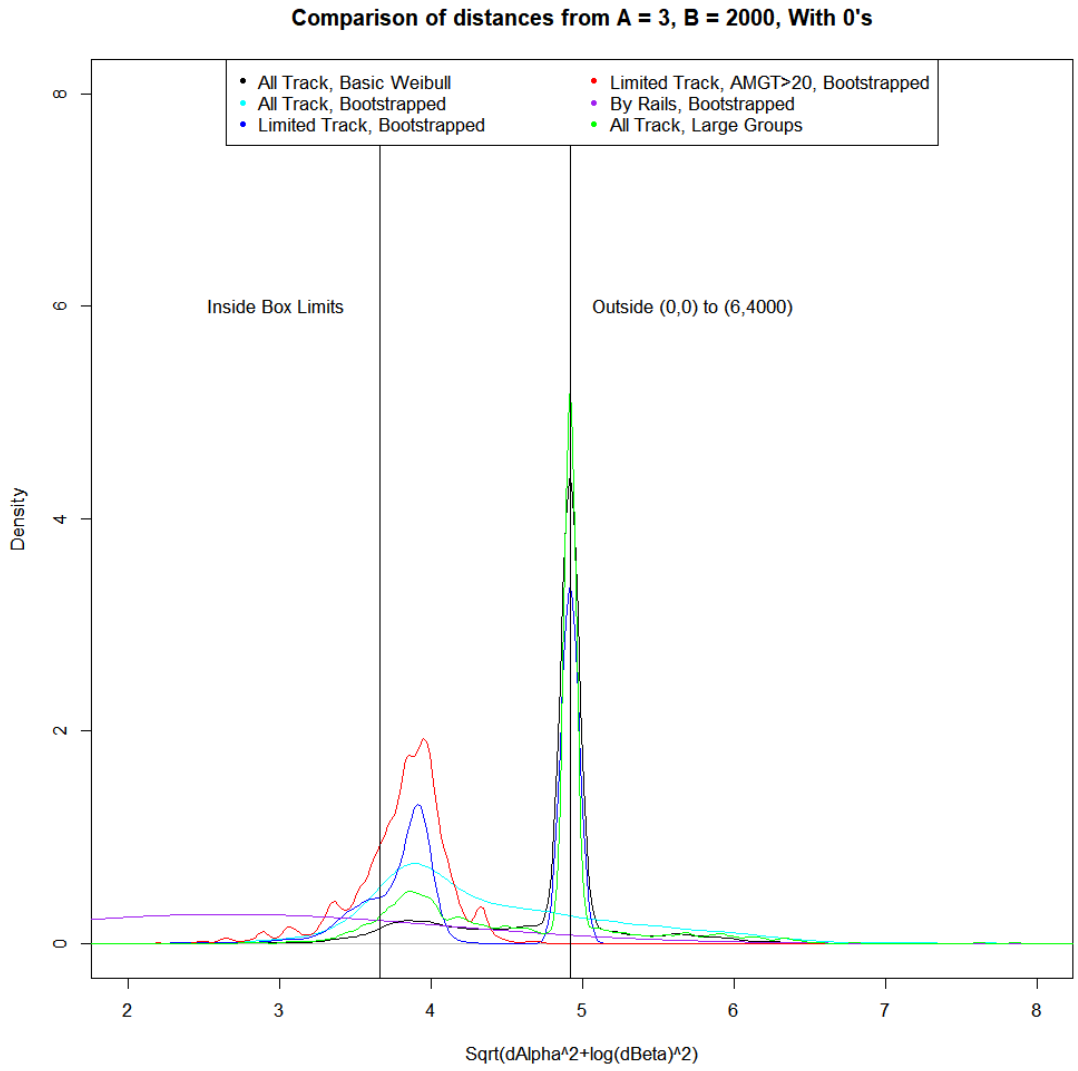


Figure 64: Density graph of Basic Weibull and Bootstrapped Weibull parameter results with 0's counted, Log Beta distance formula

Compared to the "No Zero" graph preceding this, the major spike in density has shifted to the outer boundary mark. As the only change is the inclusion of "0,0" entries, this showcases how skewed the data is to this category of non-data.

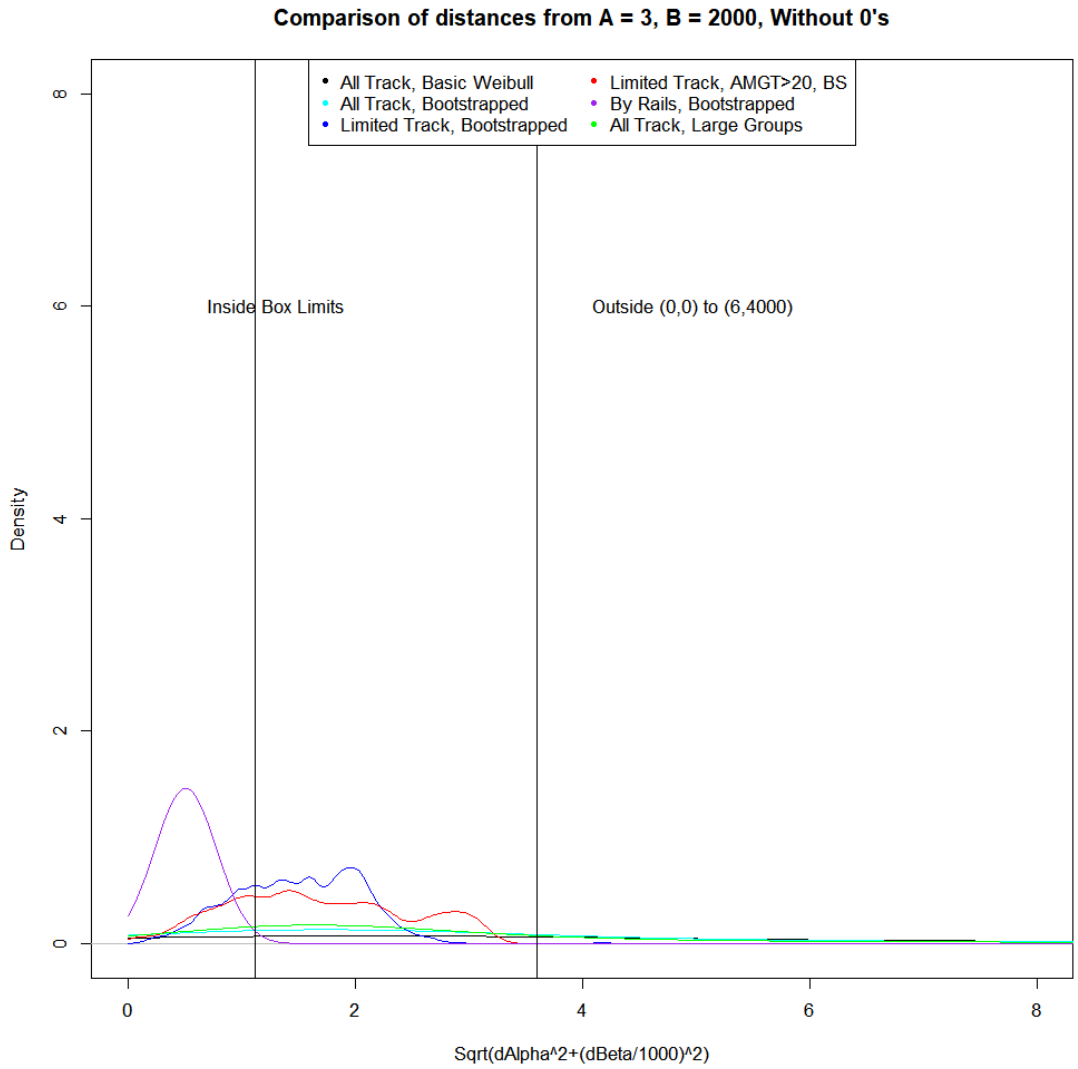


Figure 65: Density graph of Basic Weibull and Bootstrapped Weibull parameter results, Reduced Beta Scale distance formula

The main difference between the following images is that the way the X axis is calculated has been rescaled. From the start, it was apparent that using the plain Beta value would skew the X axis, causing any change in Alpha to be inconsequential compared to the magnitude of Beta. Figure 73 and Figure 74 use the standard distance formula, but takes the Log of the Beta difference first in order to rescale it closer to the Alpha parameter. In Figure 75 and Figure 76, the difference from the Ideal Beta is divided by 1,000, a basic linear scale of the Beta parameter. Figure 77 and Figure 78 alter the whole equation to use the absolute value of the difference, instead of the square root of the square of the difference, but keeps the square root function on the Beta parameter.

Comparing all of these alternative scales for the X axis, the same general trends exist: that including 0 parameter entries showcases how skewed the data is; and that the Parametric Bootstrapping Weibulls result in median Weibull fit parameters closer to the "ideal" parameters, but not strictly within the bounds.

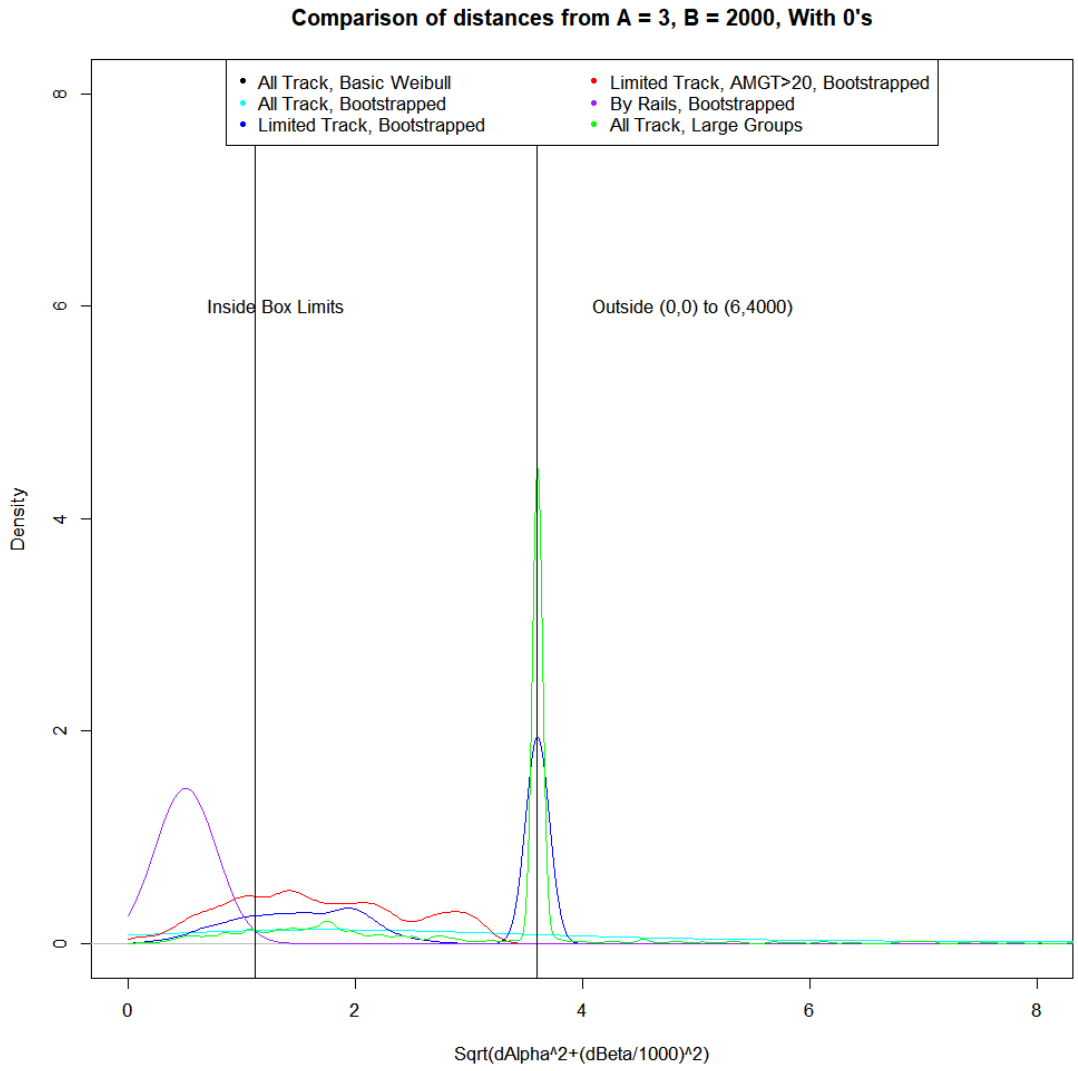


Figure 66: Density graph of Basic Weibull and Bootstrapped Weibull parameter results with 0's included, Reduced Beta distance formula

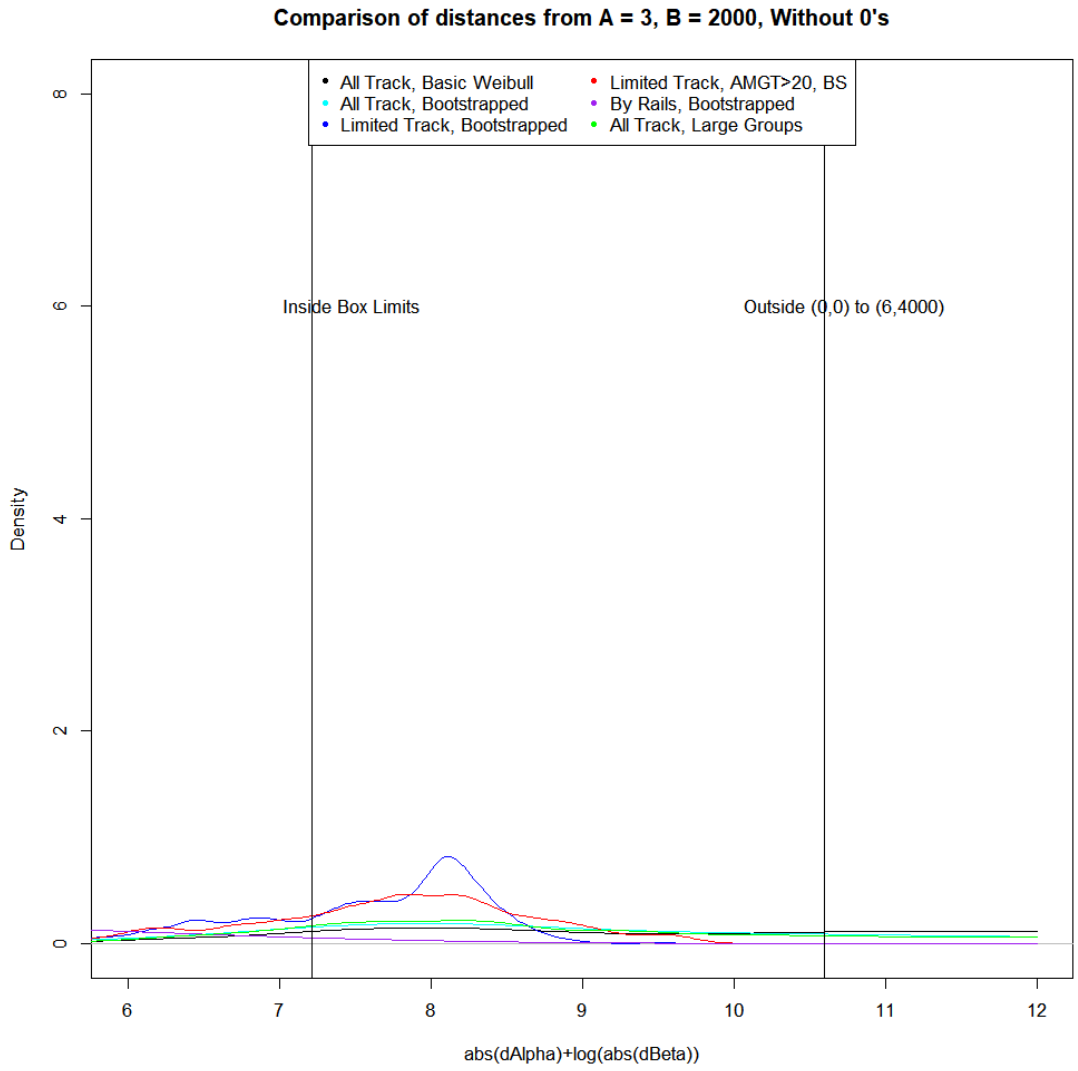


Figure 67: Density graph of Basic Weibull and Bootstrapped Weibull parameter results, Absolute Log Beta distance formula

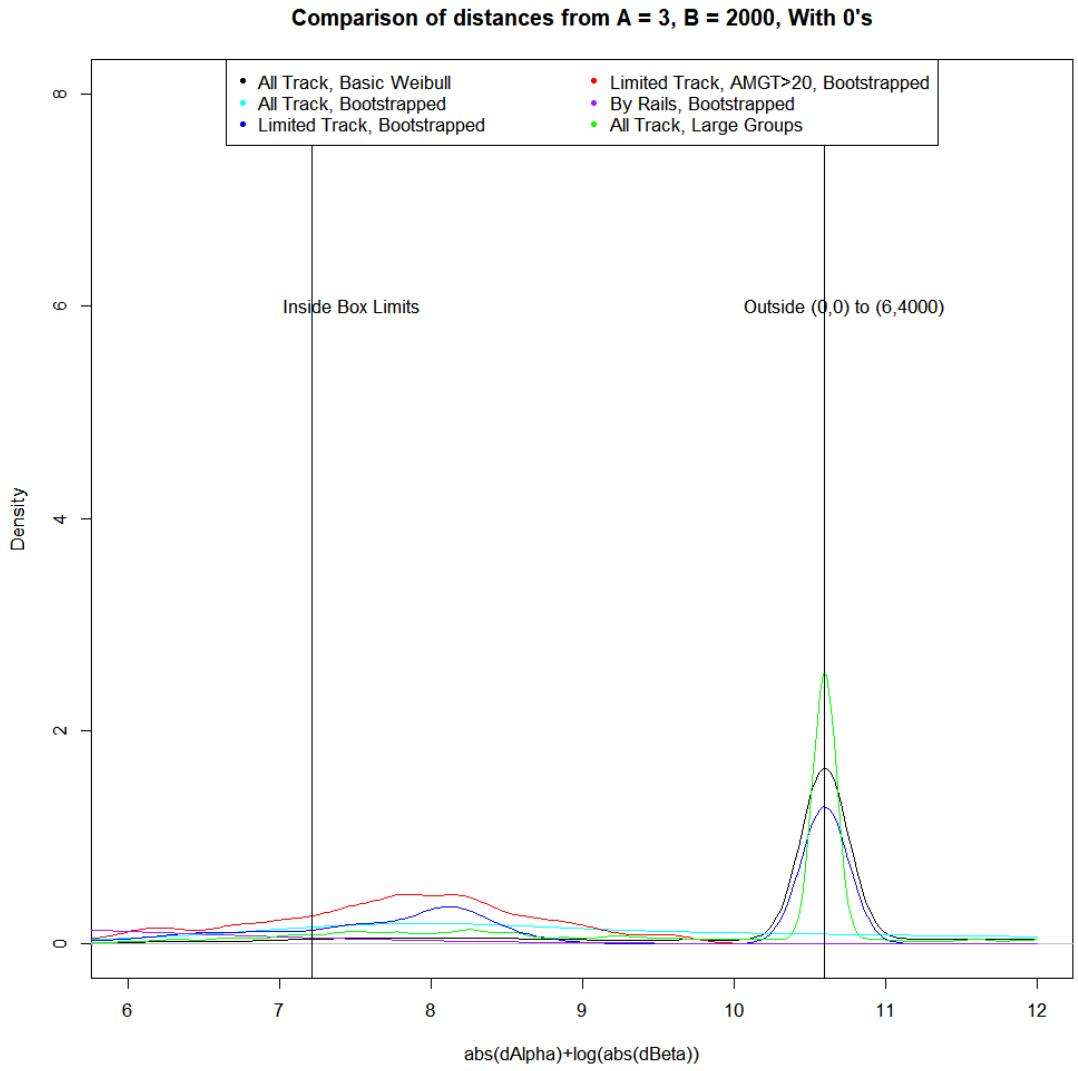


Figure 68: Density graph of Basic Weibull and Bootstrapped Weibull parameter results with 0's included, Absolute Log Distance formula

CHAPTER 4 DEVELOPMENT OF PREDICTIVE MAINTENANCE PLANNING METHODOLOGY

4.1 Basic Weibull vs Parametric Bootstrapped Weibull Analysis

One of the biggest differences between the normal Weibull method and the Bootstrapped method is that the bootstrapped method provides reasonable estimates of the rate of defects for track segments that do not have any significant prior defect data. This allows far more track segments to be analyzed, and to be accounted for in maintenance planning efforts. Adding to this, there is a range of values to use in the prediction, instead of a single value; it now becomes possible to estimate a "best case" and "worst case" scenario.

4.2 Outline of Parametric Bootstrapping Weibull Analysis

From start to finish, a step-by-step outline of the Parametric Bootstrapping Weibull Analysis approach is provided in the following:

1. Acquire Data, including but not limited to Rail, Defect, and MGT data.
2. Clean and Format the data for analysis
3. Divide the rail data into homogeneous segments
    i. Divide rail data into one-mile segments.
4. Perform a Weibull analysis and calculate 2-Parameter Weibull values (Alpha and Beta) for all track segments with sufficient data
5. Start the Parametric Bootstrapping Analysis
    i. For each rail segment in the data, find all similar track within designated parameter bounds, such as +/-10% of Rail Weight, that have 2-Parameter Weibull values
    ii. Fit a distribution to the Alpha and Beta frequency distributions of the similar track
    iii. Randomly generate pairs of Weibull values from the fit distribution
    iv. Calculate the Median and designated boundary values, such as every 10%.
    v. Plot these generated Weibull value pairs, overlaid on each other (Use a low alpha channel value when plotting so that overlaid lines show up, while individual lines are not so obvious)
    vi. Take "slices" of the Weibull plot data, such as all the Weibull lines points passing through a designated Cumulative MGT or Probability of a Defect
    vii. Plot these Slices as a density graph
6. Depending on selected "slices" (Horizontal or Vertical), graphs can be used to predict when certain thresholds will be reached.
7. By taking the area under the density curve, user can calculate the probability of reaching a certain threshold, as the density curve area is equal to 1.

4.3 Example of Parametric Bootstrapping Weibull Analysis

This example will start from the point after the data has been cleaned and combined into suitable datasets. The processes involved were written based on the datafiles detailed in Chapter 3, but can be changed to suit whatever dataset the user has.

Now, given that the datasets have been cleaned and combined, it is now possible to calculate the Weibull parameters for some of the data. The process will go through an example of track data, collect the corresponding defect data, and then calculate the Weibull fit.

For this example, the Bootstrapped Weibull output of an example track segment will be used to give recommendations on maintenance planning. The source data is shown in the following two figures, Figure 69 and Figure 70, which display the original data, the bars, and the parametric fit, the dotted line, from which the Parametric Bootstrapping will pull values from.

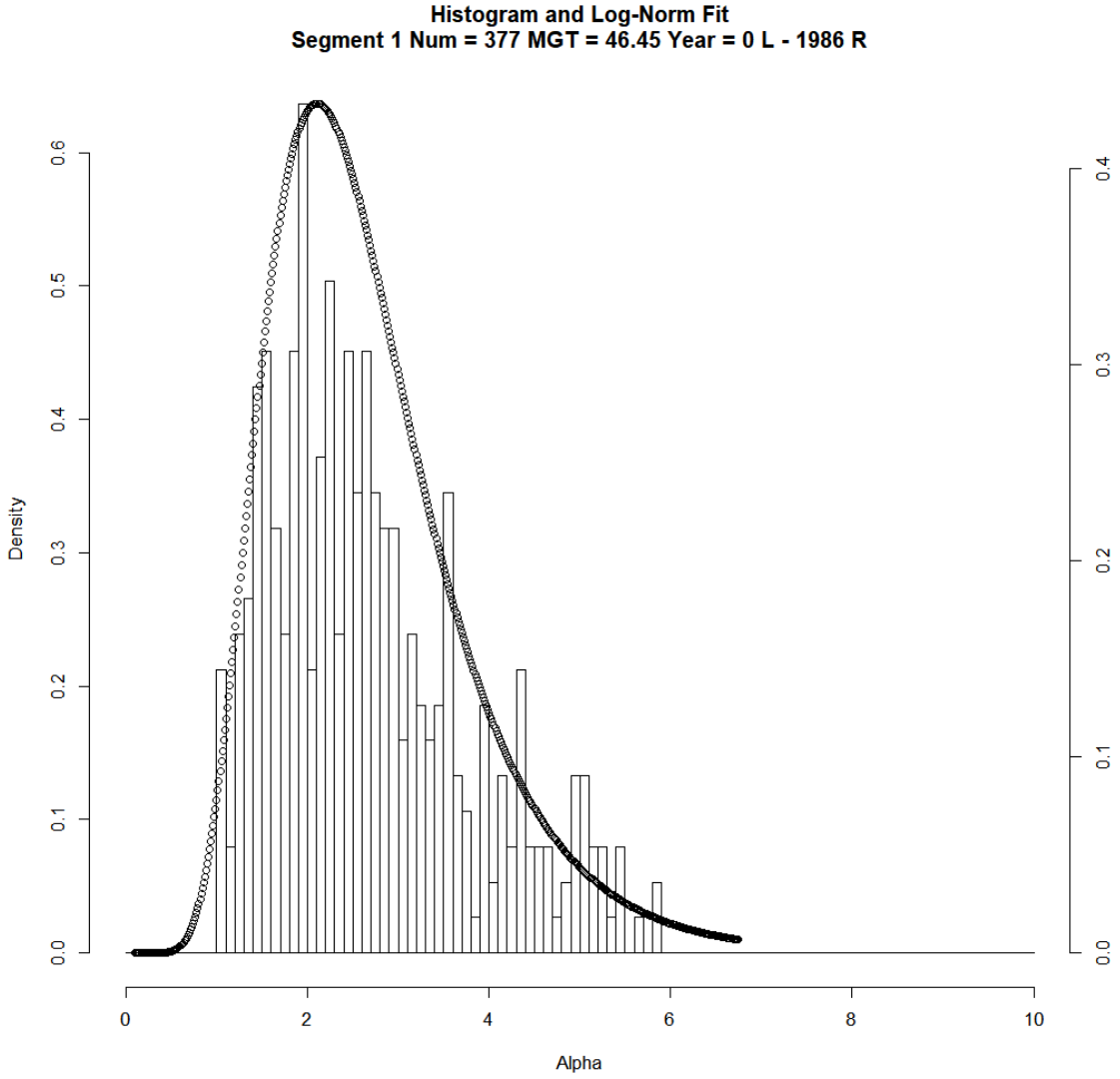


Figure 69: Example Problem Alpha Distribution Source

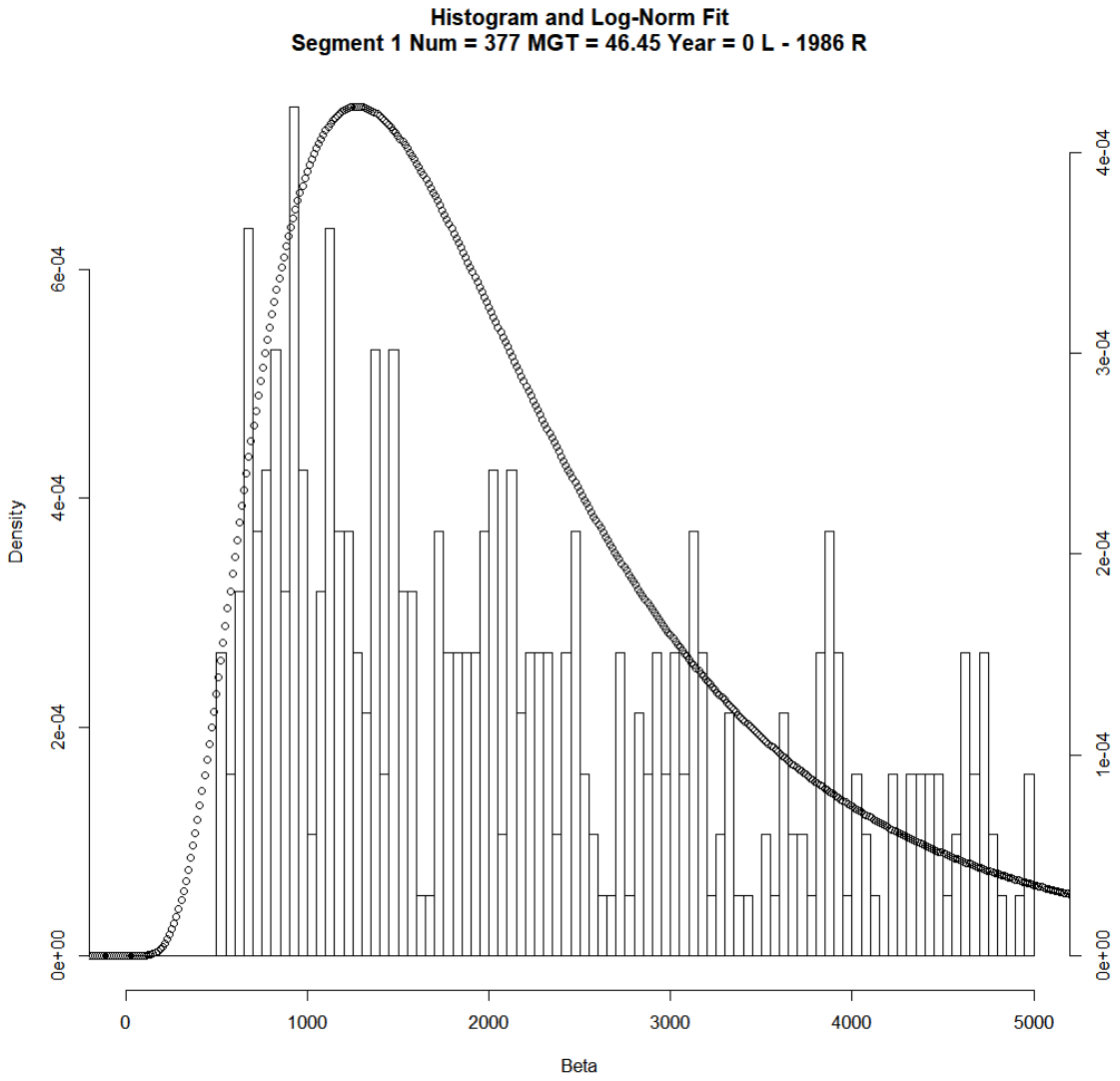


Figure 70: Example Problem Beta Distribution Source

This data is then used to generate the Bootstrapped Weibull plot, Figure 71, by pulling, at random, 1000 values from the Alpha and Beta distributions. If these values fall outside of the expected range, 1 to 6 for Alpha, and 500 to 5000 for Beta, new values are pulled at random from the distributions to replace them; on the rare event the replacement is also outside the boundary, the process is repeated. Figure 82 shows the horizontal density cuts taken at a range of defect rates (defects/mile/year) for this example. Note the shift to the right, with the maximum frequency going from around 300 Cumulative MGT for 1 defects/mile/year, to 600 Cum. MGT for 2 d/mi/yr, 900 Cum. MGT for 3 d/mi/yr, 1000 Cum. MGT for 4 d/mi/yr, and finally 1100 Cum. MGT for 5 d/mi/yr. These points of maximum frequency are directly linked to the median Weibull fit as shown by the red line in Figure 71. By taking the area of each cut between two Cum. MGT values, or more specifically 0 and the Cum. MGT of interest, it is possible to report the likelihood that the rail has reached that threshold of defects per mile per year.

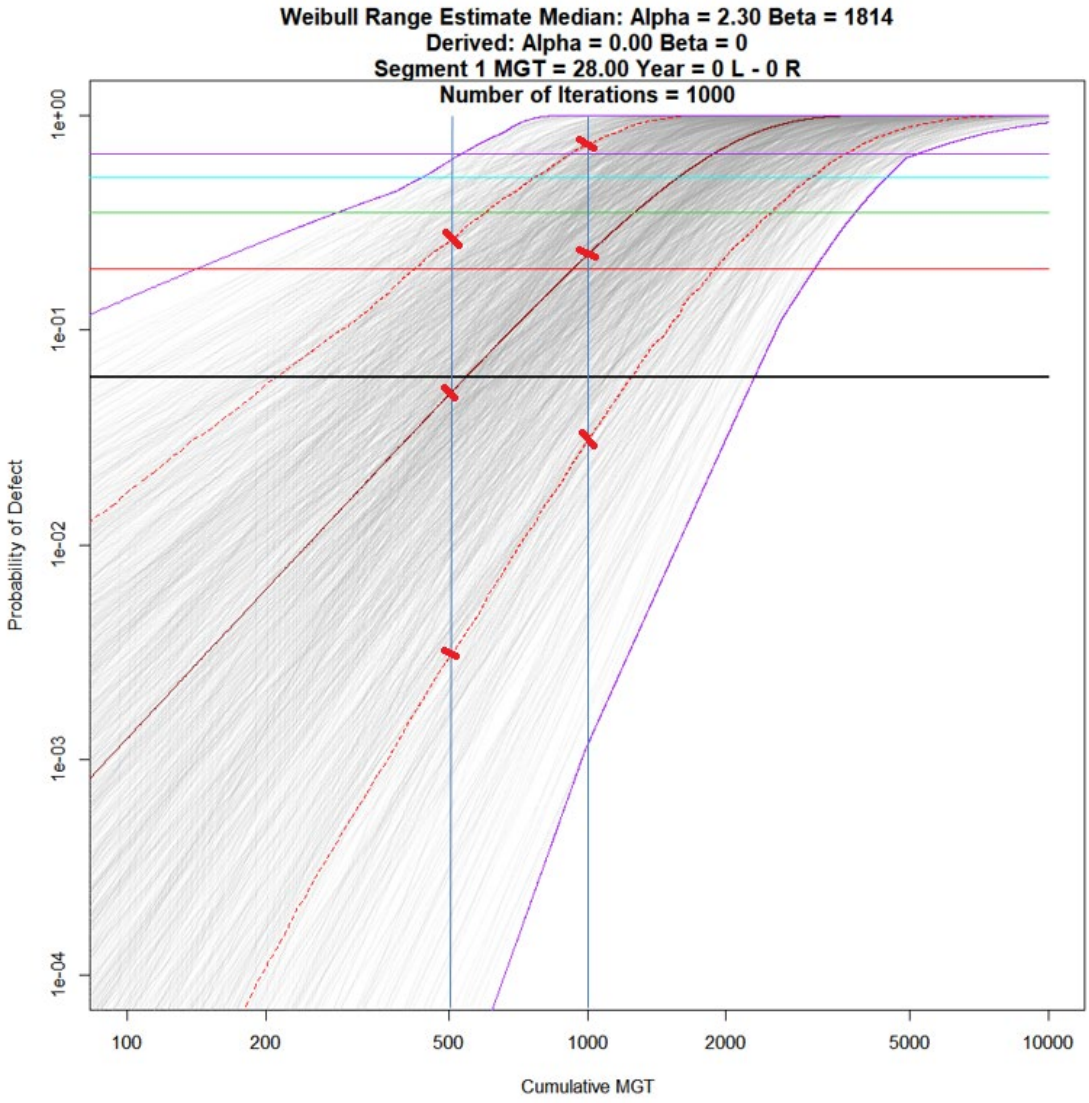


Figure 71: Bootstrapped Weibull overlay for Example

Where:

The Vertical Blue Lines are the Cum. MGT of interest
The Dark Red Line is the Median Weibull Fit
The Dashed Red Lines are the +/-40% boundaries
The Purple Lines are the extreme value lines
The Red Marks are indicating the intersections of interest between the Cum. MGT lines, and the probability intervals

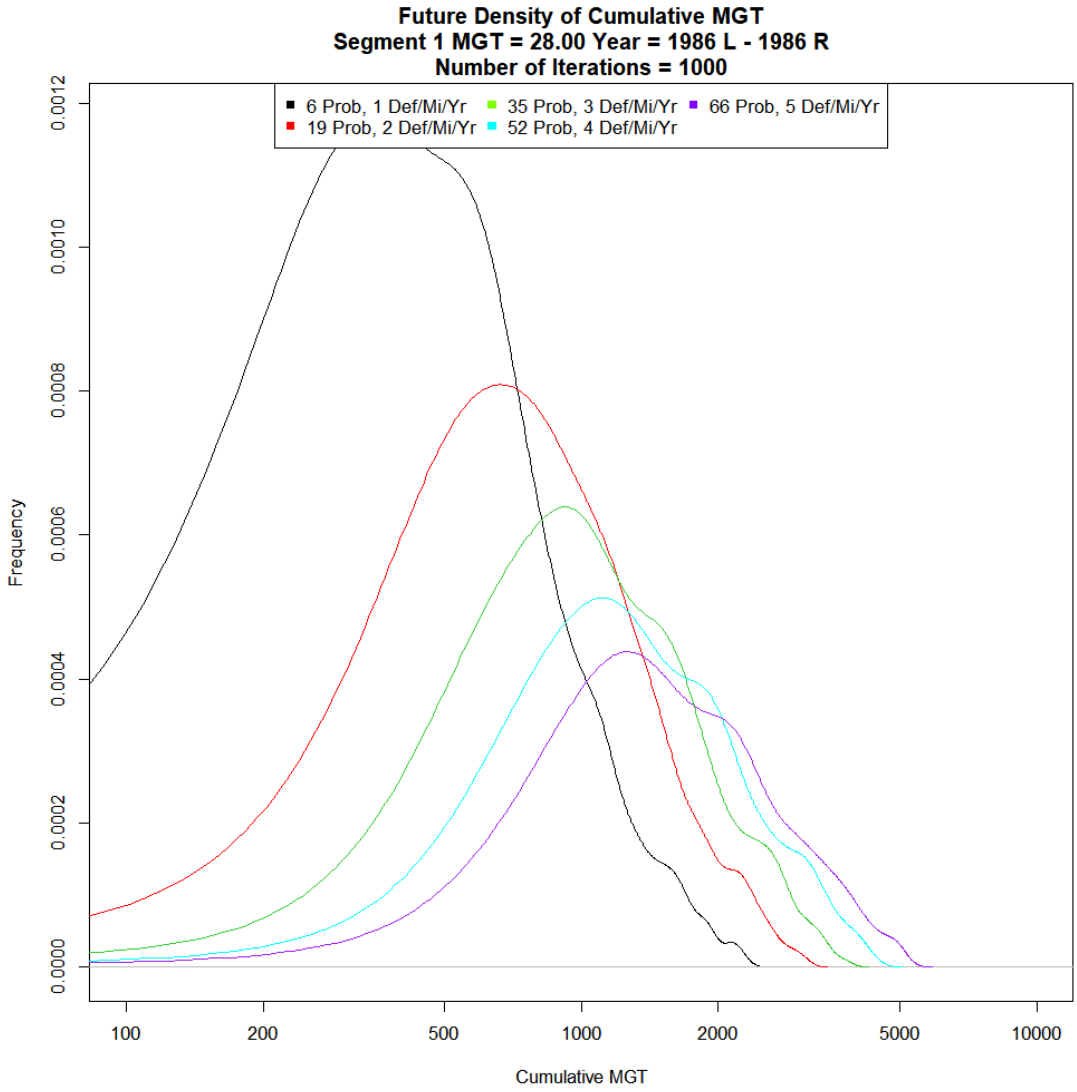


Figure 72: Density Cut of Bootstrapped Weibull Example

For the purpose of this example, let us assume that the track is currently at 200 Cumulative MGT, and the user wants to forecast what is the likely defect rate at 500 Cumulative MGT. From Figure 81, going up from the 200 Cum. MGT mark, it intercepts the 90% mark, the lower dashed red line, first at approximately 0.0001, or 0.001% probability of a defect in a rail. Next, it "hits" the Median (50%) mark, the solid red line, around 0.002, or 0.2% probability of a defect in a rail. Going further, it then intercepts the 10% mark, the upper dashed red line around 0.02, or 2% probability of a defect in a rail. These three probabilities give the user the High, Median, and Low expected defect occurrence values; given the Bootstrapped distribution. Thus, it can be assumed that 90% of all similar track segments will have a probability of a defect in a rail that exceeds 0.001%, that 50% of all similar track segments will have a probability of a defect in a rail that exceeds 0.2%, and that 10% of all similar track segments will have a probability of a defect in a rail that exceeds 2%. Now, looking at the 500 Cum. MGT probabilities, the 90% value is at 0.005, or 0.5%, the 50% value at 0.02, or 2%, and the 10% value at 0.5, or 50%.

Once these probabilities have been determined, it is possible to combine them with maintenance cost to help make maintenance decisions. For the purpose of this example, assume that the expected cost of a defect that results in a derailment is $1,000,000, and that maintenance efforts will cost 20,000 dollars if applied right now, and would effectively prevent any defect from occurring for the rest of the time period. From the difference in probabilities over time, it can be seen that there is a strong probability of a defect in the 10% case, specifically a 48% chance of a defect in a rail, while the 90% case is relatively low, 0.5%, and the median case shows a 1.8% chance of a defect in the given time period. Applying the defect and maintenance costs, as illustrated in Table 22, shows that at the higher probability range; it is worth repairing the defect.

Table 19: Data used for the Example Problem

Cost of a Defect	$1,000,000		
Cost of Maintenance	$20,000		
	90%	50%	10%
Initial Chance of Defect	0.0001	0.002	0.02
Ending Chance of Defect	0.005	0.02	0.5
Probability of a defect over the period	0.0049 or 0.5%	0.018 or 1.8%	0.48 or 48%
Estimated Cost given No Maintenance- obtained by multiplying probability of a defect occurring by the defect cost	$4,900	$18,000	$480,000
Should it be Repaired now?	No	No	Yes

This repair/wait method depends on the user defining which case should be used as the decision factor. Bypassing this choice, it is possible to develop an "expected value" that takes into account all of the possibilities, resulting in an expected cost if no maintenance happens. Table 23 continues from Table 22, adding in the Weighted Probability of Occurrence, which comes from the Low, Median, and High expected occurrence rates. Assuming the Low and High occurrence rates have a 20% likelihood (based on +/-10% around their 10% from the extremes definition), and the Median is 60% likely, the Probability of a Defect multiplied by this likelihood results in the Weighted Probability of a Defect over the period.

Table 20: Expected cost based on expected probabilities

Probability of a defect over the period	0.0049 or 0.5%	0.018 or 1.8%	0.48 or 48%
Weigh of occurrence	20%	60%	20%
Weighted Probability (Previous two rows multiplied)	0.00098	0.0108	0.096
Weighted Cost	$980	$10,800	$96,000

Expected Cost (Sum of Weighted Cost)		$107,780	
Cost of Maintenance		$20,000	
Decision		Repair Now	

Of course, a more accurate result can be obtained by splitting the data into more breaks, but this takes more computational time, nor is suitable for an example in this paper. However, such expansion is relatively straightforward and follows the approach shown here.

CHAPTER 5 CONCLUSION AND RECOMMENDATIONS

5.1 Review of Results

As discussed in this report, the Parametric Bootstrapping approach was developed to alleviate many of the problems identified with the traditional Weibull analysis approach. This included calculation of unrealistic Weibull parameters for a very large number of segments. The Parametric Bootstrapping methodology was able to address these shortcomings and create realistic output forecasts for virtually all the 200,000 segments analyzed. This was accomplished using the data from a "core" group of approximately 8000 segments which generated 'reasonable" output parameters and realistic forecast results.

From the 8,000 "traditional" Weibull plots, it became possible to create Parametric Bootstrapped Densities for all track. i.e. all 200,000 segments, with minor exceptions[12]. From these densities, the model is then able to generate meaningful and useful outputs, such as estimation of when track will reach a certain threshold, such as a rail replacement threshold defined in terms of defects/mile/year. Similarly, the risk or probability of a defect appearing "next year" based on the rail's median Cumulative MGT can be calculated. These can also be used when track segments are grouped together, providing an estimation of the entire track group's properties, which can then be used by maintenance programs to prioritize areas that will exceed designated thresholds and require maintenance actions.

This analysis approach was then applied to a large subsegment of the 200,000 segments available, however because of the sheer size of the railroads, not all segments were so analyzed.

5.2 Review of New Methodology

The Parametric Bootstrapping method expanded upon and went beyond the traditional Weibull analysis by allowing confidence interval-based estimations of when defects will occur. This now allows more flexibility, such as deciding on a "minimize risk" approach instead of using the median value, as well as warning, such as showing that other similar track have quicker generation of defects, in track maintenance planning. It also shows how multiple Weibull plots can be combined to produce a weighted resulting plot, further expanding the use of the methodology into large scale analysis.

The Parametric Bootstrapping method, and the resulting density results and corresponding defect prediction analyses, e.g. prediction of Cumulative MGT when certain defect thresholds will be reached, represents a new set of tools that will assist railroads in their maintenance efforts. By expanding the traditional Weibull plot into a more comprehensive defect probability density "map", projections of rail replacement points are more accurate and extensive. These projections tie into the already used metrics such as defects per mile per year, which are currently used by the railroads in maintenance planning efforts. By doing so, it is expected that the methodology is more

[12] Exceptions include track which is so unique as to not have any other track be similar in any category, usually as a result of odd data inputs, or to have the "similarity" so narrow that no other track can be matched.

readily accepted by maintenance planners, as it directly corresponds to metrics they have already used, thus avoiding a stumbling block of many new methods which require extensive retraining or drastic changes to operational behavior. By using the densities, predictions can be made as to when the track in question will reach a certain maintenance threshold, both for individual segments or in combined segments thus giving maintenance planners additional flexibility in their local, regional and system maintenance plans.

5.3 Review of Lessons Learned

The first lesson learned from the research is that while railroads have been using the traditional Weibull method for decades, they have been using it only on a "spot" basis, since the traditional Weibull method fails to find solutions for a large percentage of railroad track. This means that maintenance planning efforts are being extrapolated based upon those tracks that do have Weibull results. This often results in generalized time (or MGT) based "rules" for the railroad to use in their replacement decision[13].

The second is that the railroads need a unified system of reporting defects and where in the track is the defects are located. While the current system works well for rail defects, a there are still shortcomings such as due to missing historical data or incomplete current data entry values. This can cause problems in the computational phase, which is what occurred with traditional Weibull analysis. Methodologies such as the Bootstrapping approach allows for the corrections of this class of missing data using an accurate data subset to develop parameters that can be used in a more extended system analysis.

As the data is used for more extensive computational work than ever before, it becomes important that said data is written down in a uniform accurate and comprehensive way, in order to make sure any unique entry is not discarded because it is not known how to translate it into an already established category. This is of major importance if railroads start hiring people to do the analyses which are not aware of railroad terminology, and thus determine similar things to be different because of how they are recorded.

5.4 Recommendations for Future Research

One potential avenue for future research is revisiting machine learning. Now that each piece of track can have a Weibull plot developed, they could be used as input into a methodology, such as image processing and prediction. Instead of depending on something like the Weibull function, an algorithm could be developed to "complete" the right half of a Weibull plot from a given left half. The algorithm wouldn't be given the Weibull function, but instead trained on complete images, then told to "solve" for the missing parts. Another possibility would be something like a Markov Chain Monte Carlo simulation, using nodes representing the probability of a defect, the probability of finding it or having an in-service fault, corresponding costs associated with repair or derailment

---

[13] For example: replacing track every 5 years, or every 500 MGT regardless of how worn out it is.

cleanup, etc. By using this, variations on the impact a minor change to one variable, such as the probability of a defect, can be expressed as actual dollar value savings (or costs).

5.5 Recommendations for Data Collection

While the processes shown here were applied to existing datasets to ease their introduction into industry at a quicker pace, having more details and complete data can help develop more theoretical applications which then can be reduced into applicable industry standards. Mainly, the acquisition of rail history needs to be improved; as it is now, there are questions as to the actual cumulative MGT of the rail, when the rail was laid, history of the rail if it was a re-laid rail. By treating each rail like a locomotive or wagon, and keeping a detailed list of where it is, when it was last maintained, the particulars of defects found in it, etc. it becomes possible to build a detailed history of the rail which aids in any predictions based on the rail. In addition, more frequent scanning of the rails, to eliminate the "defect cliffs" where sometimes dozens of defects are detected at once, will be needed.

5.6 Conclusion

This research started with an extensive data set from a large class 1 railroad with approximately 30,000 miles of track, and an undefined goal of improving on the Weibull methodology. Through multiple analyses, it was shown that the traditional Weibull method does not work as well as desired when given the wide variety of data present on a railroad. Instead of "improving" Weibull, by increasing its accuracy or explanatory variables as originally conceived, the focus shifted to broadening the analysis approach to deal with track segments with limited or inadequate data and thus expanding the application of Weibull to more track segments. This was accomplished through the use of Parametric Bootstrapping in order to take a limited subset of "good traditional" Weibull results and creating density parameters for track that allowed for the extension of the analysis approach. This changed the balance of information; now all track had some probabilistic level of Weibull parameter values and allowed for the calculation of the rail defect failure rate on a probabilistic basis, using these initially assumed values. This also opened up new avenues for establishing metrics to be used by railroads in their maintenance efforts; such as by using the density graphs to obtain estimations as to when the track would reach certain thresholds of defects per mile per year. Converting these estimations into metrics railroads are already familiar with, acceptance of the new methodology by railroad personnel will be easier than the case where entirely new metrics are used which are not familiar to those in charge of maintenance planning. Finally, this approach allows for the combining of several track segments together instead of just focusing on one track segment at a time. This allows maintenance planning for multiple levels of a large railroad to include individual segments, combined territories, or the entire railway system so as to permit more efficient use of resources, allowing a for greater reduction in risk for the same cost.

REFERENCES

1. United States Department of Transportation, Bureau of Transportation Statistics, https://www.bts.gov/content/us-ton-miles-freight Last Accessed on 02-14-2019.
2. National Railroad Passenger Corporation (AMTRAK)http://media.amtrak.com/wp-content/uploads/2015/10/Amtrak-FY16-Ridership-and-Revenue-Fact-Sheet-4_17_17-mm-edits.pdf Last Accessed on 02-14-2019
3. P. M. Besuner, D. H. Stone, M. A. DeHerrera, K. W. Schoeneberg, R-302 Statistical Analysis of Rail Defect Data (Rail Analysis – Volume 3), Association of American Railroads Track-Train Dynamics, June 1978
4. Zarembski, A.M., Palese, J.W., "Characterization of Broken Rail Risk for Freight and Passenger Railway Operations", 2005 AREMA Annual Conference, Chicago, IL, September 25-28, 2005
5. AAR Annual Spending 2016, update 7-15-16: https://www.aar.org/_layouts/15/download.aspx?SourceUrl=/Fact%20Sheets/Safety/AAR%20Annual%20Spending_2016%20Update_7.15.16.pdf
6. Zarembski, A. M., "Forecasting of Track Component Lives and its Use in Track Maintenance Planning", International Heavy Haul Association/Transportation Research Board Workshop, Vancouver, B.C., June 1991
7. Zarembski, A. M., "Development and Implementation of Integrated Maintenance Planning Systems", Transportation Research Board Annual Meeting, Washington, DC, January, 1998
8. Armstrong, Wells, Stone, & Zarembski, "Impact of Car Loads on Rail Defect Occurrences", Second International Heavy Haul Railway Conference, Colorado Springs, CO, September 1982
9. Waloddi Weibull, "A Statistical Distribution Function of Wide Applicability", ASME Journal of Applied Mechanics, Transactions of the American Society of Mechanical Engineers, September 1951
10. Fréchet, Maurice, "Sur la loi de probabilité de l'écart maximum", Annales de la Société de Mathematique, Cracovie 6, pg 93~116, 1927
11. Rosin, P., Rammler, E., "The Laws Governing the Fineness of Powdered Coal", Journal of the Institude of Fuel, Volume 7, pg 29~36, 1933
12. Stone, D.H., Comparison of Rail Behavior with 125-Ton and 100-Ton Cars, Association of American Railroads Report Number R-405, Chicago, Illinois, (1980)
13. Zarembski, A. M., Rail Life Analysis and its Use in Planning Track Maintenance, Railway Technology International, (1993)
14. Jeong, D.Y., Analytical Modelling of Rail Defects and Its Applications to Rail Defect Management, Volpe National Transportation Systems Center, Cambridge, MA, in support of the UIC World Executive Council Joint Research Project on Rail Defect Management, (2003)
15. Palese J.W., Wright T.W., "Application of a Risk Based Ultrasonic Test Frequency Scheduling System on Burlington Northern Santa Fe". AREMA Proceedings of the 2000 Annual Conference, (2000)
16. Zarembski, A. M., Palese, J. W., & Martens, J. H., The Effect of Improved Rail Manufacturing Process on Rail Fatigue Life, American Railway Engineering Association, Bulletin 733, Volume 92, (1991)
17. Fredy Castellares, Artur J. Lemonte, "A new generalized Weibull Distribution generated by gamma random variables", Journal of the Egyptian Mathematical Society (2015) 23, 382-390

18. Eisa Mahmoudi, Afsaneh Sepahdar, “Exponentiated Weibull-Poisson Distribution: Model, properties and applications”, Mathematics and Computers in Simulation 92 (2013) 76-97
19. Jalmar M.F. Carrasco, Edwin M.M. Ortega, Gauss M. Cordeiro, “A generalized modified Weibull Distribution for lifetime modeling”, Computational Statistics and Data Analysis 53 (2008) 450-462
20. Saralees Nadarajah, Gauss M. Cordeiro, “The Exponentiated Weibull Distribution: A survey”, DOI 10.1007/s00362-012-0466-x
21. Gauss M. Cordeiro, Artur J. Lemonte, “On the Marshall-Olkin Extended Weibull Distribution”, DOI 10.1007/s00362-012-0431-8
22. Alice Lemos Morais, Wagner Barreto-Souza, “A compound class of Weibull and power series distributions”, Computational Statistics and Data Analysis vol 55, (2011)
23. Giovana O. Silva, Edwin M. M. Ortega, Gauss M. Cordeiro, “The Beta Modified Weibull Distribution”, DOI 10.1007/s10985-010-9161-1
24. Jingshu Wu, Stephen McHenry, Jeffrey Quandt, “An Application of Weibull Analysis to Determine Failure Rates in Automotive Components”, National Highway Traffic Safety Administration, United States Department of Transportation, Paper No. 13-0027
25. J.Z. Yi, P.D. Lee, T.C. Lindley, T. Fukui, “Statistical modeling of microstructure and defect population effects on the fatigue performance of cast A356-T6 automotive components”, Materials Science and Engineering A 432 (2006) 59-68
26. V N A Nalkan, S Kapur, “Reliability Modelling and Analysis of Automobile Engine Oil”, Reliability Engineering Centre, IIT Kharagpur, West Bengal, India
27. Junling Wang et al 2019 J. Phys.: Conf. Ser.1213 022010
28. Ahmed Z. Al-Garni et al, “Reliability Analysis of Aeroplane Brakes”, Qual. Reliab. Engng. Int.15: 143–150 (1999)
29. R. Danzer, P. Supancic, J. Pascual, T. Lube, “Fracture Statistics of Ceramics – Weibull Statistics and deviations from Weibull Statistics”, Engineering Fracture Mechanics 74 (2007) 2919-2932
30. W. A. Curtin, “Tensile Strength of Fiber-Reinforced Composites: III. Beyond the Traditional Weibull Model for Fiber Strengths”, Journal of Composite Materials vol 34, (2000)
31. Andrew Flor, Nicholas Pinter, Jonathan W.F. Remo, “Evaluating levee failure susceptibility on the Mississippi River using logistic regression analysis”, Engineering Geology, (2010)
32. Miriam Anderejiova, Anna Grincova, Daniela Marasova, Failure Analysis of rubber composites under dynamic impact loading by logistic regression”, Engineering Failure Analysis, (2018)
33. Dalia M. Atallah, et al, Predicting kidney transplantation outcome based on hybrid feature selection and KNN classifier, Multimedia Tools and Applications, (2019)
34. Bashir Mohammed, et al, Failure prediction using machine learning in a virtualized HPC system and application”, Cluster Computing, (2019)
35. Antonio Altavilla, Laura Garbellini, Risk Assessment in the aerospace industry, Safety Science vol 40 pg. 271-298, (2002)
36. Mehmet Firat, Recep Kozan, Murat Ozsoy, O. Hamdi Mete, Numerical modeling and simulation of wheel radial fatigue tests”, Engineering Failure Analysis 16 (2009) 1533-1541
37. Dick Veldkamp, “A Probabilistic Evaluation of Wind Turbine Fatigue Design Rules”, Wind Energy vol 11, (2008)
38. John Quigley, et al, “Estimating rate of occurrence of rare events with empirical bayes: a railway application”, Reliability Engineering and Systems Safety vol 92, (2007)

39. J. Oliver, et al, “A probabilistic Risk Modelling Chain for Analysis of Regional Flood Events”, Stochastic Environmental Research and Risk Assessment, (2019)
40. Nagaraja Iyyer, et al, “Aircraft life management using crack initiation and crack growth models – P-3C Aircraft experience”, International Journal of Fatigue, vol 29, (2007)
41. Zhen Hu, et al, “Fatigue reliability analysis for structures with known loading trend”, Struct Multidisc Optim (2014)
42. S. R. Ignatovich, “Probabilistic model of multiple-site fatigue damage of riveting in airframes”, Strength of Materials, vol 46, no. 3, (May 2014)
43. J. Curley, et al, “Predicting the service-life of adhesively-bonded joints”, International Journal of Fracture, vol 103, (2000)
44. Simon P. Wilson, “Hierarchical modelling of orthopaedic hip replacement damage accumulation and reliability”, Journal of the Royal Statistical Society, Series C, Vol. 54, (2005)
45. Jianguang Fang, et al, “Multiobjective robust design optimization of fatigue life for a truck cab”, Reliability Engineering and System Safety vol 135, (2015)
46. S. Greuling, “Approaches for Fatigue assessment of welded joins in automotive industry”, DOI 10.1002/mawe.200800350
47. Zarembski, A. M., Attoh-Okine, N, Einbinder, D. “Using Multiple Adaptive Regression to Address the Impact of Track Geometry on Development of Rail Defects”, Journal of Construction and Building Materials, Volume 127 pp 546-555, (2016)
48. Zarembski, A. M., Attoh-Okine, N, Einbinder, D., Thompson, H., Sussman, T. “How Track Geometry Defects Affect the Development of Rail Defects”, American Railway Engineering Association Annual Conference, Orlando, FL, (2016)
49. Zarembski, A. M., Yurlov, D., Palese J. W., Attoh-Okine N, and Thompson, H, “Relationship between Track Geometry Degradation and Subsurface Condition as Measured by GPR”, American Railway Engineering Association Annual Conference, Chicago, IL, (2018)
50. Yurlov, D, Zarembski, A. M., Attoh-Okine, N, and Palese, J. W., “Probabilistic Approach for Development of Track Geometry Defects as a Function of Ground Penetrating Radar Measurements” Journal of Transportation Infrastructure Geotechnology, 6(1), 1-20, (2019), DOI .1007/s40515-018-0066-x
51. Zarembski, A. M., Yurlov, D, Palese, J. W. and Attoh-Okine, N., Determination of Probability of a Track Geometry Defect based on GPR Measured Subsurface Conditions Using Data Analytics, 2019 World Congress of Railway Research, (2019), Tokyo, Japan
52. Zarembski, A. M., “Big Data in Railroad Engineering”, IEEE Big Data Conference, Washington DC, (2014)
53. Efron, B., Tibshirani, R., “An Introduction to the Bootstrap”, Boca Raton, FL, Chapman & Hall/CRC, ISBN 0-412-04231-2, 2012
54. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
55. Policy and Economics Department, Association of American Railroads, “Railroad Facts 2017 Edition”, September 2017
56. Aaron S. Hess, John R. Hess, “Logistic Regression”, Transfusion, (2019)

ACKNOWLEDGEMENTS

The authors wish to thank and acknowledge the US Department of Transportation, University Transportation Center Program (RailTEAM UTC) for funding support for this research.

ABOUT THE AUTHORS

John J. Cronin, PhD

John Cronin was a graduate research assistant for his Ph.D. degree when he worked on this project.

Dr. Allan M. Zarembski, P.E., Hon. Mbr. AREMA, FASME

Dr. Zarembski is an internationally recognized authority in the fields of track and vehicle/track system analysis, railway component failure analysis, track strength, and maintenance planning. Dr. Zarembski is currently Professor of Practice and Director of the Railroad Engineering and Safety Program at the University of Delaware's Department of Civil and Environmental Engineering, where he has been since 2012. Prior to that he was President of ZETA-TECH, Associates, Inc. a railway technical consulting and applied technology company, he established in 1984. He also served as Director of R&D for Pandrol Inc., Director of R&D for Speno Rail Services Co. and Manager, Track Research for the Association of American Railroads. He has been active in the railroad industry for over 40 years.

Dr. Zarembski has PhD (1975) and M.A (1974) in Civil Engineering from Princeton University, an M.S. in Engineering Mechanics (1973) and a B.S. in Aeronautics and Astronautics from New York University (1971). He is a registered Professional Engineer in five states. Dr. Zarembski is an Honorary Member of American Railway Engineering and Maintenance of way Association (AREMA), a Fellow of American Society of Mechanical Engineers (ASME) , and a Life Member of American Society of Civil Engineers (ASCE). He served as Deputy Director of the Track Train Dynamics Program and was the recipient of the American Society of Mechanical Engineer's Rail Transportation Award in 1992 and the US Federal Railroad Administration's Special Act Award in 2001. He was awarded The Fumio Tatsuoka Best Paper Award in 2017 by the Journal of Transportation Infrastructure Geotechnology

He is the organizer and initiator of the **Big Data in Railroad Maintenance Planning Conference** held annually at the University of Delaware. He has authored or co-authored over 200 technical papers, over 120 technical articles, two book chapters and two books.