RELIABILITY AND CLUSTERING TECHNIQUES FOR INSPECTION OPTIMIZATION OF LARGE POPULATIONS

By

Brant Arthur Stratman

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Civil Engineering

August, 2007

Nashville, Tennessee

Approved:

Professor Sankaran Mahadevan

Professor Prodyot K. Basu

Professor Gautam Biswas

Professor Bruce Cooil

To my amazing mother for her boundless love, support and encouragement

and

To Kim, Grandma Jan, and Grandpa Grana for making my years at Vanderbilt possible

and

To my twin brother, Tyler, for always being by my side

ACKNOWLEDGEMENTS

This work would not have been possible without the input of many people. I thank all of my Dissertation Committee members for their advice and suggestions. I would especially like to thank Dr. Sankaran Mahadevan, my advisor and chairman of my committee. I thank him for teaching me and supporting me throughout my career at Vanderbilt, both graduate and undergraduate.

I am truly thankful for the friendship and guidance my mentor, Dr. Yongming Liu, has provided me throughout my graduate years. I also appreciate the help Liming Liu has provided during this research.

The research reported in this dissertation was supported by funds from Union Pacific Railroad and Meridian Railroad, as well as the National Science Foundation's Integrative Graduate Education and Research Traineeship (IGERT) Program. The support is gratefully acknowledged. I would also like to thank Dr. Todd Snyder and Rex Beck for their advice and guidance throughout the completion of this work and during my summer internships at Union Pacific.

I am grateful for everyone at Vanderbilt who has contributed to my education and research; especially Michael Strong and all of the members of the IGERT Reliability and Risk Engineering and Management Doctoral Program.

I would especially like to thank my family and friends. Without their support and camaraderie this could not have been completed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

There are many industrial applications in which periodic inspections need to be performed on large populations of components or devices. However, because of time and economic constraints, frequent inspection of the entire population is not economically feasible. Therefore, in order to ensure overall safety and performance, it is desirable to identify the critical samples that pose the largest risk of failing, and select them for more detailed (expensive) and frequent inspection than other, less critical samples. It is then possible to optimally schedule the detailed inspection of critical samples by balancing reliability targets and cost constraints.

One such industry in which this work is applicable is the railroad industry. In the railroad industry, higher train speeds and increased loads have led to larger wheel/rail contact forces. This evolution has changed the major wheel rim damage from wear to fatigue. Unlike the slow deterioration process of wear, fatigue causes abrupt fractures in wheels or material loss in the tread surface. These failures can cause damage to rails, damage to the train suspensions and, in a few cases, serious derailments of the train. These failures may be very expensive in terms of human and economic loss. Since there are millions of railroad wheels that need to be inspected, but limited inspection resources, it becomes necessary to identify the critical wheels on which inspections should be focused. However, it is economically infeasible to inspect the critical wheels frequently using a detailed inspection method, therefore it is also important to determine an

inspection schedule for the critical wheels in order to minimize costs while at the same time maintaining or exceeding the minimum target reliability level.

For large populations, as in the railroad wheel application, a methodology for identifying critical samples using an unsupervised clustering technique is pursued in this study. The railroad wheel data includes both numeric and nominal features; therefore a clustering methodology that is appropriate for mixed numeric and nominal data is pursued.

Completion of the methodology for optimizing the inspection schedule for large populations required two more steps: (1) calculation of reliability, and (2) optimization of the inspection schedule of the critical samples. This study develops methods for both steps, and proposes a complete methodology for optimizing inspection schedules of large populations by focusing specifically on combining clustering, reliability, and reliability-based inspection optimization techniques into an overall methodology (see Figure 1). A detailed illustrative example applying this methodology toward railroad wheel inspections is presented to demonstrate the method's effectiveness toward practical applications.

The reliability and quality of the railroad wheels in service is generally a major concern of engineers. Large variations and uncertainties in the loading, material properties, and environmental conditions cause much variation in the wheel reliability and quality. Accurate and efficient reliability calculation is a key factor in developing a reliability-based inspection planning methodology.

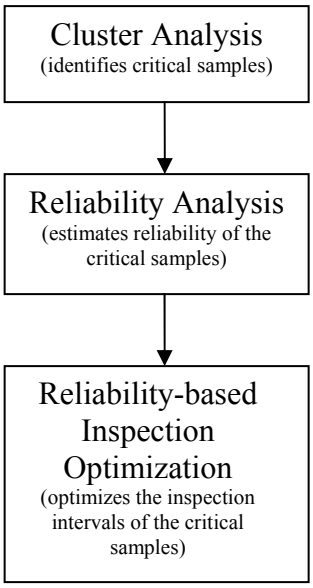


Figure 1. Methodology flow chart for optimizing inspection intervals of large populations

Research Objectives

The research in this dissertation consists of four objectives in order to ultimately develop an inspection planning methodology for railroad wheels:

1. Analyze existing manufacturing, testing, inspection, and failure data to collect appropriate statistics for reliability estimation and inspection planning.
2. Cluster the wheels into categories based on their similarities, for inspection planning.
3. Investigate the effect different wheel feature values have on the life prediction of the wheels using mechanical stress analysis and fatigue life prediction.
4. Develop an inspection planning methodology based on a reliability-based inspection optimization approach.

Scope and Organization

An examination of the manufacturing, testing, inspection, and failure data provides the statistics of significant features that influence wheel failures. Using collected railroad wheel data, the entire population of railroad wheels was divided into clusters based on similarities represented through feature values. Similarity-Based Agglomerative Clustering (SBAC) was used for this analysis. A detailed discussion of the cluster analysis is given in Chapter III.

In Chapter IV, a simulation-based method is used to estimate the reliability of railroad wheels under rolling contact fatigue loading. Numerical simulation results from a finite element analysis are used to construct the response surface of the fatigue damage index with respect to the geometry, loading, material properties and other random variables. The response surface is then used in a Monte Carlo analysis in order to ultimately develop a total lifetime estimate for the crack fatigue life within the railroad wheels. An analysis of the existing field and testing data is used to validate the lifetime estimates calculated from the Monte Carlo simulations and fatigue model.

Previously, the effects the wheel diameter, applied loading, crack length, and the crack depth below the tread surface had on the fatigue life of railroad wheels had been studied (Liu, 2006). The proposed methodology includes analyses of the rim thickness and plate design to determine their effect on the railroad wheels' fatigue life.

Traditional inspection techniques, such as drive-by inspections where all of the wheels on the train are glanced at while an inspection vehicle drives by, are not as accurate and reliable as more rigorous and quantitative identification methods. This study is concerned with adding periodic ultrasonic inspections (UT testing) to the railroad

wheel inspection practice. A reliability-based inspection optimization (RBIO) methodology that will optimize the inspection intervals of the ultrasonic inspections is proposed in Chapter V. The RBIO methodology optimizes the inspection intervals for the critical wheels that were identified in Chapter III by including the reliability of those wheels (calculated in Chapter IV) in the optimization algorithm.

The overall proposed methodology for optimizing inspection schedules of large populations is applied to the full population of railroad wheels, in Chapter VI, in order to demonstrate the method's effectiveness in an actual industry application.

CHAPTER II

WHEEL FAILURE MODES

Introduction

Although the overall research objective is to focus on the development of a methodology for optimizing the inspections of large populations, with specific application to railroad wheels, it is important to understand the different railroad wheel failure modes and how they occur since their occurrence is what is trying to be prevented. Therefore, this chapter provides a literature review focusing on the research on different wheel failure modes as well as explanations as to why and how these failures occur. After the literature review of railroad wheel failure modes in this chapter, the methodology for preventing those failures, specifically preventing shattered rim failures, is discussed in the succeeding chapters.

Fewer locomotives dispatched hauling more cars and/or hauling more tons means fewer wheels doing more work. Figure 2 illustrates this trend line (Glamorize and Gilmore, 2001). It is seen that the failure of train wheels is becoming more and more prevalent.

Three main types of failure modes are observed in the United States:

1. Shattered Rim Failure (65%)
2. Vertical Split Rim Failure (VSR) (30%)
3. Thermal Cracks (5%)

The percentage estimates of the above failures are approximate and informal by railroad industry personnel.

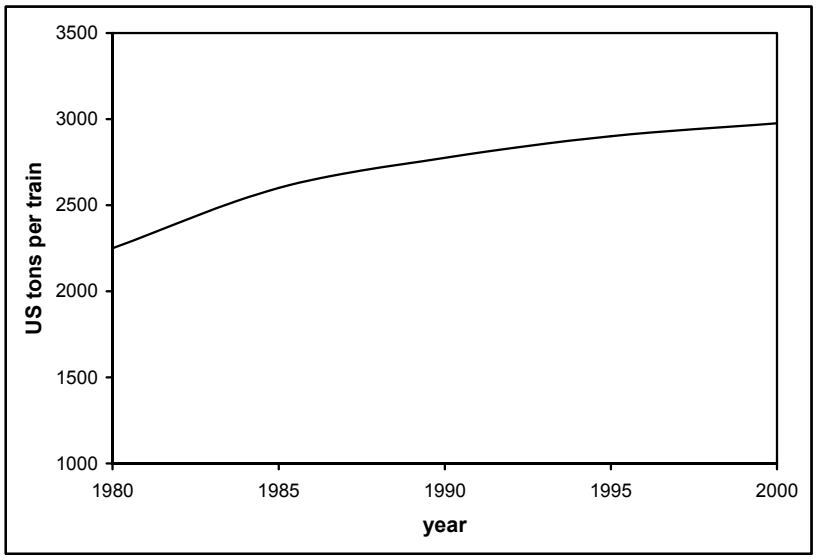


Figure 2. Average tons per train, as reported by AAR data

The following literature review focuses on the three main failure modes listed above, as well as explanations as to why and how these failures occur. Because each of the failure modes listed above start from cracks that eventually propagate, the first section of this review will focus on crack initiation.

Ekberg and Bjarnehed (1995) have summarized the research on crack initiation during rolling contact fatigue. It appears that fatigue cracks can be initiated both at the surface and below the surface (Galliera, 1995). It also appears that the mechanics behind these different phenomena are quite different.

Initiation of Surface Cracks

The initiation of surface cracks seems to be highly influenced by the presence of thermal loads due to block braking (Ekberg and Bjarnehed, 1995). Moyar and Stone (1991) use a multiaxial fatigue criterion developed by Fatemi and Socie (1988) to

quantify fatigue damage induced at the running surface. According to Moyar and Stone (1991), no fatigue damage is induced at the surface during free running of a cold wheel. When the brakes are applied and the temperature rises, the fatigue strength of the material drops. Also, the induced shear stress range and maximum normal stress are increased. This will increase the fatigue damage (Ekberg and Bjarnehed, 1995).

The hypothesis of a strong thermal influence on surface crack initiation is strengthened by the observation reported by Bartley (1988) of martensite formation near fatigue cracks. This indicates a previous history of high temperatures and fast cooling (Ekberg and Bjarnehed, 1995).

According to experimental work by Marais and Pistorius (1994), the thermal fatigue is due to the development of a tensile cyclic stress near the running surface of the wheel.

Propagation of Surface Cracks

Giménez and Sobejano (1995) analyze the propagation of surface cracks by the use of a fracture mechanics approach. According to them, the thermal cycles play a fundamental role in crack nucleation and in the growth of the crack until the threshold value of the equivalent stress intensity factor is reached. It has also been documented that thermal cycles play an important role in the generation of residual stress fields (Ekberg and Bjarnehed, 1995). Also, residual stresses have a very strong influence on the propagation of surface cracks as shown by Giménez and Sobejano (1995) and Lundén (1992).

Once the threshold value of the equivalent stress intensity factor is exceeded, the main cause of continued crack growth is the influence of mechanical loads (Giménez and Sobejano, 1995). This propagation due to mechanical loads is very fast (Hirakawa et al 1985). The fracture toughness seems to have only a slight influence on the total fatigue life (Giménez and Sobejano, 1995).

According to Ekberg et al (2002), once a crack initiates in the surface layer, it will propagate at a shallow angle to the surface, deviating first into an almost radial and later into a circumferential direction of growth. The final direction is reached at a depth of some few millimeters. In rails, the propagation of surface cracks is promoted by lubrication (grease, water, etc.) and there are indications that this is also the case for wheels. Fracture will finally occur as a branching of the crack toward the surface breaks off a piece of the wheel tread.

Initiation of Subsurface Cracks

The depth of crack initiation observed in shelling is reported to be at about 4 mm (Mutton et al, 1991) below the tread surface; Figure 3 shows a schematic of a railroad wheel and its tread surface. Lundén (1992) defines a critical region from the surface to a depth of approximately 6 mm. The initiation of subsurface cracks seems to presume very high load levels according to Ekberg et al (1995) stemming, for instance, from impact loads due to rail irregularities or joints (Ekberg and Bjarnehed, 1995).

Ekberg et al (2002) contend that subsurface cracks initiate at depths of more than 3 mm below wheel tread surface. They continue to state that at about 10 mm below the

tread surface fatigue resistance is totally governed by the presence of macroscopic inclusions.

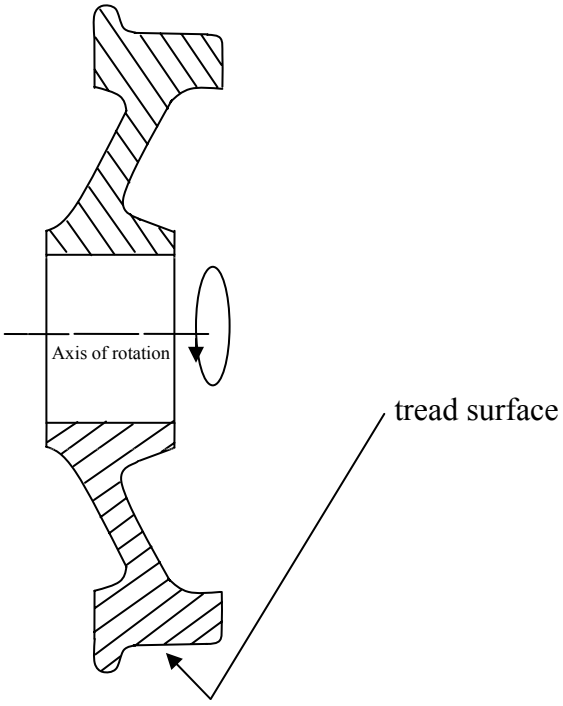


Figure 3. Cross-section of a train wheel which shows the tread surface

As in the case of surface initiated cracks, the majority of the life of subsurface cracks is spent in initiating the crack. The presence of defects or inclusions in the steel will decrease this time of initiation. According to Lundén (1992), the admissible size of a defect is strongly dependent on the crack friction coefficient (the friction coefficient between the two surfaces of the crack). If a rather high crack friction coefficient is used, a defect length of 1-2 mm would be "safe" (i.e. the influence of the defect is negligible) (Ekberg and Bjarnehed, 1995). However, if a low crack friction coefficient is used then a defect length of 1-2 mm is not necessarily considered negligible.

Ekberg et al (2002) concluded that the phenomenon of 'subsurface-initiated fatigue' is a result of high-cycle fatigue (HCF), following elastic shakedown, caused by a combination of high vertical loading, 'bad' contact geometry (giving a small contact patch) and locally low fatigue resistance of the material. Typical features of 'subsurface-initiated fatigue' are (Ekberg and Marais, 1999):

1. No signs of macroscopic inclusions or voids at the point of initiation;
2. Crack initiation at a depth of 3-10 mm below the wheel tread;
3. Crack propagation in an angle downward to a depth of some 20mm;
4. Final fracture toward the surface;
5. Circumferential crack length of 15-100 mm at fracture (in extreme cases up to 250-300 mm).

Propagation of Subsurface Cracks

According to Lundén (1992) and Mutton et al (1991), subsurface cracks propagate toward the tread surface and therefore the probability of wheel failure is small. However, Galliera (1995) shows cracks propagating in a radial direction, which can lead to catastrophic failures. According to Giménez and Sobejano (1995) cracks grow in a radial direction under the influence of thermal loads, whereas the mechanical loads make the crack grow in an axial direction. Also, a crack nucleated outside the running tread tends to grow in a direction that will position itself under the running tread (Ekberg and Bjarnehed, 1995).

Initiation and Propagation of Deep Surface Cracks

The partition between "subsurface-initiated fatigue" and "fatigue initiated at deep defects" is somewhat unclear (Ekberg et al, 2002). Typical features of fatigue initiated at deep defects are (Ekberg and Marais, 1999):

1. Fatigue crack initiation at a depth around 10-25 mm below the wheel tread;
2. Cracks initiated at defects or voids with major dimension of about 1 mm;
3. Crack propagation at an almost constant depth below the wheel tread surface (corresponding to depth of initiation) until fracture occurs;
4. Final failure as a result of branching of the circumferentially growing crack;
5. Circumferential crack length of 25-135 mm at failure.

Now that crack initial and propagation has been discussed for cracks located at different depths below the tread surface, the literature review continues by discussing how these cracks result in the three main failure modes of railroad wheels. The first failure mode discussed is shattered rim failures, which is reported to be the most dominant and detrimental failure mode. Secondly the vertical split rim failure mode is discussed, followed lastly by wheel shelling.

Shattered Rim Failures

Stone et al (2002) identify that shattered rim failures are the result of large fatigue cracks that propagate roughly parallel to the wheel tread surface. According to Stone et

al (2002), shattered rim failures form and grow 12-20 mm below the tread surface. Once a shattered rim crack is formed, it grows under normal rolling loads. Therefore, the prevention of shattered rims is best accomplished by preventing crack initiation. More importantly, the "new" wheels have a better resistance to shattered rim failures with the recent changes that have been made to ultrasonic test requirements in AAR Specification M-l07/208 ("Wheels, Carbon Steel") which reduces the acceptable size of discontinuities. This will help reduce the occurrence of some shattered rims, but will not prevent the formation of all of them (Stone et al, 2002), especially in the millions of "old" wheels that are still in use. Figure 4 schematically represents the failure mode of a shattered rim.

Consistent with Stone et al (2002), Lonsdale et al (2004) state that shattered rims are fatigue cracks that initiate and grow roughly parallel to the wheel tread surface and typically form at depths of 12 to 20 mm below the tread surface. They conclude that at this depth only stress components σ_z, τ_{oct} and τ_{max} are significant and thus logically must be responsible for shattered rims (Lonsdale et al, 2004).

Stone et al (2001) agree with the initiation depth mentioned above for shattered rims (12 to 20 mm). They also state that shattered rim fatigue cracks historically have initiated at voids and porosity in cast wheels and at aluminum oxide inclusions in forged wheels.

Giammarise and Gilmore (2001) disagree with the crack initial depth mentioned above (12 to 20 mm). However, they do agree with the fact that shattered rims are the result of large fatigue cracks which propagate parallel to the wheel tread surface. They believe that in locomotive applications, these cracks originate at depths from 22 to 32 mm below the active tread surface. From their experience, the range of mileage observed

on shattered rim failures has been on wheels with between 75,000 and 225,000 miles (120,000 to 362,000 kilometers). They also believe that when the failure mode is a shattered rim, fracture initiation on forged wheels is always associated with nonmetallic inclusions. Their experience with shattered rims has been that the failures are not random, but clustered about steelmaking deficiencies that escape vendor surveillance (Giammarise and Gilmore, 2001).

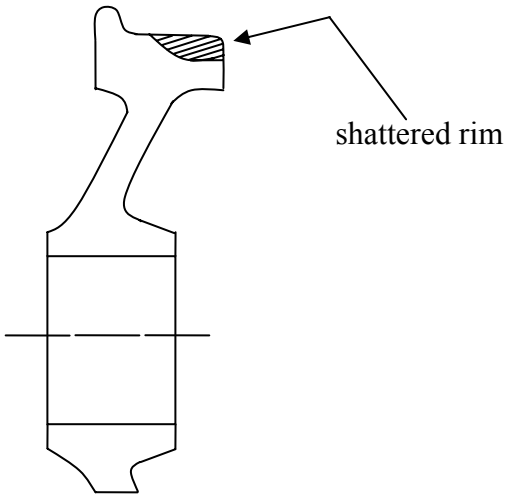


Figure 4. Typical shattered rim failure; the region with the angled lines represents the failed (missing) portion of the tread surface

Berge (2000) suggests that higher wheel loads, likely from wheel/rail impacts, are responsible for shattered rim initiation. Stone et al (2001) concur that in order to initiate cracking, a large load such as an impact may be required. They also add that once shattered rims initiate in the wheel rim, propagation occurs rapidly under normal wheel loading. Stone (2000) has also suggested that a large stress, perhaps wheel/rail impacts, is responsible for the initiation of shattered rims. He applied the Murakami criterion to the shattered rim problem in an attempt to determine fatigue life (Lonsdale et al, 2004).

With regard to τ_{oct}, present in the 12-20 mm depth range at significant magnitude, the following equation can be used to describe when yielding (permanent deformation) will occur in the wheel rim (Lonsdale et al, 2004):

$$\tau_{oct} = \frac{\sqrt{2}}{3}\sigma_{yield} \quad (1)$$

where σ_{yield} is the yield strength of the wheel rim steel in uniaxial tension. For Class C steel, this value is about 100,000 psi. Thus, yielding occurs when octahedral shear stress exceeds approximately 47,000 psi. Based upon this value, any significant wheel/rail impact load above the nominal load, or any discontinuity in the matrix (pore, inclusion, arrested crack, etc.) will act as a stress multiplier and can lead to localized yielding. Because the visual appearance of shattered rim crack fracture surfaces suggests that growth occurs due to shear stress, and the only shear stresses that are present at the relevant depth are τ_{oct} and τ_{max}, Lonsdale et al (2004) suggest these are the stresses that logically drive crack growth for a shattered rim.

Shattered rims are often noted when cracking exits the front rim face. Complete fracture of a section of the wheel rim can also occur. An examination of shattered rim fracture surfaces normally reveals obvious fatigue crack "beach marks," also referred to as a "clamshell" or "bulls-eye" pattern. If service damage is not too extensive, the crack initiation site is often clearly visible as Figure 5 shows (Stone et al, 2001).

Past technical studies have suggested that the critical defect size necessary to initiate a shattered rim is 1 mm in diameter (Marais, 1998; Lundén, 1992). Marais applied a local strain approach to the problem of shattered rim growth in cast wheels (Marais, 1998). Marais's field experience has been that shattered rims have been caused

by 1 mm voids, and a tightening of ultrasonic requirements will reduce their occurrence. Lundén (1992) calculated the size of a "safe" crack length in the wheel rim to be 1 mm in diameter. His crack growth analysis was based upon use of stress intensity factors and took into account crack surface friction and wheel/rail friction. Xing et al (1998) proposed a crack initiation process for shattered rims in forged wheels using the Murikami criterion (Murikami and Usuki, 1989) for crack initiation calculations. In their initiation model, the wheel rim crack first forms at the inclusion/matrix interface of a ball-shaped aluminum oxide inclusion (Stone et al, 2001).

Figure 5. Photograph of wheel section showing the "bulls-eye" pattern of a shattered rim; arrow points to origin

Additional research has also centered upon the size of discontinuities that are sufficient to initiate a shattered rim and to the types of stresses that are responsible for the crack propagation (Lundén, 1992; Marais, 1998; Stone and Dahlman, 2000; Xing et al, 1998).

Vertical Split Rim Failures

A vertical split rim failure is believed to be due to strength failure. The failure is a sudden failure characterized by vertical crack propagation where the tread of the wheel is detached. Figure 6 schematically represents the failure mode of a vertical split rim. In addition to occurring as a separate failure mode, vertical split rims often occur with shattered rim failures. Generally when this occurs a crack resulting in a shattered rim failure is initiated and propagates; the crack propagation can result in the branching out of the crack which can lead to both failure modes occurring simultaneously. In this case the vertical split rim would ultimately be initiated at the same deep surface initial crack as the shattered rim.

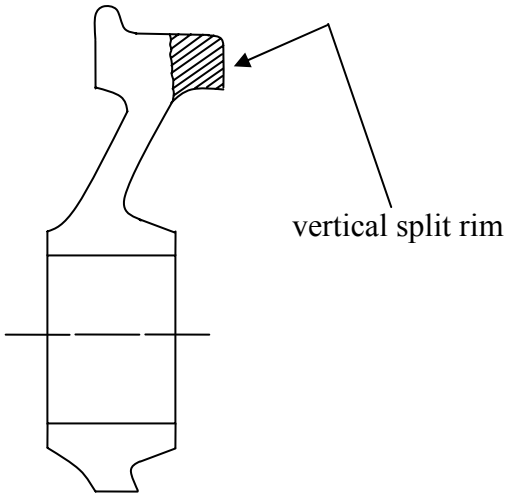


Figure 6. Typical vertical split rim failure; the region with the angled lines represents the failed (missing) portion of the tread surface

In the literature there are many references (Rail, 1994; Rail, 2000) to a vertical split head being the cause of derailment. A common definition found for a "Vertical Split Head" is a vertical split through or near the middle of the head, and extending into or

through it. A crack or rust streak may show under the head close to the web, or pieces may be split off the side of the head (Rules Respecting Track Safety, 1997). A vertical split head seems to be a very similar failure mode to the vertical split rim except that the vertical split head is a failure of the rail, not the wheel. Figures 7 and 8 schematically represent the rail and the vertical split head.

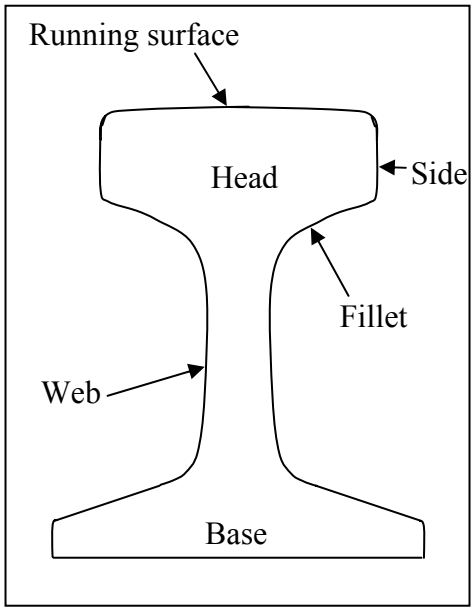


Figure 7. Schematic of the cross-section of a rail

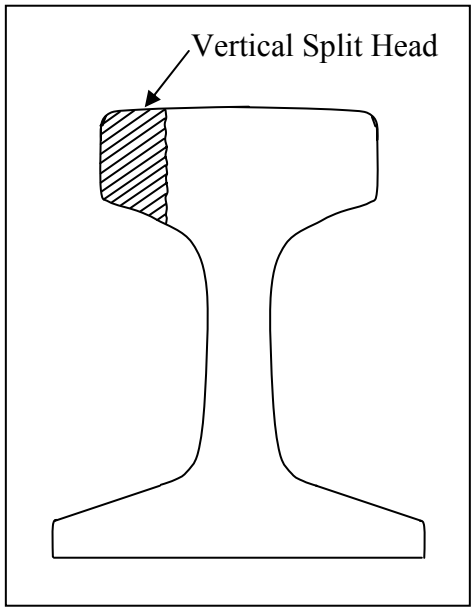


Figure 8. Typical vertical split head failure

From Figures 6 and 8, it is seen that the geometry of the contact surface of the rail and wheel are very similar. It also seems logical that both the rail and wheel are subject to the same loads since they are the two bodies exerting force on each other. It is also reasonable to assume that the wheel and rail are made from similar materials in most cases. Therefore, plausibly the failure mode for the vertical split head would be similar to the failure mode for the vertical split rim. A discussion of the failure mode for a vertical split head follows.

A Canadian derailment report (Rail, 1994) determined that the cause of their accident was a closure rail fracture initiated by a vertical split head defect. The vertical split head defect may have been initiated and propagated by wheel impacts at a loose bolted joint. The operation of the train conformed to company instructions and government safety standards. The train derailment and separation was sudden and without warning. Evidence of considerable wear damage on the underside of the rail head indicated considerable vertical and longitudinal stresses likely as a result of the joint being loose and the repetitious impacts of the passing wheels on the rail ends. These stresses and impacts may have initiated or propagated the development of the vertical split head defect (Rail, 1994).

Another Canadian derailment report (Rail, 2000) stated that failure started with a crack originating in an elongated inclusion of manganese sulfide, then propagated transversely to the rail web, then longitudinally. They concluded that this type of failure is typical of a vertical split head (VSH). Oxidation in the rail section indicated that the vertical separation had existed for several months. The VSH originated as cracks initiated by inclusions. The crack was not recent, and had progressed over a period of time; however, it was not detected because the crack was probably not long enough when the last rail test was done about seven months before the derailment. Once vertical split heads are initiated, they propagate rapidly (Rail, 2000).

It is clear from the examination of the above cases that vertical split heads propagate rapidly and without warning. It has also been stated that high vertical and longitudinal stresses may be responsible for the vertical split head failures. As stated at

the beginning of the section, by similarity, the vertical split rim is assumed to be caused by similar stresses and impacts, and have a similar failure mode as the vertical split head.

Shelling/Spalling

Ekberg and Bjarnehed (1995) refer to shelling as subsurface induced cracks, and refer to spalling as surface cracks joining to produce losses of small pieces of tread material. Shelling and spalling combine to produce the failure mode of concern. The presence of thermal cracks and spalling on the tread of a railroad wheel can be seen in Figures 9 and 10.

Principal stress components are present at and below the tread surface in the *x*, *y* and *z* directions and they can be calculated using general contact equations found in Norton (2000). The *x* direction is oriented along the length of the rail and the *y* direction is across the rail head. The *z* direction is taken as normal to the wheel tread surface, and therefore can be considered as the "depth" below the tread surface (Lonsdale et al, 2004). The equations for surface principal stresses are as follows (Norton, 2000):

$$\sigma_x = -\left(2\nu + (1-2\nu)\frac{b}{a+b}\right)P_{\max} \tag{2}$$

$$\sigma_y = -\left(2\nu + (1-2\nu)\frac{a}{a+b}\right)P_{\max} \tag{3}$$

$$\sigma_z = -P_{\max} \tag{4}$$

where σ_x is the principal stress in the *x* direction, σ_y is the principal stress in the *y* direction, σ_z is the principal stress in the *z* direction, ν is Poisson's ratio, *a* and *b* are the

semi-major and semi-minor elliptical axes of the contact patch, respectively (Lonsdale et al, 2004).

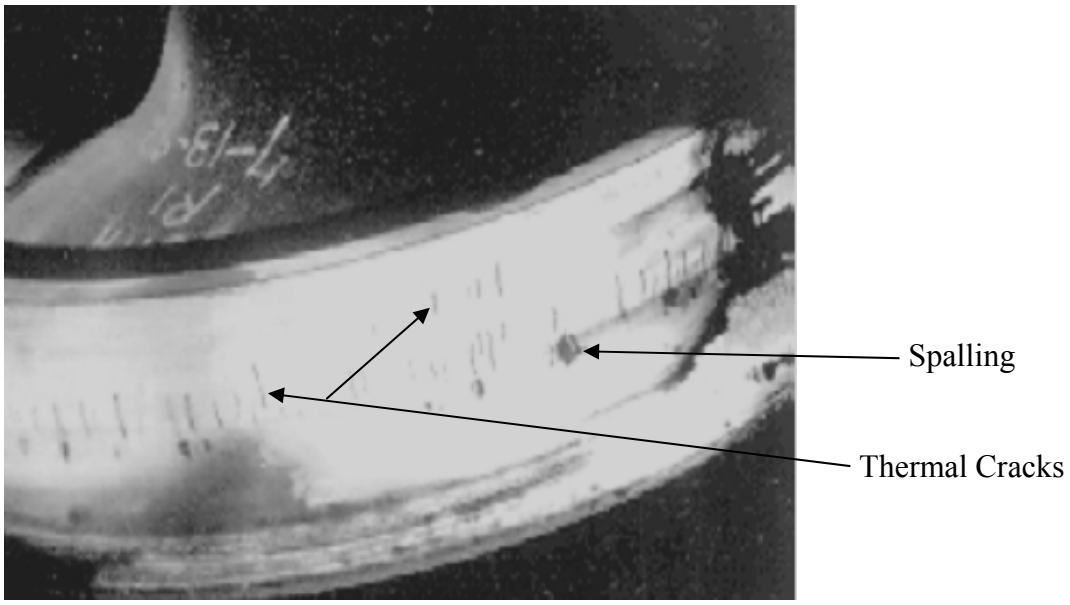


Figure 9. Thermal cracking and spalling

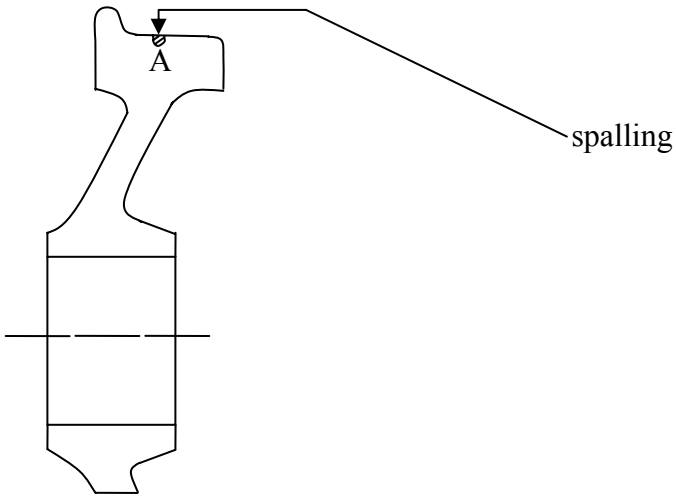


Figure 10. Typical spalling failure; region A (with the angled lines) represents the failed (missing) portion of the tread surface

The maximum shear stress τ_{max} can be calculated at the wheel/rail surface using the following equation from Norton (2000):

$$\tau_{\max} = \left| \frac{\sigma_x - \sigma_z}{2} \right| \tag{5}$$

Norton (2000) then estimates the change in the various stress components with a change in depth, z, below the wheel tread surface using Equations 6 to 9. It is important to note that the equations for a spherical contact case have only a radius a, while an elliptical contact patch which is appropriate for the general contact case has different radii a and b. For σ_x vs. depth calculations Norton (2000) uses the semi-minor axis b and for the σ_y vs. depth calculations uses the semi-major axis a. Thus, the true stress distribution below the surface may be slightly different if more exact general contact "stress vs. depth z" equations are used (Lonsdale et al, 2004).

$$\sigma_x = P_{\max}\left[-(1+2\nu) + 2(1+\nu)\left(\frac{z}{\sqrt{b^2+z^2}} \right) - \left(\frac{z}{\sqrt{b^2+z^2}} \right)^3 \right] \tag{6}$$

$$\sigma_y = P_{\max}\left[-(1+2\nu) + 2(1+\nu)\left(\frac{z}{\sqrt{a^2+z^2}} \right) - \left(\frac{z}{\sqrt{a^2+z^2}} \right)^3 \right] \tag{7}$$

$$\sigma_z = P_{\max}\left[-1 + \frac{z^3}{\left(a^2+z^2\right)^{3/2}} \right] \tag{8}$$

$$\tau_{\max} = \frac{P_{\max}}{2}\left[\frac{(1-2\nu)}{2} + (1+\nu)\left(\frac{z}{\sqrt{a^2+z^2}} \right) - \frac{3}{2}\left(\frac{z}{\sqrt{a^2+z^2}} \right)^3 \right] \tag{9}$$

It is important to emphasis that the maximum sub-surface shear stress τ_{max} occurs at a depth of about 5-6 mm. Other shear stress components include τ_{xy} (subsurface shear stress caused by *x* and *y* components) and τ_{oct} (octahedral shear stress). τ_{xy} is a very small stress below the surface and is the shear stress created by principal stresses in the *x* and *y* directions (Lonsdale et al, 2004).

The octahedral shear stress (τ_{oct}) is the stress component responsible for plastic deformation and yielding (Dieter, 1986). It is slightly less in magnitude than τ_{max}, and a sub-surface maximum occurs at the same depth. The following equation (Dieter, 1986) is used to calculate the octahedral shear stress (Lonsdale et al, 2004):

$$\tau_{oct} = \frac{1}{3}\left[(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2\right]^{1/2} \quad (10)$$

Summary

The summary of the three failure modes and their causes are:

1. Shelling occurs at initial cracks about 4 mm below the tread surface.
2. There is some disagreement about the depth of the initial cracks responsible for shattered rims, but the cracks are deep surface initiated cracks at a depth about 10-25 mm below the tread surface.
3. The maximum sub-surface shear stress τ_{max}, occurs at a depth of about 5-6 mm, which subsequently results in shelling (initial cracks for shelling are at 4-6 mm.)
4. For shattered rim failure, τ_{oct} and τ_{max} are the stresses that drive crack growth.

5. It appears that impact loading is the major factor in the initiation of shattered rim defects in wheels.
6. The failure mode that causes the wheel to fail (i.e. spalling, shattered rim, or vertical split rim) depends on where the initial crack that propagates is located. This is true because all of the failure modes start out as cracks which eventually propagate to failure; and as stated in conclusions one and two, the individual failure modes consistently initiate at cracks at a range of depths.
7. Vertical split rims result because of high vertical and longitudinal stresses, and they occur rapidly.
8. Shattered rim failure occurs slowly as a crack propagates over time. The critical defect size necessary to initiate a shattered rim is 1 mm in diameter.

This chapter discussed the three most common failure modes of railroad wheels. However, the proposed methodology in subsequent chapters to optimize the inspection schedule of railroad wheels is designed to prevent shattered rim failures as they are the most common and detrimental railroad wheel failure mode. Chapters III to V use a smaller illustrative example for railroad wheel inspection optimization to develop the overall methodology. The following chapter develops a methodology for identifying critical samples out of a large population in which inspections should be focused.

CHAPTER III

CLUSTER ANALYSIS

Overview

Since there are millions of railroad wheels which need to be inspected, but there are limited inspection resources, it becomes necessary to identify the critical wheels for which the inspections should be focused. This chapter develops a methodology for identifying such critical wheels. The general methodology is demonstrated on a small illustrative example in this chapter to demonstrate the method's effectiveness toward practical applications. The method is then applied toward a large population of railroad wheels in Chapter VI.

In the past, the railroad industry has used inspection methods that lack robustness. A commonly used method is drive-by inspections, where all of the wheels on the train are glanced at while an inspection vehicle drives by, to identify critical wheels. Crude methods such as these are not as accurate and reliable as the more rigorous and quantitative identification methods to identify critical wheels. The railroad industry has attempted to identify critical wheels by comparing individual features using common statistical techniques; however this did not produce much success in improving the identification accuracy.

A more informed way to look for differences between known critical and non-critical wheels is to employ statistically relevant methods for the extraction of implicit, previously unknown and potentially useful information in the data. More recently, such

methods have come to be known as knowledge discovery techniques. There are two types of primary discovery schemes that one may apply to derive classificatory information from data: (i) supervised classification schemes, and (ii) unsupervised classification or clustering schemes. Supervised classification schemes assume that the partitioning identifiers of the data samples being analyzed are known. The goal is to derive feature-value sets that define the classes of interest. When new data observations with unknown labels are presented, their feature values are used to assign them into one of a set of known categories. In our example, this would imply that all wheels used for deriving the classifier are labeled "critical" and "non-critical" wheels. Unsupervised classification or clustering schemes make no assumptions about the class structure; they use objective criterion functions to define similarity or dissimilarity among pairs of objects. The goal of the clustering scheme is to partition the set of samples into groups such that the samples that are more similar tend to fall into the same group and samples that are relatively distinct tend to separate into different groups.

In our application domain, we assume that the features of the critical railroad wheels are not known, and, therefore, need to be identified using an unsupervised clustering technique. This is done by initially clustering the wheels into groups; the groups that contain the highest percentage of known failed wheels are then identified as the critical wheels, and their feature values are used to identify other potentially critical wheels.

The proposed method for identifying critical wheels uses static features, i.e., features that don't change with time, to determine which wheels are most likely to fail. Another method to identify critical wheels would be to observe a large sample of wheels

over time to determine which wheels have the greatest deterioration. The wheel samples that have the most deterioration would be considered critical and would, therefore, identify the set of critical wheels for future inspections. However, the railroad industry does not collect data for individual wheels over large periods of time, so identifying critical wheels by looking at the wheels' deterioration is currently not feasible. Therefore, a methodology for identifying critical wheels using only static features has to be employed.

It is also important to note that railroad wheel data, like many industries, include both numeric and nominal features. Thus, the clustering methodology used must be designed for mixed numeric and nominal data. This chapter pursues the Similarity-Based Agglomerative Clustering (SBAC) technique (Li and Biswas, 2002) to identify critical samples since it can cluster data that contains both numeric and nominal features.

Railroad Wheel Inspection

In addition to visual inspections and ultrasonic testing, the U.S. railroad industry uses Wheel Impact Load Detectors (WILDS), to identify critical wheels[1]. A WILD system is composed of a series of strain gages welded to a rail. The strain gages measure the force applied by a wheel as it goes over the rail using a mathematical function that uses the applied load and the deflection at the foot of the rail. These impact forces are used to monitor locomotive and rail car wheel health to ensure safe train operations. More specifically, if an individual wheel surpasses the impact load limit, as detected by WILDS, that wheel can be (it's not required) taken out of service. Currently the

[1] Author worked in the industry and knows this from experience.

Association of American Railroads' (AAR) wheel impact load limit for wayside detectors is 400.3 kN (90,000 pounds) (AAR, 2005); meaning any wheel with an impact higher than 400.3 kN (90,000 pounds) can be taken out of service. However, even though the impact limit may be surpassed by an individual wheel, the railroads still have the choice to leave the wheel in service or to remove it. The 400.3 kN (90,000 pound) impact load limit is set so that railroads don't remove their competitor's wheels, and charge their competitor, until a wheel surpasses the 400.3 kN (90,000 pound) impact load limit. Therefore, the 400.3 kN (90,000 pound) impact load limit is used to identify critical wheels but it's not required that those critical wheels be taken out of service. This means, because of time constraints involved in removing a wheel, among other reasons, many wheels are left in-service even after they surpass the 400.3 kN (90,000 pound) impact load limit. It is also important to note that many of the catastrophic wheel failure modes, namely shattered rim failures and vertical split rim failures, occur below the 400.3 kN (90,000 pound) condemning limit making the WILD system ineffective when looking at individual impacts.

Stratman et al (2007) proposed two additional criteria for the removal of wheels with high likelihood of failure. The study observed two real-time structural health monitoring trends using the WILD system in data collected from in-service trains. These impact trends can identify, in real-time, the wheels that have a high probability of failure from high impact wheels.

Railroad wheels fail in different ways that can be attributed to different failure mechanisms, as discussed in Chapter II. Subsurface defect-initiated failure, such as shattered rim failures, is the critical wheel failure mode that needs to be prevented.

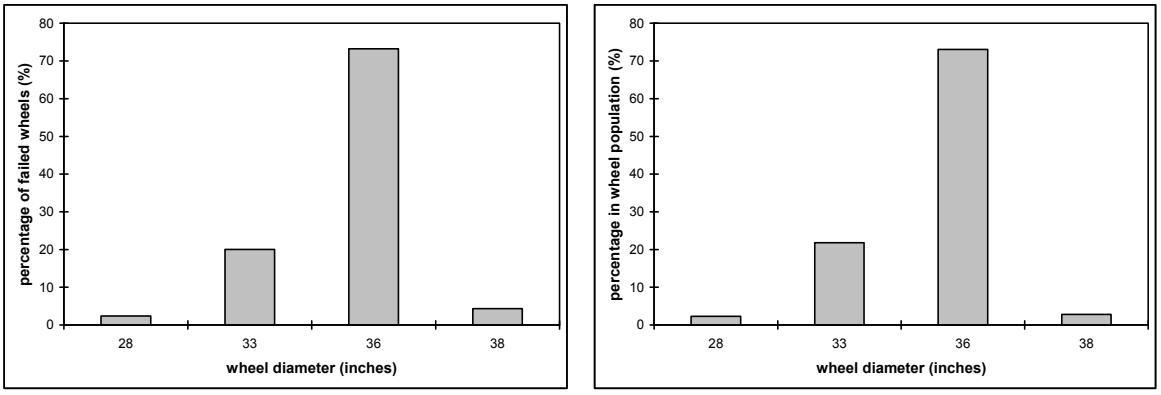
Therefore, wheels with the highest probability of resulting in a shattered rim failure need to be inspected more regularly than the general population of wheels.

Results of a preliminary analysis to identify critical wheels in the entire wheel population using individual features are shown in Figures 11 to 21. When comparing the individual feature values of failed wheels to the same feature values for the entire wheel population, there does not seem to be individual feature value differences that separate the failed wheels from the rest of the wheel population.

For further discussion, it is seen that when comparing the percentage of wheels in the failed population (Figures (a)) to the percentage of wheels in the total population (Figures (b)) that their distributions are very similar. As an example, approximately 70 percent of wheels in the total population have a wheel diameter of 36 inches (as seen in Figure 11(b)). Also, approximately 70 percent of wheels in the failed population have a wheel diameter of 36 inches (Figure 11(a)). The percentage similarity holds true for each wheel diameter which means that there isn't a specific diameter that occurs in failures more often than it occurs in the total population. If, on the other, had only 3 percent of wheels in the total population had a wheel diameter of 28 inches while 20 percent of wheels in the failed population had a wheel diameter of 28 inches, then that would indicate that wheels with a 28 inch wheel diameter fail at a greater rate compared to their total population than other wheel diameters; indicating a critical feature. However, as seen in Figures 11 to 21, there are no individual features or feature values with significant differences between total and failed populations that could indicate a critical wheel.

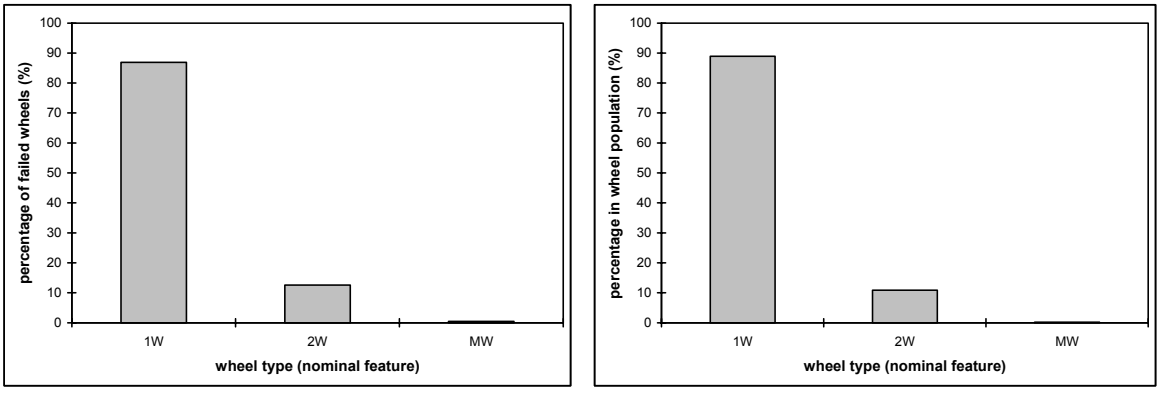
Common statistical techniques based on numerical methods have not been very successful in identifying the critical wheels with a high probability of failure either. A

primary reason for this is that a number of important features are symbolic-valued and not numeric. A systematic clustering approach for mixed numeric and symbolic features is used to overcome this problem, to create groups of wheels that have similar characteristics, and then study the groups with dominant failed wheels to derive the feature values that define critical wheels.



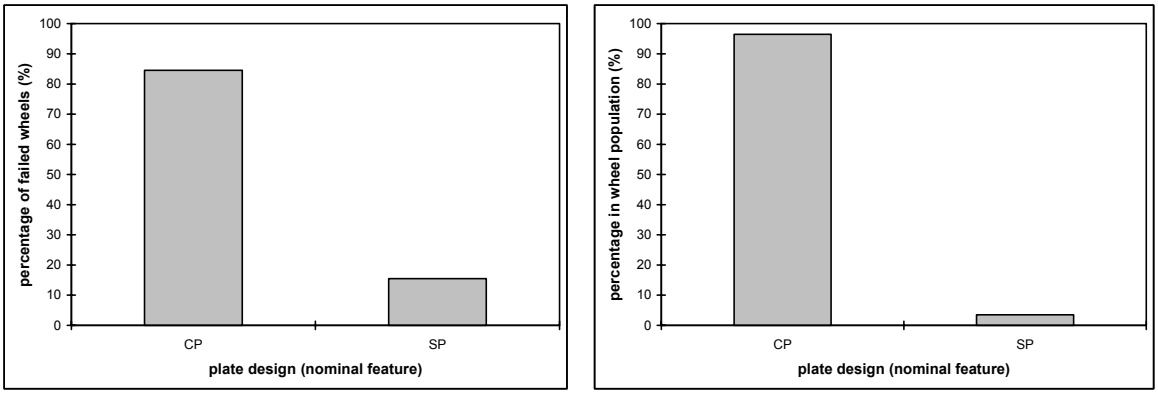
a) Failed Wheels b) Entire Wheel Population

Figure 11. Wheel diameter feature comparison between entire wheel population and failed wheels



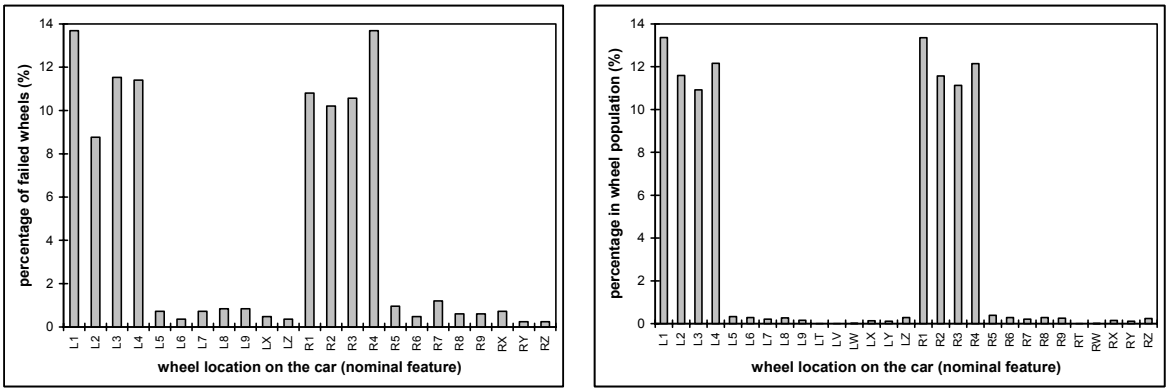
a) Failed Wheels b) Entire Wheel Population

Figure 12. Wheel type feature comparison between entire wheel population and failed wheels



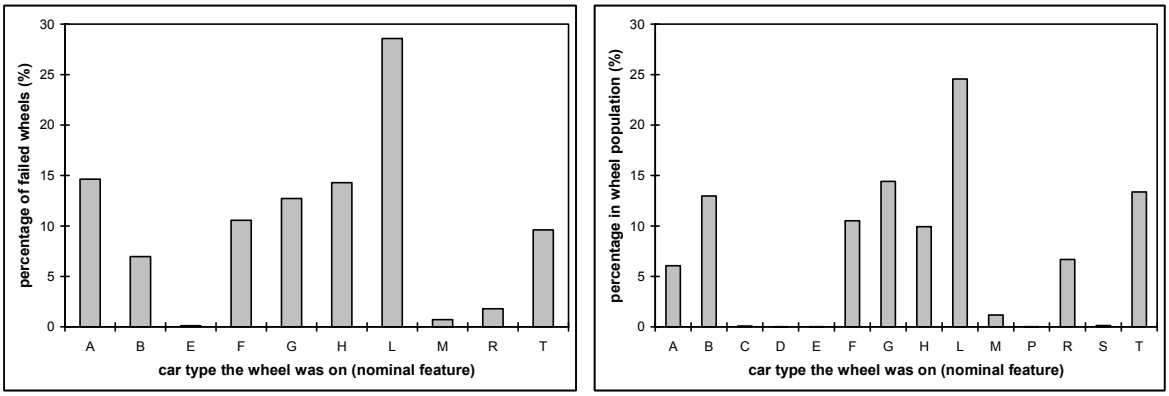
a) Failed Wheels b) Entire Wheel Population

Figure 13. Plate design feature comparison between entire wheel population and failed wheels



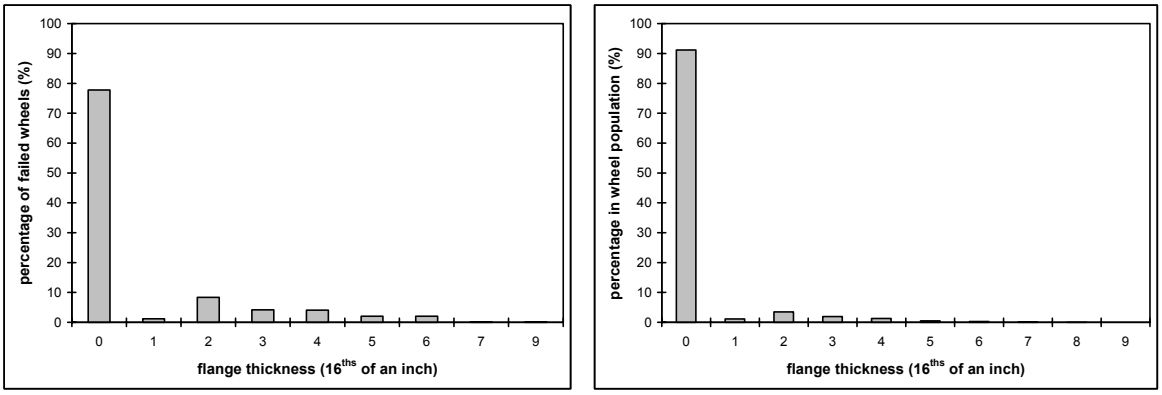
a) Failed Wheels b) Entire Wheel Population

Figure 14. Wheel location feature comparison between entire wheel population and failed wheels



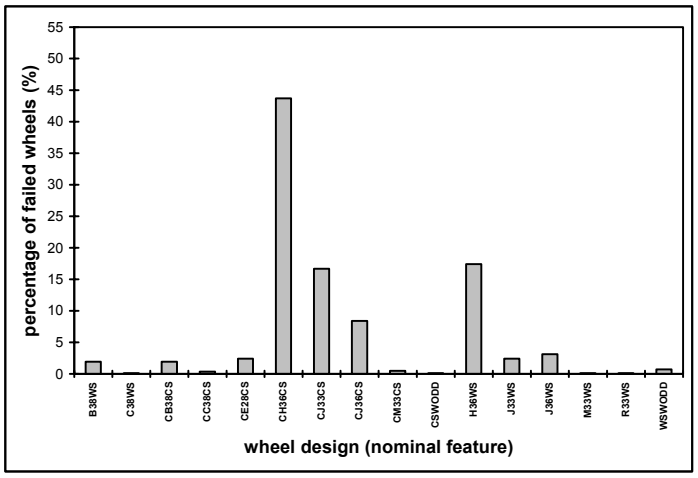
a) Failed Wheels b) Entire Wheel Population

Figure 15. Car type feature comparison between entire wheel population and failed wheels

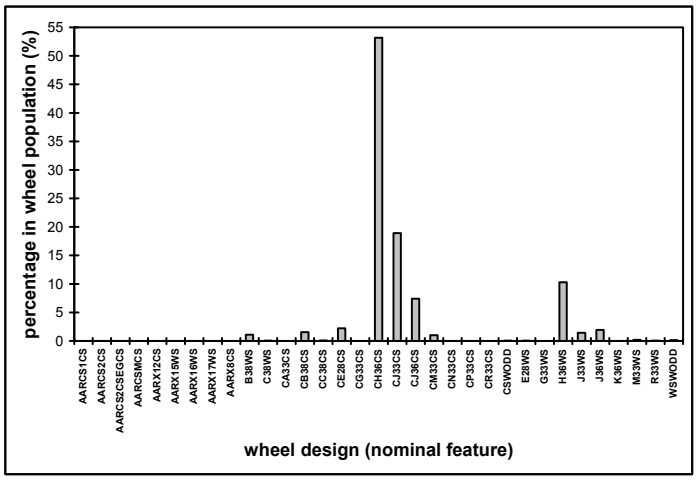


a) Failed Wheels b) Entire Wheel Population

Figure 16. Flange thickness feature comparison between entire wheel population and failed wheels

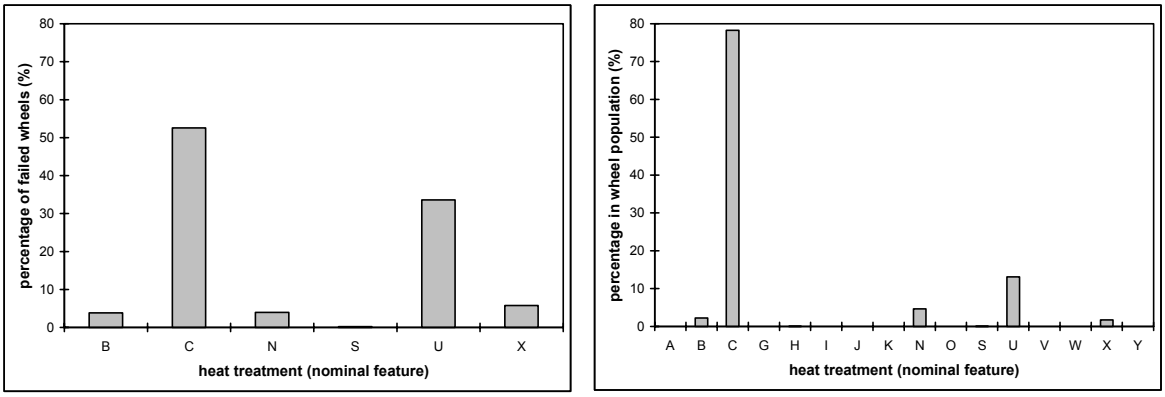


a) Failed Wheels



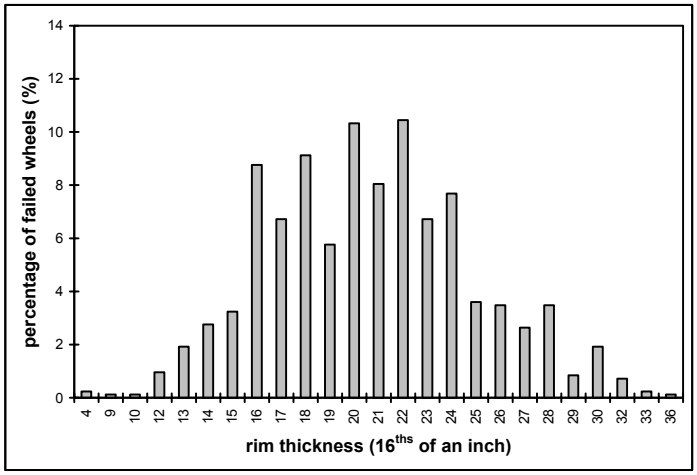
b) Entire Wheel Population

Figure 17. Wheel design feature comparison between entire wheel population and failed wheels

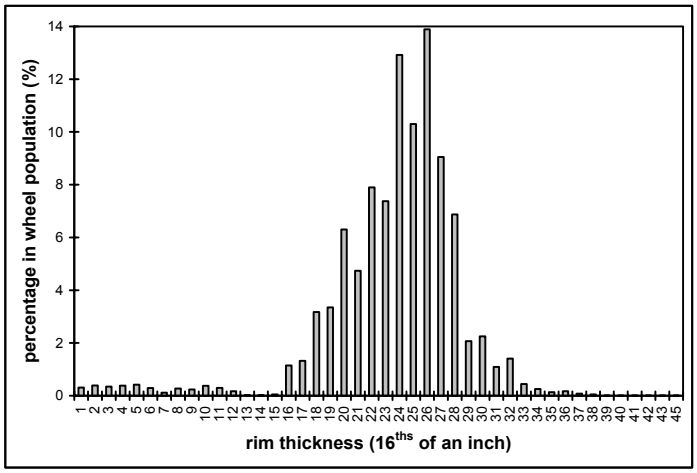


a) Failed Wheels b) Entire Wheel Population

Figure 18. Heat treatment feature comparison between entire wheel population and failed wheels

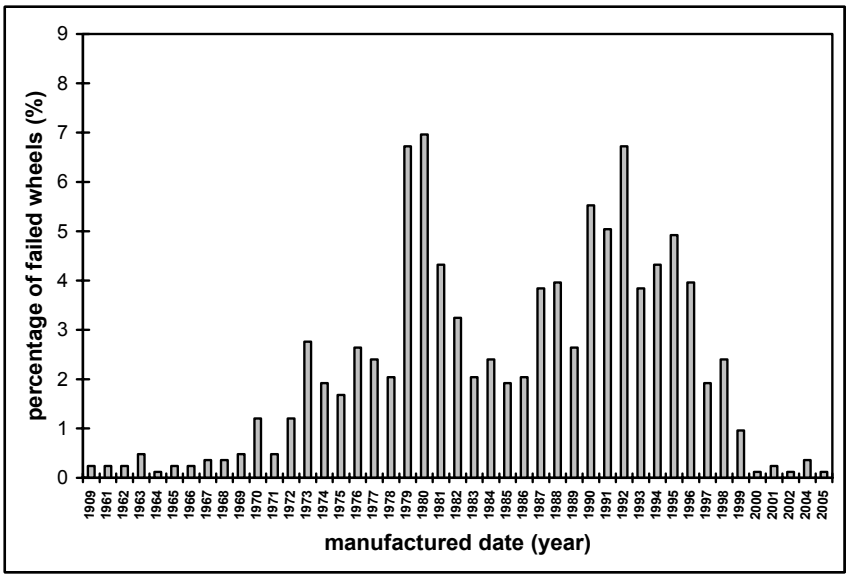


a) Failed Wheels

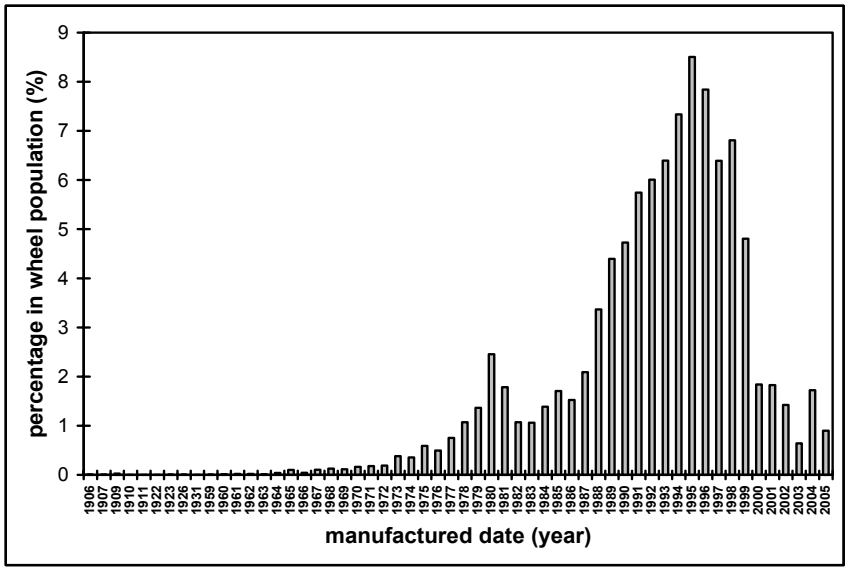


b) Entire Wheel Population

Figure 19. Rim thickness feature comparison between entire wheel population and failed wheels



a) Failed Wheels

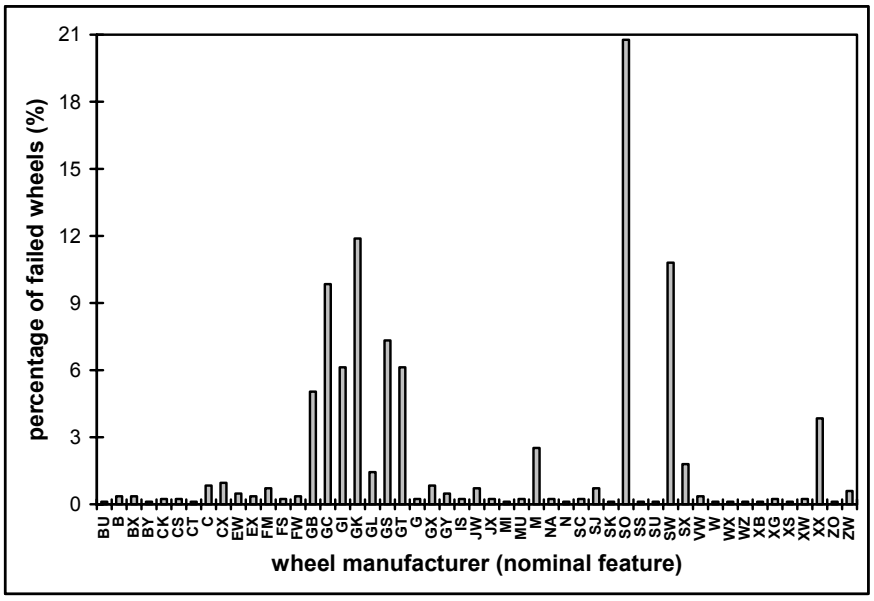


b) Entire Wheel Population

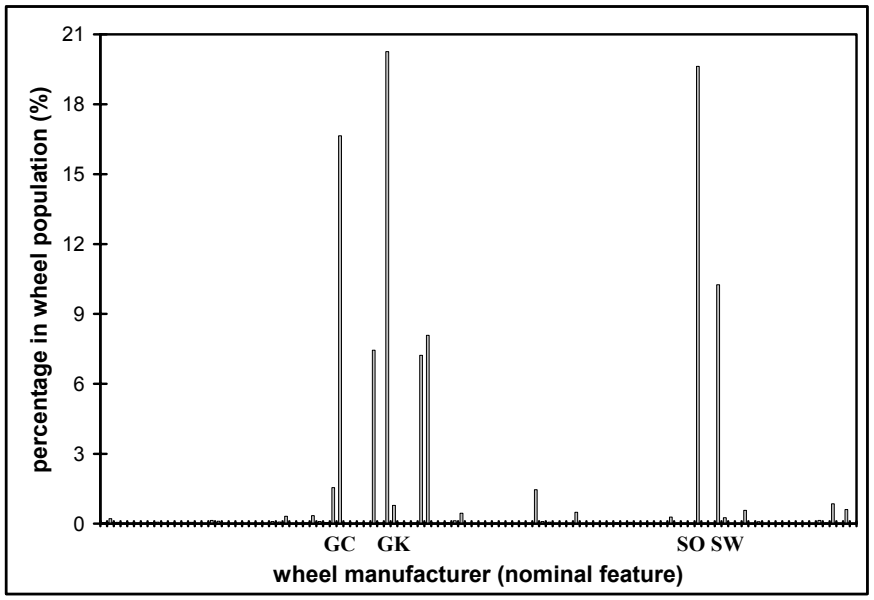
Figure 20. Manufacturing date feature comparison between wheel population and failed wheels

In Figure 19 it is seen that there are differences between the failed population and total population when comparing individual rim thicknesses. This is expected because as a wheel wears due to usage it becomes more likely to fail. Therefore, although the differences between the total and failed populations are significant, they are expected and

don't indicate critical feature values. The same logic holds true for Figure 20; as wheels age they are expected to fail at greater rates and therefore the differences between failed and safe populations don't indicate critical feature values.



a) Failed Wheels



b) Entire Wheel Population

Figure 21. Wheel manufacturer feature comparison between wheel population and failed wheels

A clustering approach was used as opposed to classification, because the purpose was to group, or "cluster", the most similar samples in the feature space without using previous knowledge of how to group the samples. Classification techniques assume class labels, such as safe and failed wheels, and then employ information-theoretic measures to the feature sets to create the "best" separation between the pre-defined classes. For complex domains, this may produce decision schemes that are non-intuitive and hard to interpret. On the other hand, clustering schemes group by context that is defined by the multi-dimensional space of feature values, and within different contexts attempt to separate the critical wheels from those that are less likely to fail. In this case, the identification of critical wheels in the multi-dimensional space is further governed by contexts that may be attributed to different classes of wheels.

Cluster Analysis

Clustering is a useful exploratory tool for multi-dimensional data analysis. When the underlying structure of the data is not readily apparent (i.e., the class structure or the number of groups in the data are unknown), cluster analysis may be applied to uncover this knowledge. Clustering schemes require data objects to be defined in terms of a predefined set of features. Features represent properties of the object that are relevant to the problem solving task. For example, if we wish to classify automobiles by speed and power, then body weight, body shape, and engine size are relevant features, but the color of the car body is not. Selection of appropriate features is an important pre-processing step for cluster analysis.

There are three components that primarily govern the clustering process: (i)

distance or similarity metric, (ii) the control algorithm, and (iii) the criterion function. The distance metric typically uses an objective measure, e.g., the Euclidean distance or the Mahalanobis distance, to define the proximity between pairs of objects in the data set. The control algorithm can be *agglomerative*, where the partition structure is constructed bottom-up through successive merging of atomic clusters into larger groups, or *divisive*, where clusters are formed top-down from one large cluster that includes all of the data objects, and then is successively subdivided into smaller groups. The criterion function is used for evaluating the goodness of a partition structure once it is formed; typically the mean square error is used in numeric partitioning schemes.

Regardless of the distance metric, control algorithm or criterion function, the most defining characteristic of a clustering scheme is the type of data the scheme is intended to cluster. For static data clustering, e.g., each data is represented as a vector of feature values with one value per feature[2], clustering schemes have been developed for numeric, nominal and mixed data types.

Traditional clustering algorithms focus on numerical data whose inherent geometric properties can be exploited naturally to define distance functions between data points. These algorithms include K-means (MacQueen, 1967), DBSCAN (Ester et al, 1996), BIRTH (Zhang et al, 1996), C^2P (Nanopoulos et al, 2001), CURE (Guha et al, 1998), CHAMELEON (Karypis et al, 1999), and WaveCluster (Sheikholeslami et al, 1998). The illustrative railroad wheel example data set includes numeric features, such as the manufactured date, rim thickness, flange thickness, and wheel diameter. Distance metrics like the Euclidean distance can be applied to such feature sets, and a number of

[2] Clustering schemes developed for time series data and multimedia data, for example video, audio, or image data, are not discussed in this study.

numeric data clustering algorithms can be applied to form the groups or clusters. However, the numeric distance metrics cannot be used to compute the proximity for data objects with nominal features, e.g., wheel manufacturer, or the type of heat treatment the wheel was subjected to. Therefore, the traditional numeric clustering algorithms are not suitable for clustering data with mixed numeric and nominal features.

On the other hand, for data sets that contain only nominal-valued features, frequency counts of feature value matches or conditional probability estimates form the basis for defining the similarities/dissimilarities between the data objects. The illustrative railroad wheel data set includes a number of nominal-valued features: wheel manufacturer, heat treatment, wheel design, wheel type, plate design, car type the wheel was on, and the wheel location on the car. Algorithms have been proposed in the past for clustering nominal data (e.g., He et al, 2002; Guha et al, 1999; Gibson et al, 1998; Zhang et al, 2000; Han et al, 1997; Ganti et al, 1999; Huang, 1997; Wang et al, 1999), however all of these algorithms work only with nominal-valued features, and they cannot be easily extended to clustering mixed types.

The railroad wheel data includes both numeric and nominal-valued features, as seen in Figures 11 to 21. Thus, a clustering methodology that is appropriate for mixed numeric and nominal data is pursued here.

Clustering Mixed Numeric and Nominal Data

Some research efforts have been reported (e.g., Huang, 1997; Wang et al, 1999; Huang, 1998; Chiu et al, 2001; Fisher, 1987; Gluck and Corter, 1985; Hanson and Bauer, 1989; McKusick and Thompson, 1990; Reich and Fenves, 1991; Reich, 1991; Cheesman

and Stutz, 1995; Li and Biswas, 2002) for the problem of clustering mixed type features. Huang (1998) presents two algorithms, *k*-modes and *k*-prototypes, which extend the *k*-means paradigm to nominal domains and domains with mixed attributes. A new distance measure for nominal features based on the total mismatches of the nominal features of two data records is proposed in the *k*-modes algorithm. For mixed attributes, the *k*-prototypes algorithm used a weighted sum of Euclidean distance for numeric attributes and the *k*-modes distance measure for nominal attributes. However, the weights have to be determined *a priori*. Improper weight assignments result in biased treatment of different attribute types, and the clusters formed are very sensitive to the weights chosen.

The clustering algorithm presented by Chiu et al (2001) is available commercially in the Clementine 6.0 data mining tool. The distance measure is derived from a probabilistic model, where the distance between two clusters is equivalent to the decrease in the log-likelihood function as a result of merging the two clusters. This algorithm is based on the framework and distance measure proposed in the BIRTH system (Zhang et al, 1996). The BIRTH algorithm has the drawback that it may not work well when clusters are not "spherical" and it is sensitive to the order in which the data objects are considered by the clustering algorithm. The same problems carry over to Chiu et al (2001) clustering algorithm.

Methods, such as COBWEB (Fisher, 1987) use the *Category Utility* (*CU*) measure (Gluck et al, 1985) to partition a data set in a manner that maximizes the probability of correctly predicting a feature value given group C_k. The category utility measure represents a tradeoff between intra-class similarity and inter-class dissimilarity of samples. Methods like WITT (Hanson and Bauer, 1989) use *correlation* measures to

define the similarity between pairs of objects. These measures are tailored for nominal attributes, though variations, such as COBWEB/3 (McKusick and Thompson, 1990) and ECOBWEB (Reich and Fenves, 1991) use modifications of the *CU* measure to handle numeric attributes. In COBWEB/3 the numeric *CU* measure does not take into account an important piece of information, the actual distances between object values, in determining class structure; also, the accuracy of the *CU* measure for numeric features is difficult to bound. In ECOBWEB the clustering results are very much dependent on two user defined parameters: (i) the *method* for calculating $2I_i$, and (ii) n, the "expected number of distinct intervals of property attribute A_i." This may not be an issue when a hierarchical structure is used for prediction purposes (Reich, 1991), but it is not a desirable characteristic in knowledge discovery tasks.

AUTOCLASS (Cheesman and Stutz, 1995) uses a *finite mixture model* and derives groupings of objects that locally maximize the posterior probability of individual clusters given the feature distribution functions. Unfortunately, AUTOCLASS suffers from the over-fitting problem associated with the maximum likelihood optimization methods for probabilistic models. Also, the computational complexity required by the nested three-level search process is extremely high making AUTOCLASS inconvenient for handling large datasets.

The Similarity-Based Agglomerative Clustering (SBAC) algorithm (Li and Biswas, 2002) works well for data with mixed numeric and nominal features. SBAC is based on a similarity measure proposed by Goodall (1966) for biological taxonomy that gives higher weight to less common feature values. It makes no assumptions about the underlying distributions of the feature values. An agglomerative algorithm is employed

to construct a classification tree and a simple distinctness heuristic is used to extract a partition of the data. The performance of SBAC has been studied on real and artificially generated data sets (Li and Biswas, 2002). Results demonstrate the effectiveness of this algorithm in unsupervised discovery tasks. Comparisons with other clustering schemes illustrate the superior performance of this approach.

Similarity-Based Agglomerative Clustering (SBAC)

A brief overview of the Similarity-Based Agglomerative Clustering (SBAC) algorithm (Li and Biswas, 2002) is presented below. A similarity measure inspired by biological taxonomy proposed by Goodall (1966) provides the basis for defining the proximity of objects described by both numeric and symbolic features. The control algorithm uses an agglomerative approach (Li and Biswas, 2002) to build the partition structure.

The Similarity Measure

The Goodall similarity measure (Goodall, 1966) inspired by biological taxonomy, provides a unified framework for handling both nominal and numeric features. The SBAC algorithm (Li and Biswas, 2002) adopts this similarity measure into a general framework for clustering in many different domains. A pair of objects (i, j) is considered more similar than a second pair of objects (l, m), if and only if the objects i and j exhibit a greater match in feature values that are less common in the population. In other words, similarity among objects is decided by the uncommonality of their *feature value* matches. The similarity measure computed using this heuristic helps to define more cohesive, tight

clusters where objects grouped into the same cluster are likely to share special and characteristic feature values.

In order to calculate the similarity value for a pair of categorical features, the *More Similar Feature Value Set*, $MSFVS((V_i)_k)$ is defined first. This is the set of all pairs of values for feature k that are equally or more similar to the pair $((V_i)_k, (V_i)_k)$. A value pair is more similar if it has a lower frequency of occurrence. The square of the probability of picking a pair $((V_l)_k, (V_l)_k) \in MSFVS((V_i)_k)$ at random is:

$$\left(p_l\right)_k^2 = \frac{\left(f_l\right)_k \cdot \left(\left(f_l\right)_k - 1\right)}{n \cdot \left(n-1\right)}, \tag{11}$$

where $(f_l)_k$ is the frequency of occurrence of value $(V_l)_k$ in the population and n is the total number of objects in the population. Summation of the squares of the probabilities of all such pairs, by definition, gives the dissimilarity score of the pair, $(D_{ii})_k$. Thus, the similarity of the pair $((V_i)_k, (V_i)_k)$ is computed as:

$$\left(S_{ii}\right)_k = 1 - \left(D_{ii}\right)_k = 1 - \sum_{l \in MSFVS\left((V_i)_k\right)} \left(p_l\right)_k^2. \tag{12}$$

The similarity score for a numeric feature value pair is calculated in a similar fashion as for a nominal-valued pair. For numeric feature values, the similarity measure takes both the *magnitude* of the feature value difference and the *uniqueness* of the feature value pair into account. The *magnitude* of the feature value difference is considered first. The smaller the *magnitude* of the difference between the values $((V_i)_k, (V_j)_k)$, the less likely it is that a pair of values picked at random will fall in the segment defined by the endpoints $(V_i)_k$ and $(V_j)_k$, therefore, the more similar this pair of objects.

When the *magnitudes* of difference for two pairs of values are equal, the similarity value is influenced by the *uniqueness* of the segment defined by the values.

The *uniqueness* of a segment is computed by summing up the frequency of occurrence of all other values in the population that are between the pair of values that define the segment. For two pairs of values of equal length but different endpoints, the segment which includes the lower cumulative frequency of occurrence of other values within its endpoints is defined to be more *unique* and, therefore, the more similar the pair of values that define the segment.

For numeric feature k, to calculate the similarity score for a pair of values $((V_i)_k, (V_j)_k)$, first the *More Similarity Feature Segment Set, MSFSS*$((V_i)_k, (V_j)_k)$ is determined. This is the set of all pairs of values $((V_l)_k, (V_m)_k)$ for feature k that are equally or more similar to the pair $((V_i)_k, (V_j)_k)$; comparing similarities of value pairs $((V_l)_k, (V_m)_k)$ to the pair $((V_i)_k, (V_j)_k)$ follows the procedures discussed above. The probability of picking two objects from the population having values $(V_l)_k$ and $(V_m)_k$ for feature k, where $((V_l)_k, (V_m)_k) \in MSFSS((V_i)_k, (V_j)_k)$, is:

$$\begin{cases} 2(p_l)_k(p_m)_k = \dfrac{2(f_l)_k(f_m)_k}{n(n-1)}, & (p_l)_k \neq (p_m)_k, \\ (p_l)_k(p_m)_k = \dfrac{(f_l)_k((f_l)_k - 1)}{n(n-1)}, & (p_l)_k = (p_m)_k, \end{cases} \tag{13}$$

where $(f_l)_k$ and $(f_m)_k$ are the frequencies of occurrence of values $(V_l)_k$ and $(V_m)_k$, respectively, and n is the total number of objects in the population. Summing up the probabilities of all value pairs in $MSFSS((V_i)_k, (V_j)_k)$ gives the dissimilarity contribution of feature k, $(D_{ij})_k$. Thus the similarity of the pair $((V_i)_k, (V_j)_k)$ is computed as:

$$(S_{ij})_k = 1 - (D_{ij})_k = 1 - \sum_{l,m \in MSFSS((V_i)_k, (V_j)_k)} 2(p_l)_k(p_m)_k. \tag{14}$$

Up to this point, the similarity score computations for individual features have

been discussed. Now the similarity score computations for individual features are extended to a collective similarity score for a pair of objects described by multiple, mixed type features.

First however, the combined similarity score for the numeric features is calculated using Fisher's (Fisher, 1963) χ^2 transformation:

$$(\chi_c)_{ij}^2 = -2\sum_{k=1}^{t_c} ln\left((D_{ij})_k\right), \tag{15}$$

where t_c is the number of numeric features in the data. χ_c also follows a χ^2 distribution with t_c degrees of freedom. The similarity test scores for the nominal features are then combined using Lancaster's *mean value* (Lancaster, 1949) χ^2 transformation:

$$(\chi_d)_{ij}^2 = 2\sum_{k=1}^{t_d}\left(1 - \frac{(D_{ij})_k ln(D_{ij})_k - (D_{ij})_k' ln(D_{ij})_k'}{(D_{ij})_k - (D_{ij})_k'}\right), \tag{16}$$

where t_d is the number of nominal attributes in the data, $(D_{ij})_k$ is the dissimilarity score for nominal attribute value pair $((V_i)_k, (V_j)_k)$, and $(D_{ij})'_k$ is the next smaller dissimilarity score in the nominal set. χ_d^2 is χ^2 distributed with t_d degrees of freedom.

The addition of two χ^2 distributions is still χ^2 with the number of degrees of freedom equal to the sum of the two degrees of freedom values. Therefore, the sum of the two types of features is χ^2 distributed with $(t_c + t_d)$ degrees of freedom. Now, the collective dissimilarity score for the pair of objects with mixed feature types is calculated from the expression:

$$D_{ij} = e^{-\frac{\chi_{ij}^2}{2}(t_d + t_c - 1)} \sum_{k=0} \frac{\left(\frac{1}{2}\chi_{ij}^2\right)^k}{k!}, \tag{17}$$

where $\chi_{ij}^2 = (\chi_c)_{ij}^2 + (\chi_d)_{ij}^2$. The overall similarity score representing the set of $(t_c + t_d)$ independent similarity measures is $S_{ij} = 1 - D_{ij}$.

The Control Structure

SBAC's control algorithm is agglomerative; meaning the partition structure is constructed bottom-up through successive merging of atomic clusters into larger groups. Specifically, SBAC's agglomerative hierarchical clustering algorithm uses the Unweighted Pair Group Method with Arithmetic Average (UPGMA) method (Jain and Dubes, 1988). The algorithm starts with a pairwise dissimilarity matrix D of the set of objects to be clustered. Dissimilarity between a pair of objects is the complement of their similarity score, $D_{ij} = 1 - S_{ij}$. At any step, the clusters that have the minimum pairwise dissimilarity value are merged into a single cluster. Dissimilarity between the new cluster and the other clusters is defined as the average dissimilarity between an old cluster and the component clusters of the new cluster. The end result is a classification tree whose leaf nodes are individual objects and whose root defines a cluster or group that includes all objects. The steps of the algorithm are described in (Li and Biswas, 2002) in detail.

In the classification tree, each node has an associated dissimilarity score, which indicates the dissimilarity level at which its child nodes merge together. The decrease in dissimilarity score from a parent node to a child node indicates that the objects in the child node form a more cohesive group than the ones in the parent node. From the root of the tree to the leaves, the dissimilarity score along each path decreases monotonically. The defining clusters in the classification tree are the clusters that are cohesive yet not

fragmented. To achieve the proper balance between cohesiveness and fragmentation, the tree is evaluated top-down in a depth first fashion, comparing the dissimilarity values between parent and child nodes. The classification tree is cut off at points where the difference between the two dissimilarity values is greater than a certain percentage threshold, *t*. The set of nodes on the frontier of the traversal along the different paths define the partition structure (the clusters of interest).

Railroad Wheel Inspection Planning

This section illustrates the SBAC approach to railroad wheel analysis. The goal is to identify the set of critical wheels from the entire population in which inspections should be focused. All of the samples contain a failed/safe feature value which is used to identify the critical clusters once the clusters are formed by identifying the clusters with the highest percentages of failed wheels. The wheels were described by 12 features; four of the features were numeric-valued, and eight were nominal-valued (see Table 1). This illustrative analysis included 200 wheels that have not yet failed (denoted as safe) and 50 wheels that have failed (denoted as failed), (i.e., total population = 250).

Table 1. Wheel clustering features

Clustering Feature	Feature Type	Number of Possible Feature Values
wheel manufacturer	Nominal	28
manufactured date	Numeric	462 (month intervals)
heat treatment	Nominal	5
rim thickness	Numeric	17 (1/16" intervals)
flange thickness	Numeric	9 (1/16" intervals)
wheel design	Nominal	13
wheel type	Nominal	3
plate design	Nominal	2
wheel diameter	Numeric	4
car type the wheel was on	Nominal	9
wheel location on the car	Nominal	8
safe/failed indicator	Nominal	2

Description of the Dataset

Some of the features are stamped into the wheels when they are manufactured which means that these values do not change with time; these features include the wheel manufacturer, manufactured date, heat treatment, wheel design, wheel type, plate design, and wheel diameter. Other features that need to be recorded when the wheel is being removed from the train are: the car type the wheel was on and the wheel location. Additional features are measured and recorded when the wheel is in a wheel shop, either because the wheel failed or its mate wheel needed service. These features change over time based on the age and condition of the wheel; they include the rim thickness, flange thickness and if the wheel is considered failed or safe.

For the cluster analysis study, we used the wheel manufacturer as a nominal feature. The data set included national and international train wheel manufacturers. Some of the manufacturers have multiple plants in which the wheels are manufactured; it has been assumed in the past by industry personnel that the different manufacturing processes of each plant may influence the wheels' lifetime. To take into account the manufacturing differences, the wheel manufacturer feature was defined by the manufacturing plant, and not simply the name of the manufacturer. In this data set, we used the specified manufacturing code abbreviation (AAR, 1998, 2005) stamped on the wheel at manufacture time as the feature value. The data set included 28 possible manufacturing codes.

The manufactured date is a numeric feature that is incremented in months. The dates range from April 1965 to September 2003 for this analysis. The months (1-12) and years (1965-2003) were considered separate features during this analysis; however, the

combined month-year manufactured date for an individual wheel is maintained because of the clustering process.

The heat treatment feature is defined by the markings that designate class and the method of heat treatment used on the steel (AAR, 1998, 2005). These markings indicate the class of steel that was used to manufacture the wheel as well as the heat treatment used: (i) either the entire wheel was heat treated, (ii) only the rim was heat treated, or (iii) the wheel was not heat treated at all. There are 5 possible heat treatment codes included in the data being used for this illustrative example.

The rim thickness is measured by taking a side scale reading (AAR, 1998); the side scale measures the amount of metal on the tread surface above the measuring line, in effect supplying a rim thickness measurement. The rim thickness is given in $1/16^{th}$ of an inch increments; a thicker rim has a larger side scale reading. The rim thickness is a numeric feature ranging from 18/16" (18) to 34/16" (34) in this analysis.

A finger reading is taken in order to get an idea of the flange thickness (AAR, 1998); the measurement starts at 0 (0/16") for a new wheel and as the flange gets thinner with wear, the measurement increases in increments of $1/16^{th}$ of an inch. The flange thickness is a numeric feature ranging from 0/16" (0) to 8/16" (8) in this illustrative analysis.

Wheel design is a nominal feature. There are many wheel designs which vary according to manufacturer and wheel diameter. In this illustrative analysis 13 wheel designs are included in the samples.

Wheels can be reshaped, known as "turning" the wheel in the railroad industry, after they have been in service and considerable wear of the tread surface has taken place.

Wheels are turned to specifications in accordance with the Association of American Railroads' (AAR) guidelines. Wheels are designed so that they can be turned either once, twice, multiple times (three or more), or not at all. This design parameter is known as the wheel type. The wheel type is a nominal feature which, in this analysis, includes three of the possible designs.

The plate design can be either a straight plate or a curved plate design. Most of the wheels manufactured during and after the mid-1970s are curved plate wheels; however, there are still straight plate wheels in service, so both types of wheels are included in this analysis. The plate design is a nominal feature.

The wheels in this analysis have diameters of 28, 33, 36, and 38 inches in size. As shown in Figure 11, 36 inch wheels are the largest percentage of wheels in the population. The wheel diameter is a numeric feature.

Different types of cars carry different loads and also have different activity levels; therefore the car type the wheel was on is an important feature that is included in this analysis. There are nine different car types included in the samples for this analysis.

Generally, wheels are in the 1^{st}, 2^{nd}, 3^{rd}, or 4^{th} position on a car on either the left or right side. This means that generally a wheel has eight possible locations on a car; location is considered as a nominal feature in this analysis.

The safe/failed indicator is a nominal feature indicating the condition of the wheel when the features discussed above were recorded. This feature includes two possible scenarios, either safe or failed.

Thus twelve features are considered: wheel manufacturer, manufactured date, heat treatment, wheel design, wheel type, plate design, wheel diameter, rim thickness, flange

thickness, the car type the wheel was on, the wheel location, and if the wheel is considered failed or safe. It is evident that there are many possible combinations of these 12 features. Because some of the features imply the value of other features, i.e. the wheel design indicates the wheel diameter, the total number of combinations isn't simply a product of the numbers of the individual possible feature values; however, the total number of feature combinations is still very large. This illustrative example obviously won't include all of the possible feature combinations in the real wheel population because only 250 samples are used; however, it will be able to indicate which feature combinations among the given samples have the highest percentage of failed wheels. In order to develop the actual inspection scheme for the railroad industry a much larger sample size will need to be used for the clustering study. This is pursued in Chapter VI.

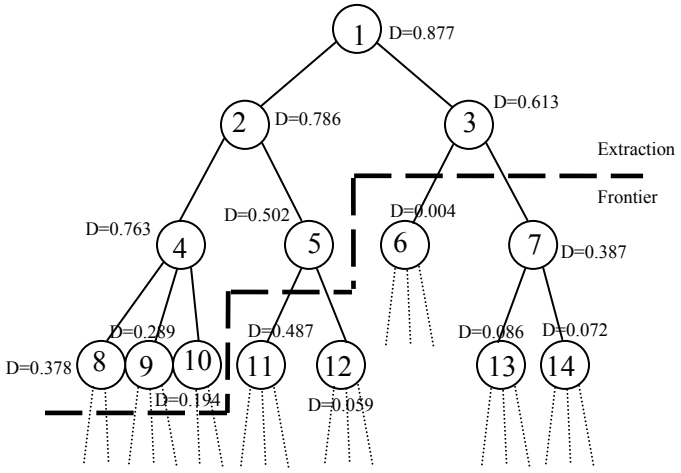


Figure 22. Illustrative example of how the threshold value (t = 0.25) determines the critical clusters

For this illustrative analysis the wheels were clustered twice, each time with a different threshold value, t. The threshold value determines the critical clusters from the

classification tree, as described above. The first analysis used a threshold valued of 0.4, and the second analysis used a threshold value of 0.25. These two threshold values were chosen because analyses with threshold values around 0.3 have been shown to produce the best results (Li and Biswas, 2002); 0.25 and 0.4 are on either side of 0.3 and were therefore included in the analysis.

The results of the first analysis with the threshold value set at 0.4 resulted in 15 clusters of interest. The results of the second analysis with a threshold of 0.25 resulted in 21 clusters of interest. It was concluded based on the partition structure that the threshold value in the first analysis (0.4) resulted in clusters that were fragmented but not cohesive to the desired level. This was determined because the clusters of interest identified from the first analysis contained more samples per cluster with lower percentages of failed wheels per cluster than the second analysis. The second analysis resulted in more cohesive clusters, some of which (the *critical clusters*) contained higher percentages of failed wheels. Therefore, the following section discusses the results from the second analysis.

Discussion of the Results

As stated above, the analysis with threshold value, t = 0.25, resulted in 21 clusters of interest (see Table 2). Had the failed wheels been evenly distributed throughout the entire wheel population, then each cluster would contain approximately 20 percent failed wheels (based on 50 failed wheels and 200 safe wheels in the overall population) and 80 percent safe wheels. Based on the results it is obvious that certain feature combinations result in more failed wheels. The corresponding clusters are the *critical clusters* in which

inspections should be focused.

Table 2. Results of clustering analysis

Cluster	Number of Samples in Cluster	Percent Failed Wheels in Cluster	Percent Safe Wheels in Cluster
A	10	0	100
B	27	0	100
C	8	100	0
D	9	0	100
E	8	0	100
F	2	50	50
G	15	0	100
H	15	0	100
I	10	30	70
J	1	0	100
K	3	0	100
L	6	0	100
M	9	11.1	88.9
N	15	53.3	46.7
O	18	44.4	55.6
P	5	0	100
Q	27	33.3	66.6
R	15	53.3	46.7
S	11	0	100
T	30	13.3	86.7
U	6	0	100

Clusters that contain considerably greater than 20 percent failed wheels were therefore considered as candidates for the critical clusters. However, it is also necessary to examine the number of total samples in a cluster to determine whether the percentage of failed wheels in that cluster is significant. For example, cluster F is composed of 50 percent failed wheels, however there are only 2 samples contained in that cluster. Because of the insignificant number of samples contained in that cluster, cluster F is eliminated as a candidate for a critical cluster.

From Table 2 and the above discussion it is seen that clusters C, N, O, and R have greater than 40 percent failed wheels in their cluster and contain a significant number of samples in their cluster; therefore they make up the critical clusters in which inspections should be focused.

It is important to note that the samples in a particular cluster will have similar feature combinations but they will generally vary slightly. Therefore each critical cluster needs to be analyzed further to determine which *feature combinations* specifically resulted in failures. Then those feature combinations in the critical clusters that resulted in failures will be the feature combinations towards which inspections are focused. To help understand this, an examination of critical clusters C, N, O and R follows.

A closer examination of critical cluster C shows that there are three *critical feature combinations* that comprise the cluster's makeup; i.e., of the eight samples contained in critical cluster C there were three critical feature combinations present. Most of the features are the same in the three critical feature combinations, which is why the samples were clustered together, but there are some features that vary per critical feature combination; the critical feature combinations for cluster C are listed in Table 3. Note, because of proprietary information, the actual results can't be provided so the numeric features have been changed and the nominal features have been disguised in Tables 3-6.

Table 3. Critical feature combinations from cluster C

Flange Thickness	Wheel Diameter	Wheel Manufacturer	Heat Treatment	Wheel Type	Plate Design	Wheel Design
2	36	BA	S	2W	CP	XY33RS
2	36	TD	S	2W	CP	XY33RS
2	36	AC	S	2W	CP	XY33RS

Note that only 11 of the 12 features can actually be used to identical critical wheels in the population because the safe/failed feature is only an indicator used to help identify the critical clusters. It is also important to point out that only 7 of the 11 features are shown in Table 3. The other four features appear to be insignificant. These four features are: the manufacturing date, car type the wheel was on, wheel location, and the rim thickness.

The manufacturing dates in cluster C covered a wide range of values, which means that feature is considered insignificant for predicting wheel failures for wheels with that specific feature combination. Practically this means a wheel with any manufacturing date is considered critical and needs to be removed if its feature combination follows one of the critical feature combinations given in Table 3.

The wheels in cluster C were on four out of the nine possible car types included in the analysis so the car type the wheel was on is also considered insignificant. Similarly, the wheel location and rim thickness are not important for the wheels in cluster C. As seen in Table 3, the flange thickness, wheel diameter, heat treatment, wheel type, plate design, and wheel design are the same for the three critical feature combinations obtained from the samples in cluster C. There were three different manufacturers included in cluster C, and because there were 28 wheel manufactures included in the overall analysis this was considered a significant feature. The only feature that is different in the three critical feature combinations for critical cluster C is the wheel manufacturer.

In summary, inspection resources should be partially focused on the three critical feature combinations identified from critical cluster C. Additional critical feature combinations in which inspections should be focused are identified from the other three

critical clusters as described below.

Analysis of clusters N, O and R follow the same procedure as above except that because clusters N, O and R contain some wheels that have not yet failed, only the samples that resulted in failures are further analyzed. Therefore, the candidate critical feature combinations in those clusters are reduced to only the feature combinations that resulted in the wheels failing and therefore need to be the focus of inspection. The critical feature combinations for cluster N, O and R are given in Tables 4, 5 and 6.

Table 4. Critical feature combinations from cluster N

Rim Thickness	Flange Thickness	Wheel Diameter	Heat Treatment	Car Type	Wheel Type	Plate Design	Wheel Design
19	2	38	S	BD	2W	CP	D38PR
18	2	36	S	BD	2W	CP	T36XX
20	2	36	S	BD	2W	CP	T36XX
19	2	36	S	BD	2W	CP	Y36FE

Table 5. Critical feature combinations from cluster O

Rim Thickness	Flange Thickness	Wheel Diameter	Wheel Manufacturer	Heat Treatment	Car Type	Wheel Type	Plate Design	Wheel Design
19	2	38	TD	S	G	2W	CP	JK38SG
20	2	38	TD	S	G	2W	CP	JK38SG
20	2	38	TD	V	G	2W	CP	JK38SG

Table 6. Critical feature combinations from cluster R

Flange Thickness	Wheel Diameter	Wheel Type	Plate Design	Wheel Design
1	33	2W	SP	J33HA
2	33	2W	SP	J33HA

Table 7. Critical feature combinations found from the SBAC analysis

Critical Feature Combination	Rim Thickness	Flange Thickness	Wheel Diameter	Wheel Manufacturer	Heat Treatment	Car Type	Wheel Type	Plate Design	Wheel Design
1	insignf.	1	33	insignf.	insignf.	insignf.	2W	SP	J33HA
2	insignf.	2	33	insignf.	insignf.	insignf.	2W	SP	J33HA
3	19	2	38	TD	S	G	2W	CP	JK38SG
4	20	2	38	TD	S	G	2W	CP	JK38SG
5	20	2	38	TD	V	G	2W	CP	JK38SG
6	19	2	38	insignf.	S	BD	2W	CP	D38PR
7	18	2	36	insignf.	S	BD	2W	CP	T36XX
8	20	2	36	insignf.	S	BD	2W	CP	T36XX
9	19	2	36	insignf.	S	BD	2W	CP	Y36FE
10	insignf.	2	36	BA	S	insignf.	2W	CP	XY33RS
11	insignf.	2	36	TD	S	insignf.	2W	CP	XY33RS
12	insignf.	2	36	AC	S	insignf.	2W	CP	XY33RS

These four critical clusters contain twelve critical feature combinations (given in Table 7) in which inspections should be focused (in Table 7, when *insign* (insignificant) is listed as a feature value it means that specific feature is insignificant for that critical feature combination and that any value for that feature is acceptable). These twelve critical feature combinations contain 52 percent of the failed wheels (26 out of the 50 failed samples). Out of the 250 samples included in the analysis, 39 had one of the critical feature combinations; i.e., 15.6 percent of the samples had a feature combination that was considered critical. This means that if inspection resources were focused on this critical 15.6 percent of the population then wheel failures could possibly decrease by approximately 52 percent.

The average yearly cost from wheel failures is approximately $6,000,000 when considering only derailment costs which were caused by shattered rim failures. Decreasing wheel failures by approximately 52% would result in annual savings of approximately $3,120,000 minus the costs of the additional inspections which will be calculated once the inspection times are optimized in future work. However, based on

preliminary calculations, the additional inspection costs will be approximately $2,000,000 per year; resulting in annual saving of approximately $1,120,000 while at the same time increasing the safety of the industry.

For results that will be implemented in industry, analyses need to be performed with much larger sample sizes and additional features. In order to get the best results large numbers of samples (on the order of hundreds of thousands to millions of samples) will need to be used. However, there are generally many more safe samples than failed samples (generally on the order of a few thousand). Because the clustering methodology's time demands grow exponentially it is infeasible to cluster the entire dataset. Therefore, a modification of the clustering methodology discussed above follows:

1. Cluster only the failed samples in order to find all of the *feature combinations*.
2. Once all of the *feature combinations* are found then they can be queried against all of the safe samples in order to determine the percentage of failed/safe samples for each *feature combination*.
3. These percentages are then used to identify the *critical feature combinations* in which inspections should be focused.

This modified methodology has been applied to the railroad wheel application, in Chapter VI, using the full dataset. As discussed in Chapter VI, this modified methodology successfully identified the critical feature combinations in which inspections should be focused in actual industry implementation.

Conclusion

Railroads generally remove and replace a large number of wheelsets per year using either visual inspections or Wheel Impact Load Detectors. However, these wheels aren't necessarily critical wheels that need to be taken out of service based on their likelihood of failing. Therefore, clustering analysis results could help the railroads focus their removals on actual critical wheels that have a higher probability of failing. The Similarity-Based Agglomerative Clustering algorithm has been shown to identify these critical wheels with much success.

It is important to note that railroads will need to inspect all of the wheels that have a specific critical feature combination. If only a portion of the wheels with that specific critical feature combination are inspected then the confidence level of the reliability assessment will fall. Further study needs to be done to explore how inspecting only a portion of the wheels with a specific critical feature combination will affect the assessment. It is also important to point out that each of the critical feature combinations will need to be optimized individually and will have individual inspection schedules, as discussed in Chapter V.

As noted above, once the critical wheels are identified, they can be taken out of service or inspected on an optimized inspection schedule. Many railroads are currently implementing technology that will allow them to identify a specific wheel while it is in-service. Being able to identify the critical wheels that are in-service will allow the railroads to focus their inspections on the critical wheels at the wheels' critical inspection times. Chapter V develops a methodology for optimizing the inspection schedules of the critical wheels, which will in effect prolong the usable lives of the wheels while

maintaining acceptable reliability levels for those wheels.

However, in order to optimize the inspection schedules of the critical feature combinations, first the reliability corresponding to each critical feature combination needs to be calculated. Therefore, the next chapter develops a simulation-based methodology for life prediction of wheels with the critical feature combinations. These life estimates are then implemented in the reliability-based inspection schedule optimization algorithm in Chapter V.

Note that while the reliability-based inspection scheduling algorithm is general, the clustering and reliability analysis are confined to shattered rim failures in this study.

CHAPTER IV

RELIABILITY ANALYSIS

The cluster analysis discussed in Chapter III was used to identify the critical feature combinations and their corresponding feature values that result in wheel failures. This chapter analyzes wheels with these critical feature combinations in order to find their corresponding reliability with respect to shattered rim failure. The reliability estimates are then used for the inspection optimization in Chapter V. The reliability is estimated through a combination of finite element-based stress analysis, multi-axial fatigue modeling and stochastic life prediction analysis.

Introduction

Damage accumulation due to fatigue, plastic deformation and wear significantly reduces the service life of railroad wheels. As discussed perviously, higher train speeds and increased axle loads have led to larger wheel/rail contact forces. Also, efforts have been made to optimize wheel and rail design. This evolution tends to change the major wheel rim damage from wear to fatigue (Tournay and Mulder, 1996). Unlike the slow deterioration process of wear, fatigue causes abrupt fractures in wheels, or tread surface material loss. These failures may cause damage to rails, damage to train suspensions and, in some cases, serious derailment of the train.

As discussed in Chapter I, railroad wheels may fail in different ways corresponding to different failure mechanisms (Stone and Moyar, 1989; Marais, 1998; Mutton et al, 1991). Shattered rim failure is the failure mode considered in this chapter.

The methods for fatigue life prediction of mechanical/structural components could be divided into several groups. Among these, fatigue crack initiation prediction models based on the *S-N* or *e-N* curve approach, and fatigue crack propagation prediction models based on fracture mechanics are predominantly used. The fatigue crack initiation models are appropriate for the analysis of components with non-crack-like geometries or without large initial defects. The fatigue crack propagation models are appropriate for the analysis of components with crack-like geometries or with large initial defects. If neither stage (initiation or propagation) dominates the entire life of the mechanical component, a total life methodology is required to accurately predict the component fatigue reliability.

There are two major difficulties in deterministic railroad wheel fatigue modeling. One is that the wheels are usually under rolling contact condition, which leads to a non-proportional multi-axial stress state within the wheels. An appropriate multi-axial fatigue model is required to handle this type of fatigue life prediction, which should be able to handle non-proportional loading conditions. The other difficulty is how to accurately describe the stress state in contact analysis. Analytical solutions and simplified 2D finite element models are not appropriate for the rolling contact analysis of mechanical components with complex geometries, such as railroad wheels (Liu et al, 2006).

A large amount of scatter has been observed in the fatigue life distribution of railroad wheels, ranging from several months to several decades. A probabilistic fatigue analysis is more appropriate in order to consider the large observed variability, due to variability in material properties, structural geometries and applied loadings. Due to the complex mechanism involved in rolling contact fatigue analysis and large number of

random variables affecting the final reliability, a direct analytical reliability calculation is impractical.

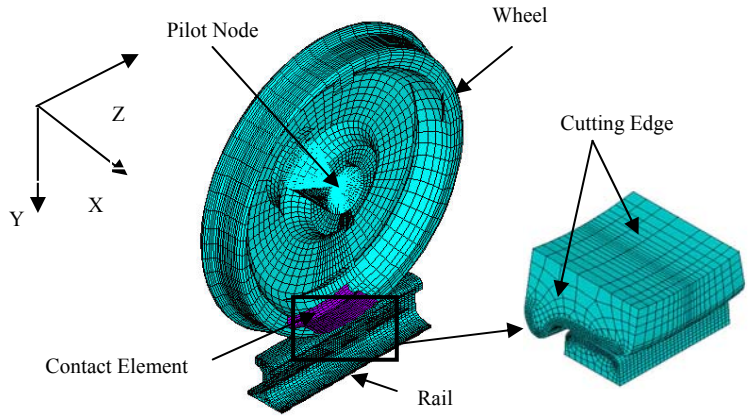
Therefore, a general methodology for rolling contact fatigue life prediction under a stochastic loading process has been proposed by Liu et al (2006). The total fatigue life of railroad wheels is separated into two parts, crack initiation life and crack propagation life

$$N_{total} = N_{initiation} + N_{propogation} \tag{18}$$

where N_{total}, $N_{initiation}$ and $N_{propogation}$ are the total fatigue life, fatigue crack initiation life and fatigue crack propagation life, respectively. The details regarding the calculation of each part of the fatigue life and the transition between crack initiation and crack propagation are discussed later in this section, and can also be found in (Liu and Mahadevan, 2005; Liu et al, 2006; Liu, 2006; Liu and Mahadevan, 2007). A brief illustration of the finite element models used for the fatigue crack initiation and fatigue crack propagation life predictions is shown below; details of the models can be found in Liu et al (2006) and Liu et al (2007).

Discussed first is the fatigue crack initiation life prediction model for railroad wheels. Initially, available profiles are used to build the geometry model of one wheel and a piece of rail. This model is called the full model. The rail length equals the length between two sleepers (which is assumed to be 600 mm). Fixed boundary conditions are applied to the two ends of the rail. Different mesh is applied to the full model using 3D elements (SOLID 45 in ANSYS). In the contact region, a relatively finer mesh is used. At the wheel center, a pilot node is connected to the wheel using some rigid link elements. All of the external loading and boundary conditions of the wheel are applied

on the pilot point. These loading and boundary conditions can be obtained through field measurements or from numerical simulation of the track system motion analysis. On the possible contact areas of the railhead and the wheel tread, area contact elements (CONTACT 174 and TARGET 170 in ANSYS) are used corresponding to the geometry mesh of the wheel. The contact algorithm is the augmented Lagrangian method (Ansys, 2003). Friction effect is included in the material properties of the contact elements. A Coulomb friction model is used in ANSYS. Friction coefficients can be calibrated using field measurement data. The material properties of the wheel and rail are described using a bilinear kinematic hardening model in ANSYS. The material properties of the rail and wheel are assumed to be the same (yielding strength = 500 MPa; Young's Modulus = 205 MPa; Friction coefficient = 0.25). No isotropic hardening is included in the current model. The finite element model is shown in Figure 23.



a) Full model b) Sub model

Figure 23. Finite element modeling of wheel/rail contact

A quasi-static analysis is performed for the full model and the results for each step are stored. Next, the geometry model of the contact region is cut out to be a sub-model. The size of the sub-model depends on the analysis objective and also on the wheel motion simulated. The same types of elements as those in the full-model analysis are used to mesh the sub-model. A very fine mesh is applied from the contact area to below the tread surface a few millimeters. The results of the full-model are interpolated on the cut boundaries of the sub-model corresponding to different calculation steps, and the interpolation results are applied as boundary conditions to the sub-model. The stress response from the sub-model is used for fatigue life prediction.

Next, a finite element model similar to the one for the fatigue crack initiation analysis is used for the fatigue crack propagation analysis. The difference is that an embedded elliptical crack is built into the model. The crack location and orientation are determined from the previous numerical prediction of the initial fatigue crack profiles. The major axis is along the track direction and the minor axis is perpendicular to the track direction. Based on the field observations of the initial fatigue crack profile, the aspect ratio of the elliptical crack is assumed to be 1.5. The subsurface crack is modeled as two contact surfaces to make sure that the two crack surfaces do not penetrate each other during the calculation. On the crack surfaces, area contact elements (CONTACT 174 and TARGET 170 in ANSYS) are used. Friction effect is also included between the two crack surfaces. A very fine mesh (the average element length is about 0.1 mm) is applied near the crack region. The finite element models of the full model, sub-model and crack are shown in Figure 24.

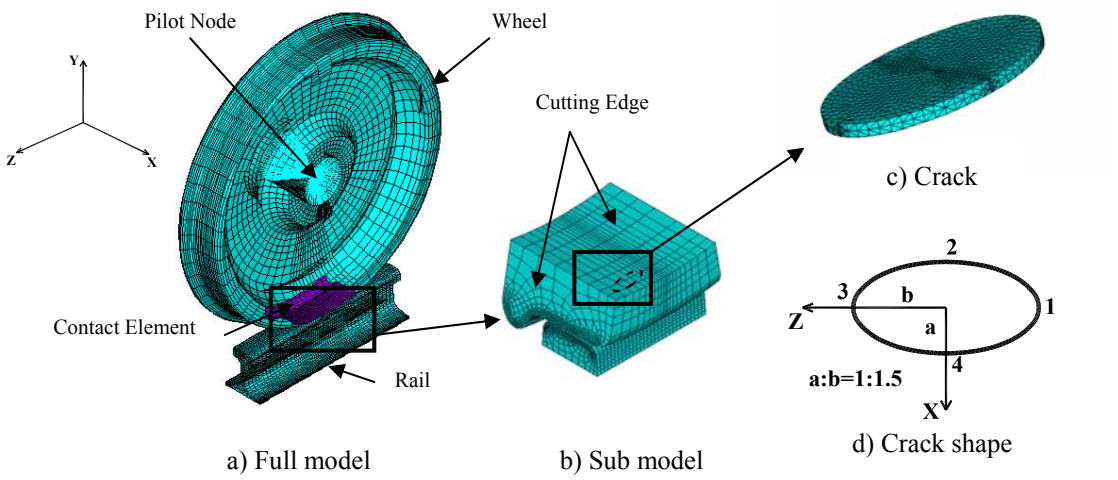


Figure 24. Finite element modeling of wheel/rail contact with subsurface crack

A quasi-static analysis is performed to simulate the rolling motion of the wheel. The calculated stress intensity factors (SIF) are used for the fatigue crack propagation analysis.

A response surface is developed from the finite element model simulations to calculate the fatigue crack initiation and propagation lives. A stress intensity factor (SIF) amplitude approximation is required for fatigue crack propagation life prediction. Liu (2006) has performed a parametric study for the fatigue crack propagation analysis of railroad wheels and found that the applied load, the crack length and the crack depth below the tread surface have significant influences on the SIF ranges. Based on first principles of fracture mechanics, a simple formula was proposed (Liu, 2006) to calculate the equivalent SIF range as

$$K_{eq} = \xi(F - F_c)\sqrt{\pi a d(2d_c - d)} \tag{19}$$

where K_{eq} is the equivalent SIF, F is the applied loading, a is the half length of the embedded crack along the axis, and F_c, d_c and ξ are regression constants. The

prediction using Equation (19) and the finite element results for the SIF along the major axis were shown to agree well (Liu et al, 2007). Therefore, the SIF equation replaced the finite element results for the fatigue crack propagation life prediction.

In addition to the features included in Liu (2006), the effects of rim thickness and plate design on the fatigue damage of the wheels were considered in this research. The plate design did not affect the fatigue damage, and was therefore excluded from the SIF equation. The rim thickness, on the other had, did significantly affect the fatigue damage, and was therefore included in the SIF equation in the form of a multiplier. The effect the rim thickness has on the SIF can be seen in Figure 25.

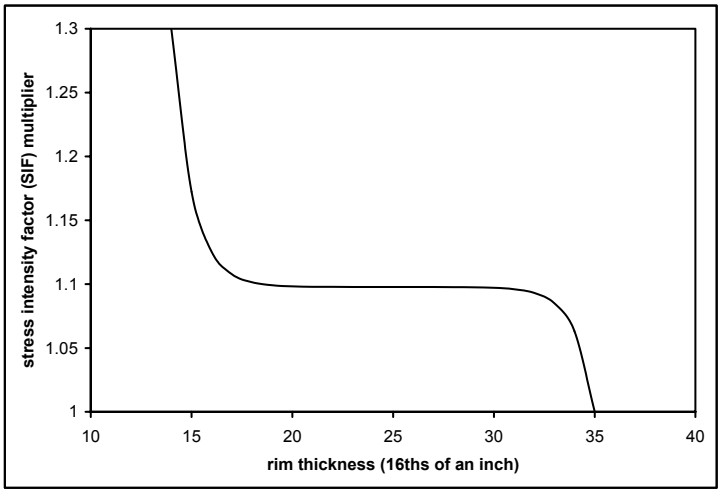


Figure 25. Stress intensity factor multiplier dependent on wheel rim thickness

The updated equation to include rim thickness in the equivalent SIF calculation is therefore

$$K_{eq} = \xi(F - F_c)\sqrt{\pi a d(2d_c - d)t_r} \tag{20}$$

where K_{eq} is the equivalent SIF, F is the applied loading, a is the half length of the

embedded crack along the axis, t_r is a multiplier based on the rim thickness (see Figure 25), and F_c, d_c and ξ are regression constants.

As mentioned above, the crack initiation life and crack propagation life are calculated using a stochastic fatigue life prediction approach proposed by Liu et al (2006). The details of the methodology are discussed below.

First consider the stress response at a specific location and time instant. The equivalent stress amplitude S_{eq} is calculated using the combination of a characteristic plane-based multi-axial fatigue theory (Liu and Mahadevan, 2005) and the 3D finite element model approach (Liu et al, 2006) discussed above. Where, at any fixed location, S_{eq} is a random variable with probability density function (PDF) of $f_{S_{eq}}(S_{eq})$, and the fatigue damage caused by the stress amplitude is usually expressed as a fraction of the total number of cycles to failure

$$D = \frac{1}{N} \tag{21}$$

where N is the fatigue life estimation from the S-N curve under constant stress amplitude S_{eq}. N represents the material resistance to fatigue loading. It is also a random variable at a specific stress amplitude. The conditional PDF of N can be found from experimental data and is expressed as $f_{N|S_{eq}}(N)$. The single cycle damage which considers both the randomness in material resistance and applied stress amplitude is a random variable whose joint PDF can be expressed as

$$f_D(D) = f_{D|S_{eq}}(D) f_{S_{eq}}(S_{eq}) = \frac{1}{D^2} f_{N|S_{eq}}\left(\frac{1}{D}\right) f_{S_{eq}}(S_{eq}) \tag{22}$$

For the fatigue damage accumulation process, a damage accumulation rule is

required. A linear damage accumulation rule, known as the Miner's rule, is used for its simplicity. Equation (23) is the general expression for the Miner's rule.

$$D_{initiation} = \sum_{i=1}^{K} D_i = \sum_{i=1}^{K} \frac{n_i}{N_i} \tag{23}$$

where K is the number of loading blocks, n_i is the ith applied loading cycle. For stochastic loading during a certain time period, not only is the stress amplitude S_{eq} a random variable, but also the number of cycles n_i at the stress amplitude S_{eq}. Nagode and Fajdiga (1998) proved that the conditional PDF ($f_{n_i|S_{eq}}(n_i)$) of number of cycles n_i at the stress amplitude level S_{eq} can be modeled by a normal distribution based on the DeMoivre–Laplace principle, with the mean and standard deviation expressed as

$$\begin{cases} \mu = 1 + (T-1)(1 - F(S_{eq})) \\ \sigma = \sqrt{(T-1)(1-F(S_{eq}))F(S_{eq})} \end{cases} \tag{24}$$

where T is the total number of load cycles during each block, $F(S_{eq})$ is the cumulative density function (CDF) of the stress amplitude which can be obtained from cycle counting techniques such as the rain-flow counting method (Liu, 2006). The joint PDF of the total damage at a specific location can be expressed as

$$f_{D_{tinitiation}}(D_{tinitiation}) = \frac{1}{D^2} f_{N_i|S_{eq}}\left(\frac{1}{D}\right) f_{n_i|S_{eq}}(n_i) f_{S_{eq}}(S_{eq}) \tag{25}$$

When the fatigue damage equals or exceeds unity, it is assumed that the initial fatigue crack is formed. The damage accumulation at different locations is checked. If fatigue damage exceeds unity at one location, the number of loading blocks is the fatigue crack initiation life of the structure. Equation (26) is a general expression for the structural fatigue crack initiation criterion

$$D_{structure} = max(D_{total,x_j}) = G(x_j, N_{initiation}, R_1, ...R_P) = 1 \quad (26)$$

where x_j is the coordinate at different locations, $N_{initiation}$ is the number of loading cycles to fatigue crack initiation, R_1 through R_P, are random variables which affect the fatigue damage in the structure. Solving Equation (26) for $N_{initiation}$ the following expression is obtained

$$N_{initiation} = f_i\left(x_j, R_1, ...R_P\right) \quad (27)$$

Equation (27) shows that the fatigue crack initiation life is a function of geometric locations and input random variables. The analytical solution for $N_{initiation}$ using Equation (27) is rather complicated and sometimes impractical. In the current study, a Monte Carlo simulation is used to calculate the probabilistic fatigue crack initiation life.

Once the fatigue crack is initiated, a fatigue crack propagation model developed by Liu (2006) is used to calculate the fatigue crack propagation life. A finite element model similar to the one for the fatigue crack initiation analysis is used. The difference is that an embedded elliptical crack was built into the model. The equivalent stress intensity amplitude, K_{eq} (Equation 20), used in this study was discussed earlier in this section. At any point at the crack tip, K_{eq} is a random variable with the probability density function (PDF) of $f_{K_{eq}}\left(K_{eq}\right)$. The crack growth rate at a specific SIF amplitude is also a random variable with the conditional PDF of $f_{\frac{da}{dN}|K_{eq}}\left(\frac{da}{dN}\right)$. Following a similar procedure as the stochastic fatigue crack initiation life prediction, the single cycle fatigue crack length increment Δa is a random variable and its joint PDF can be expressed as

$$f_D(\Delta a) = f_{\frac{da}{dN}|K_{eq}}\left(\frac{da}{dN}\right) f_{K_{eq}}(K_{eq}) \tag{28}$$

During the entire loading history, the crack propagation length is added to the initial crack length. When the crack reaches the critical length, the wheel is assumed to fail. The failure criterion is expressed as

$$a_{total} = a_i + \sum_{m=1}^{R} \Delta a_m \geq a_c \tag{29}$$

where R is the number of loading cycles, Δa_m is the crack length increment during each loading cycle. In Equation (29), a_i, Δa_m and a_c are random variables. a_i is calculated using Equation (23) and is related to material properties. a_c is obtained using field observations of failed components (failed railroad wheels in the current study). a_m is calculated using Equation (28) and is related to the applied stochastic loading and material properties. Solving Equation (29) for fatigue crack propagation, $N_{propagation}$, results in

$$N_{propagation} = f_p(a_i, a_c, R_1, \ldots R_p) \tag{30}$$

Equation (30) shows that the fatigue crack propagation life is a function of several random variables. Again, the analytical solution for $N_{propagation}$ using Equation (30) is rather complicated and sometimes impractical. In the current study, Monte Carlo simulations are used to calculate the probabilistic fatigue crack propagation life.

Substituting Equations (27) and (30) into Equation (18) obtains the total fatigue life of the wheel.

As discussed above, in the current study, a Monte Carlo simulation based on Equations (18) to (30) is used to calculate the probabilistic life distribution and reliability

degradation of railroad wheels. The statistics of material properties, wheel geometry, and applied loadings are included in the analysis.

Statistics of input random variables

Using the probability distribution functions of the input random variables a Monte Carlo simulation was used to predict the fatigue life of railroad wheels. Known feature values from the critical feature combinations were set constant and were no longer considered random variables. Each critical feature combination was analyzed separately based on its feature values. The details are discussed below.

The median fatigue S-N curve and its 90% confidence bounds are plotted with the experimental data in Figure 26. The fatigue life at a specific stress level is assumed to follow the lognormal distribution. The wheel diameter and rim thickness were set constant based on the critical feature combinations' feature values for those features. When the rim thickness feature value was *insign* (given in Table 7) then 14 (14/16ths; the minimum rim thickness generally allowed in service) was used.

The applied loading on the wheel is represented by a multinomial distribution. The applied loading depends on the car type the wheel was on, and each car type carries different cargo which affects the loading levels. The car type feature value for a specific critical feature combination determined the applied loading. No experimental data for the material hardness distribution is available. However, the hardness value for class B and C railroad wheels is bounded between 277 and 363 (AAR, 1998). In was assumed that the hardness follows a Beta distribution. The PDF of the beta distribution is plotted in Figure 27.

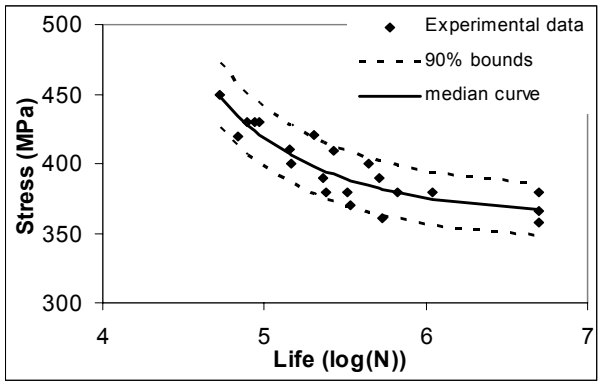


Figure 26. Fatigue S-N curve

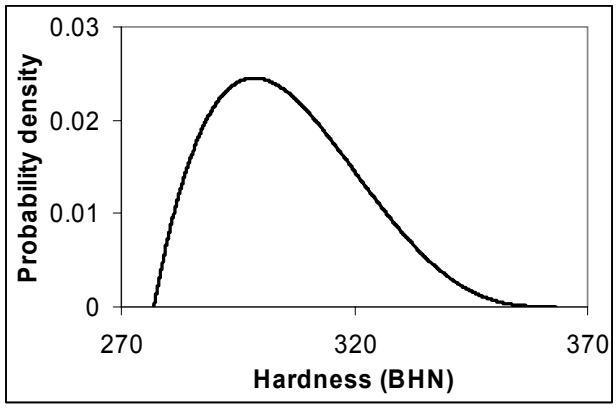


Figure 27. PDF of hardness distribution

The fatigue crack propagation curve suggested in AAR (1998) is used. The median and 90% confidence bounds are plotted in Figure 28. The crack growth rate at a specific stress intensity factor amplitude is assumed to follow the lognormal distribution. The final failure crack length (i.e., critical crack length) used field observation data and is approximated using a lognormal distribution. The histogram and lognormal fit for critical crack length are plotted in Figure 29.

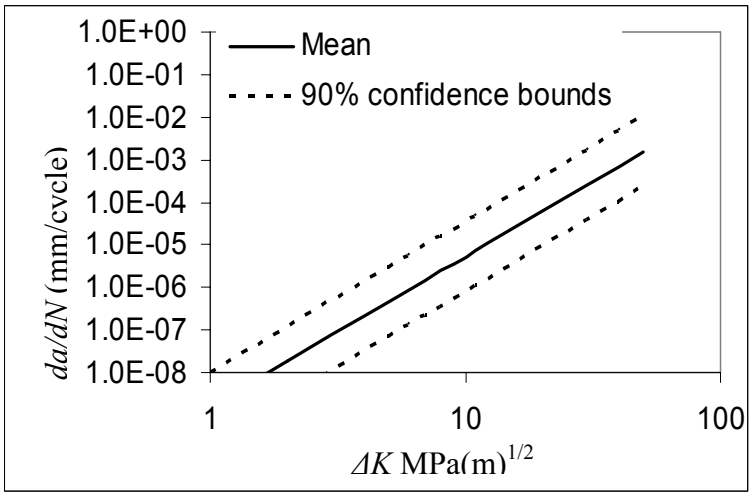


Figure 28. Fatigue crack growth curve

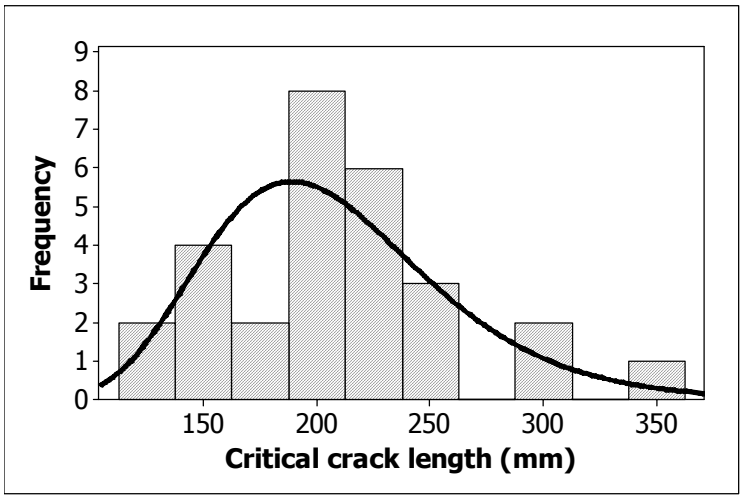


Figure 29. Histogram of critical crack length

Ultrasonic inspection (the detailed inspection technique for railroad wheels which will be used at the wheels' optimized inspection intervals) of railroad wheels only detects cracks that are 3.175 mm or greater. Therefore, the wheels' life prediction was performed with an initial crack size for all of the wheels set constant at 3.175 mm.

Probabilistic life distribution and reliability

Using the described statistics of the input random variables, Monte Carlo simulation were used to predict the fatigue life of railroad wheels. Monte Carlo simulation using the response surface is much more computationally efficient than with the finite element model; however, the computational time is still relatively large. Ten thousand Monte Carlo simulations were used to calculate the fatigue failure life of the railroad wheels.

It is important to note that the accuracy of the estimate depends on the number of simulation cycles. Haldar and Mahadevan (2000) define the percentage error, ε%, with a 95% confidence interval as

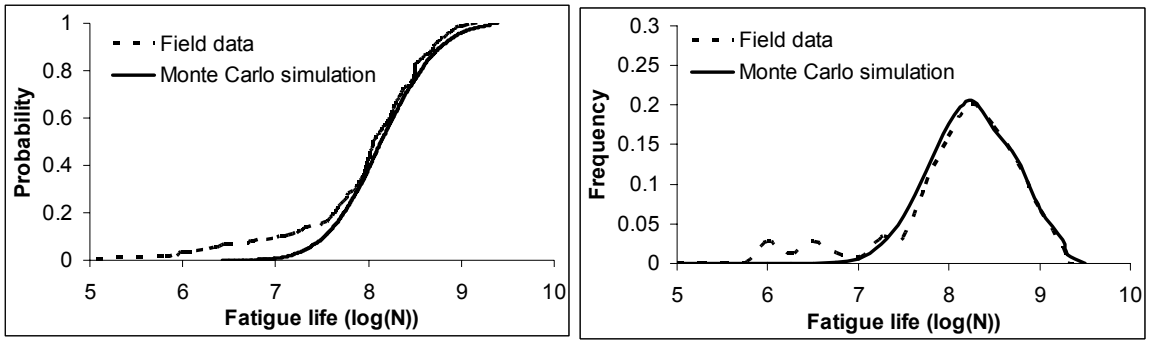
$$\varepsilon\% = \sqrt{\frac{\left(1 - p_f^T\right)}{N \times p_f^T}} \times 200\% \tag{31}$$

where p_f^T is the true probability of failure, and N is the total number of simulation cycles. As an example, if p_f^T is 0.1667 (corresponding to a reliability of 0.8333) and if 10,000 trials are used in the simulation, it indicates that there will be about 4.47% error. It can also be stated that there is 95% probability that the probability of failure, p_f, will be in the range of 0.1667 ± 0.000447 with 10,000 simulations.

For comparison purposes, a life prediction assuming no initial defects was performed. Field data regarding the number of cycles to failure of railroad wheels was collected and compared with the Monte Carlo simulation results that assumed no initial defects were present. The CDF and the frequency histogram of the numerical fatigue life prediction are plotted in Figure 30 together with the corresponding curves from field observations. The fatigue life in Figure 30 is censored at 2×10^9 cycles because it is

assumed that the wheel would fail due to other failure mechanisms by the end of this time-period.

In Figure 30(a), the numerical fatigue life predictions agree with the field observations reasonably well and capture the major trend of the life distribution. However, a large difference is observed at the early life regime, i.e. near the tail region of the fatigue life distribution. The reason is that the field data shows two ranges of fatigue life distribution. This phenomenon can be clearly seen in the frequency diagram (Figure 30(b)). It also can be seen that the number of wheels experiencing premature failure is only a small fraction of wheels (around 10%) and does not greatly affect the overall mean fatigue life. However, their effects are significant at the tail region, which affects the reliability evaluation.



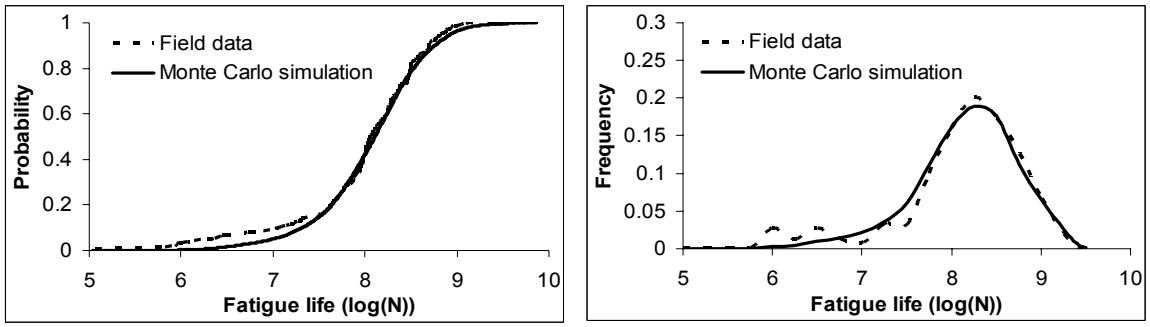
a) CDF b) Frequency histogram

Figure 30. Empirical CDF and frequency histogram of the field data and numerical predictions with no initial defects

Berge (2000) and Stone and Geoffrey (2000) suggest that the large stress, perhaps due to wheel/rail impact or material discontinuity, has an important effect on the shattered rim failure, as discussed in Chapter I. Also, the large on-tread brake loading

and the thermal stresses arising from on-tread friction braking will reduce the fatigue life of railroad wheels (Gordon and Perlman, 1998). The observed premature fatigue failure (earlier failure mode in Figure 30(b)) is possibly due to the above mentioned factors or their combinations, such as initial defects, brake loading, and thermal loading.

The brake loading and the thermal loading effects were beyond the scope of this study. Only the effect of initial defects was considered here to predict the premature fatigue failure. Based on the ultrasonic inspection capabilities, the length of the initial defect was assumed to be 3.175 mm (1/8 in). The location of the initial defect was assumed to be uniformly distributed between 5 mm to 8 mm below the tread surface. A Monte Carlo simulation was used again to calculate the fatigue life of defective railroad wheels, in which 10% of the failed wheels were assumed to be controlled by the large initial defects. The numerical prediction and field observations are plotted in Figure 31. It is seen that the numerical prediction are closer to the field observations when large initial defects were included in 10% of the wheels.

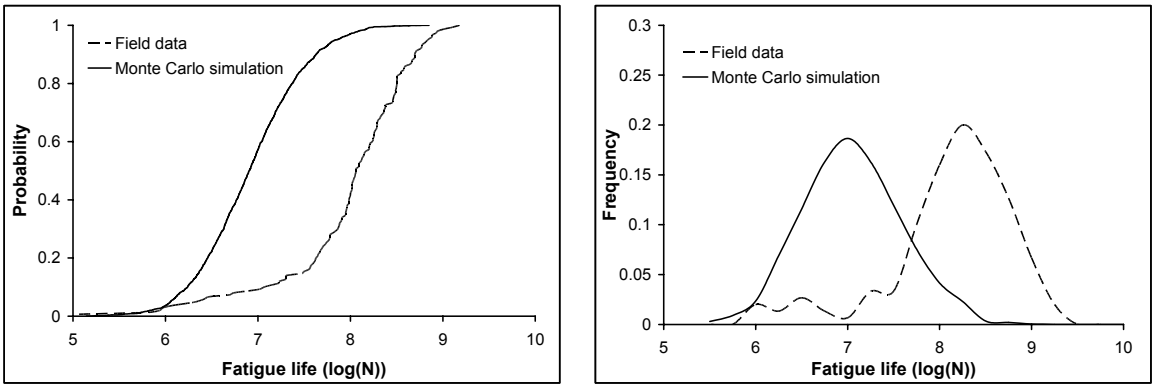


a) CDF b) Frequency histogram

Figure 31. Empirical CDF and frequency histogram of the field data and numerical predictions with 3.175 mm defects included in 10% of the wheels

As discussed in the previous section, the wheels' life prediction was performed with an initial defect size of 3.175 mm included in all of the wheels. This ultimately means the crack initiation life is not included in the total fatigue life (only the crack propagation life assuming an initial crack size of 3.175 mm is included in the total fatigue life).

This assumption is significant for actual implementation in the industry because it is possible that a wheel could pass the ultrasonic inspection at the wheel's optimized inspection time while containing a crack slightly less than 3.175 mm (because the ultrasonic inspection would not detect a crack smaller than 3.175 mm). This conservative assumption will ultimately decrease the inspection interval between the optimized inspection times, but it is necessary in order to maintain confidence in the optimized inspection schedule. The numerical prediction and field observations are plotted in Figure 32. As expected, when crack initiation is ignored in the total life of the wheel (since an initial crack of 3.175 mm was present in all of the wheels) the fatigue life predictions were drastically reduced, as can be seen in Figure 32.



a) CDF b) Frequency histogram

Figure 32. Empirical CDF of the field data and numerical predictions with 3.175 mm defects included in all of the wheels

The reliability of each critical feature combination was calculated separately, using the Monte Carol simulation method, based on the feature values of the critical feature combinations given in Table 7. The inspection intervals are in terms of days, therefore the fatigue lives were converted from cycles to failure to days to failure. This was done by first converting cycles to failure to miles to failure based on the wheel diameter. Next, the miles to failure were divided by the average miles per day the wheel traveled, as indicated by the car type feature value for the corresponding critical feature combination; this resulted in days to failure. The average miles per day each car type travels is shown in Table 8 (Guins, 2006). The fatigue life data, in terms of days to failure, was then fit to a curve using TableCurve 2D version 5.01. Therefore, each critical feature combination has a reliability curve associated with it that was calculated from its corresponding feature values.

Table 8. Data used to convert cycles to failure to days to failure

Car type feature value	Average miles per year	Average miles per day
A	19355	53
B	17300	47
C	15866	43
D	17300	47
E	7372	20
F	18192	50
G	13742	38
H	21844	60
J	62435	171
K	35468	97
L	28027	77
M	28027	77
P	60859	167
Q	60859	167
R	30637	84
S	60859	167
T	12406	34
V	49158	135

Illustrative Example's Reliability Curves

The reliability curves, calculated from the method described above, for the 12 critical feature combinations used in the illustrative example are shown in Figures 33-44. From the reliability charts (Figures 33-44) it is difficult to see the significant differences between the wheels during their early live regime; however the equations can not be given for proprietary reasons.

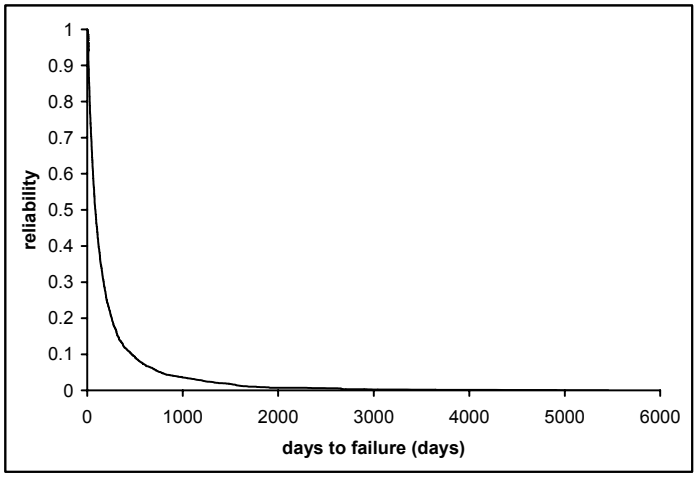


Figure 33. Reliability for critical feature combination 1 given in Table 7

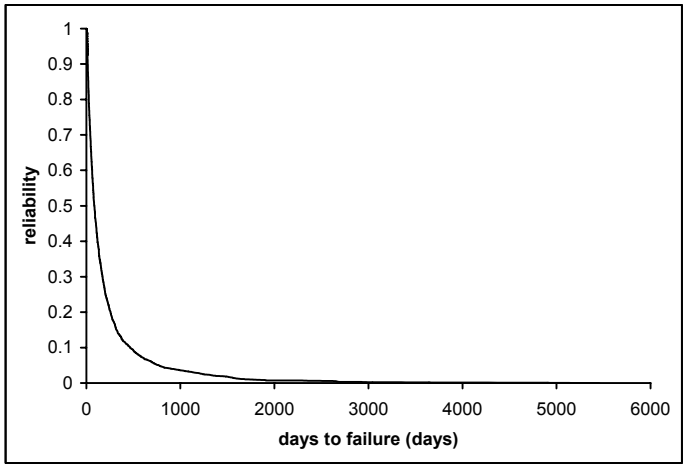


Figure 34. Reliability for critical feature combination 2 given in Table 7

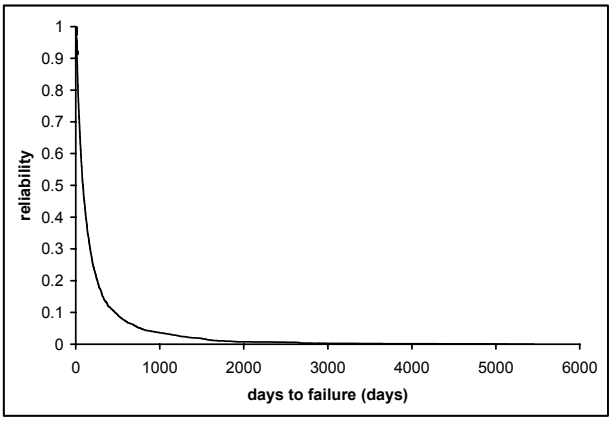


Figure 35. Reliability for critical feature combination 3 given in Table 7

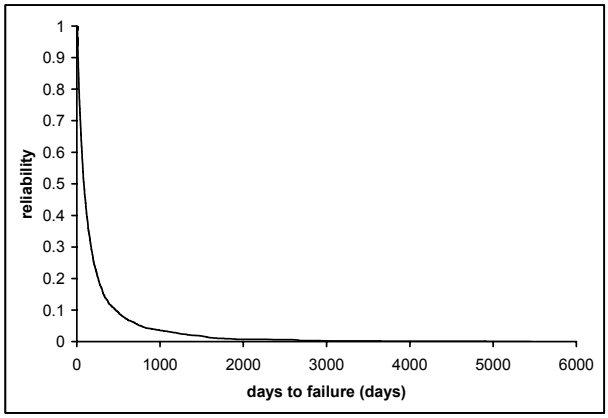


Figure 36. Reliability for critical feature combination 4 given in Table 7

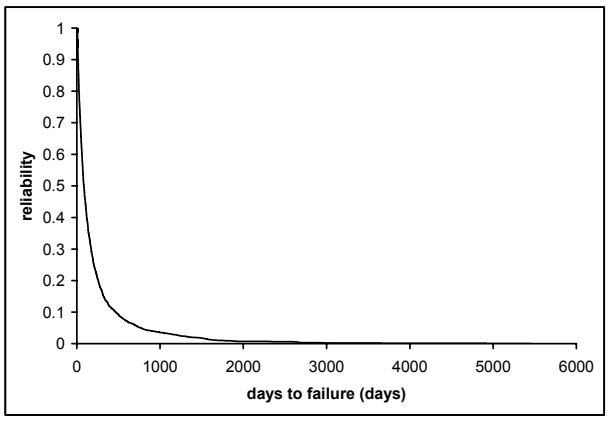


Figure 37. Reliability for critical feature combination 5 given in Table 7

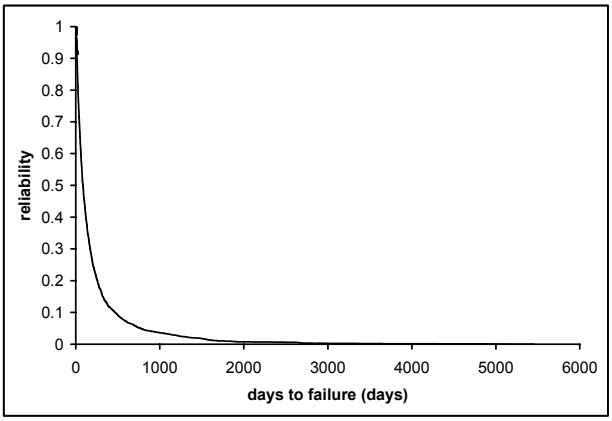


Figure 38. Reliability for critical feature combination 6 given in Table 7

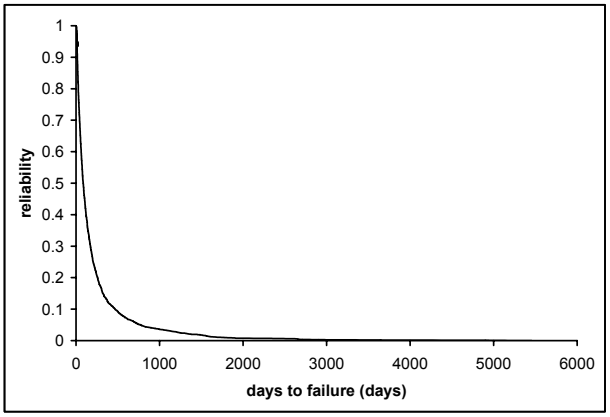


Figure 39. Reliability for critical feature combination 7 given in Table 7

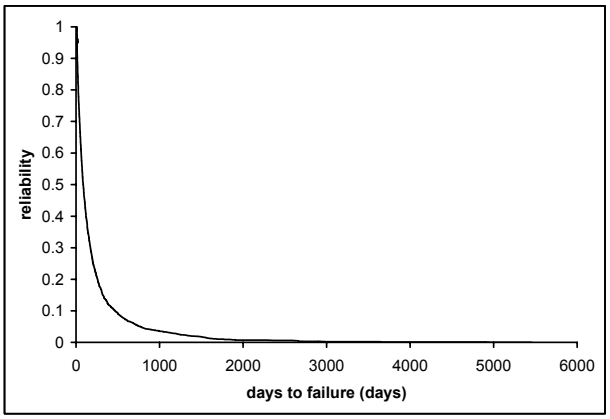


Figure 40. Reliability for critical feature combination 8 given in Table 7

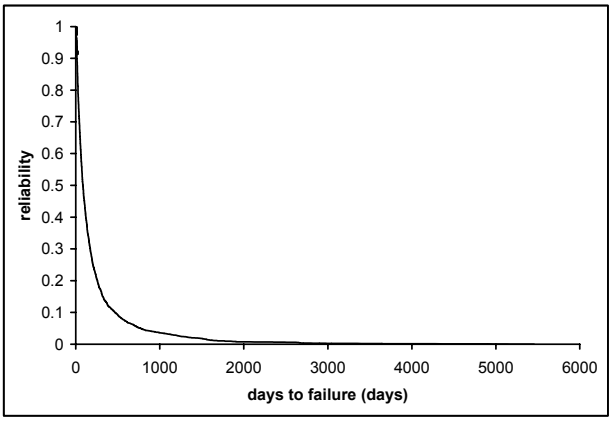


Figure 41. Reliability for critical feature combination 9 given in Table 7

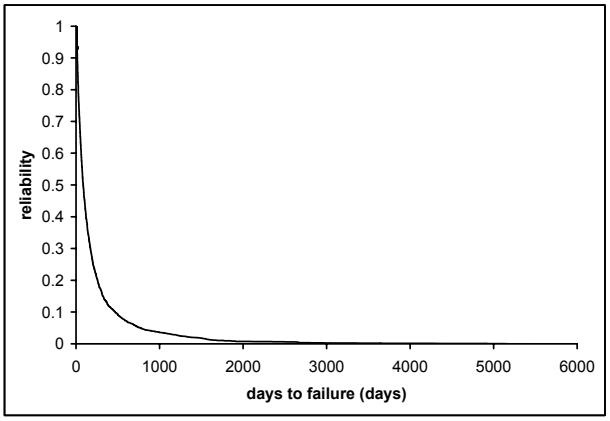


Figure 42. Reliability for critical feature combination 10 given in Table 7

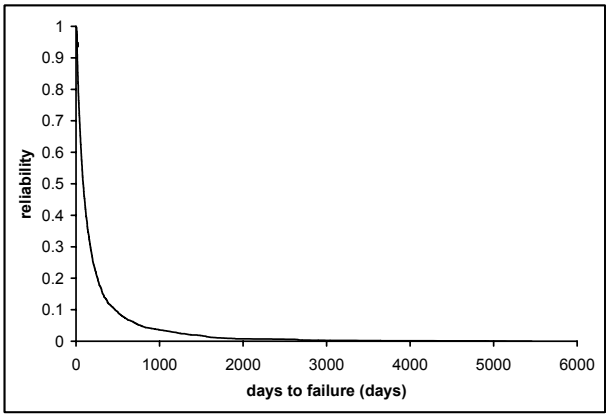


Figure 43. Reliability for critical feature combination 11 given in Table 7

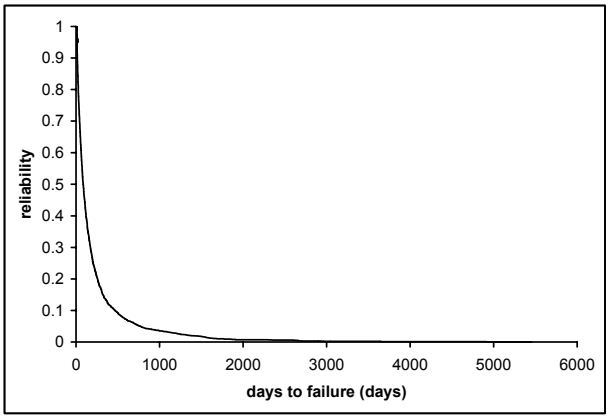


Figure 44. Reliability for critical feature combination 12 given in Table 7

Summary

The Monte Carlo simulation predictions provide a conservative life estimate, as compared to the field data, for each critical feature combination. However, a conservative life estimate is desirable in order to insure the reliability of the critical feature combinations are maintained by periodic inspections. The reliability curves for each critical feature combination are then used to optimize their inspection schedules, as discussed in the next chapter, using the inspection optimization methodology developed in the next chapter.

CHAPTER V

INSPECTION OPTIMIZATION

This chapter develops a methodology for optimizing inspection schedules of large populations, with the purpose of maintaining or exceeding the minimum target reliability level of the critical feature combinations (which were found in Chapter III) while minimizing inspection costs. The proposed methodology builds on the Similarity-Based Agglomerative Clustering (SBAC) algorithm (Li and Biswas, 2002) which was used to identify the critical feature combinations in Chapter III, and a multi-axial fatigue life prediction methodology (Liu et al, 2007; Liu and Mahadevan, 2005; Liu et al, 2006; Liu, 2006; Liu and Mahadevan, 2007) which was used in Chapter IV to calculate the reliability of wheels with the critical feature combinations. A reliability-based inspection optimization technique is developed in this chapter, in order to optimize the inspection schedules of wheels with the critical feature combinations.

Reliability-Based Inspection Optimization (RBIO)

Various tools and methodologies have been developed for fatigue reliability analysis and inspection updating (Faccioli et al, 1995; Onoufriou et al, 1994; Shetty et al, 1997). Garbatov and Soares (2001) proposed a method of optimized inspection planning for floating structures, in which inspection planning was treated as an optimization problem under reliability constraints. Zhang and Mahadevan (2001) proposed an approach for reliability-based reassessment of corrosion fatigue. The reassessment

approach incorporates the reliability of the NDI (non-destructive inspection) technique, inspection data, and prior prediction in a probabilistic framework for decisions regarding maintenance and repair. Other techniques and methodologies for reliability-based inspection optimization can be found elsewhere (Acar and Haftka, 2005; Chung et al, 2003; Estes and Frangopol, 2000).

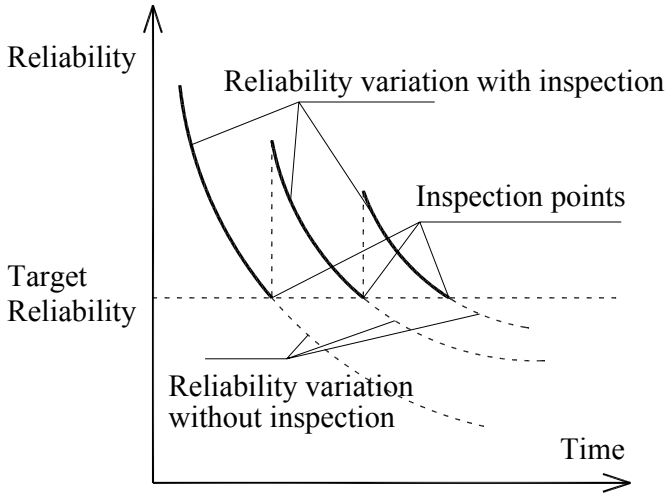


Figure 45. Reliability variation with and without inspection

For a specified target reliability (or equivalently, a minimum allowable safety level), the objective of reliability-based inspection optimization is to decide the inspection program that is the most economical while at the same time also maintains the reliability of the device or system above the target reliability. A graphical representation of this concept is given in Figure 45.

One application of RBIO was toward the inspection schedule optimization of an aircraft structure containing multiple site damage (Garbatov and Soares, 2001) in which the First Order Reliability Method (FORM) was used to calculate the probability of

failure. The initial crack lengths, crack growth equation constants, and the applied load were considered as random variables. The optimization of the inspection schedule for the structure was accomplished by minimizing the life cycle cost including materials, manufacturing, fuel consumption, and inspections, under the constraint of a pre-specified level of safety.

Another application of RBIO developed a methodology for optimizing the timing, frequency, and type of inspection over the expected useful life of a deteriorating structure such as an underwater cable (Estes and Frangopol, 2000). A decision tree analysis was used to develop an optimum lifetime inspection plan which could be updated as inspections proceeded and more data became available. This methodology included the expected life of the structure, the minimum prescribed safety level of the structure, costs of inspection and specific repairs, discount rates, the capability of the test equipment to detect a flaw, and the management approach of the owner towards making repairs.

In general, the objective function for a RBIO problem includes the cost of inspections, expected costs associated with any needed repairs, and expected costs associated with failure. The objective function to be minimized is total cost, while the reliability of the structure is kept above an acceptable level.

Railroad Wheel Inspection Optimization

The overall optimization problem may be stated as

$$\begin{aligned} &\text{Minimize} && \text{Overall cost} \\ &\text{s.t.} && \text{Reliability} > \text{Target Reliability} \end{aligned} \tag{32}$$

where inspection interval is the decision variable.

For this application, the objective function is found by summing the cost of

inspection; which includes both the cost of drive-by inspections, C_{I_d}, and the cost of ultrasonic inspections, $C_{I_{UT}}$, the cost of repairs, C_R, and the costs associated with a failure, C_F. The cost of drive-by inspections, C_{I_d}, is a sunk cost because they occur regardless of whether a failure, repair, or ultrasonic inspection takes place. In order to formulate the remaining cost terms for the objective function the four possible outcomes are considered: "no inspection and no failure", "no inspection and failure", "inspection and no failure", and "inspection and failure".

There are no costs associated with the "no inspection and no failure" case, therefore this term is left out of the objective function. For the "no inspection and failure case", the cost of repair and the cost of failure are dependent on the wheel failing. Therefore, C_R and C_F are both multiplied by the probability that the wheel will fail, p_f. Also, each term is divided by the number of days to inspection, T, because the daily cost (cost per day) is being minimized in the objective function. Therefore the costs associated for this case can be formulated as

$$\left(p_f \frac{C_R}{T}\right)+\left(p_f \frac{C_F}{T}\right) \tag{33}$$

For the "inspection and no failure" case, it is important to note that all of the critical wheels within the critical feature combination must be inspected at their optimized inspection time in order to maintain the desired reliability of that critical feature combination. As discussed previously, the detailed inspection method that is being used for this application is ultrasonic inspection, and is therefore the inspection method that is scheduled at time, T. Since all of the wheels in the critical feature

combination must be ultrasonically inspected at time T, then the cost of ultrasonic inspections, $C_{I_{UT}}$, needs to be multiplied by the number of wheels in the North American population, N, with that specific critical feature combination. Once again the cost is divided by T since the cost per day was being minimized. Also, for this case the wheels are being inspected but they don't fail, therefore the term is multiplied by $(1 - p_f)$. Therefore the costs for this case are formulated as

$$\left(\frac{C_{I_{UT}} \cdot N}{T}\right)(1 - p_f) \tag{34}$$

Finally, the "inspection and failure" case is formulated by summing the costs associated with inspecting the wheels and having a failure still occur. Included in this term are the same values associated with an ultrasonic inspection as in the "inspection and no failure" case, except this time the cost is multiplied by the probability of failure, p_f. Also included in this term are the costs associated with a repair and failure, C_R and C_F. However, for this case the wheels are inspected but they still failed, which means the crack wasn't detected by the ultrasonic inspection. Therefore, since the cracks weren't detected, the C_R and C_F terms are multiplied by $(1 - p_d)$, where p_d is the probability of detection. The cost for this case is thus

$$\left(\frac{C_{I_{UT}} \cdot N}{T}\right)(p_f) + \left(p_f \frac{C_R}{T}\right)(1 - p_d) + \left(p_f \frac{C_F}{T}\right)(1 - p_d) \tag{35}$$

The optimization problem finds the optimal interval between inspections for all of the wheels with that specific critical feature combination in the population. Based on the above discussion, the total cost per day, C_T, is the sum of the four cases and the sunk costs

$$C_T = \left(C_{I_d} \cdot N\right) + \left(p_f \frac{C_R}{T}\right) + \left(p_f \frac{C_F}{T}\right) + \left(\frac{C_{I_{UT}} \cdot N}{T}\right)\left(1 - p_f\right) + \left(\frac{C_{I_{UT}} \cdot N}{T}\right)\left(p_f\right) + \left(p_f \frac{C_R}{T}\right)\left(1 - p_d\right) + \left(p_f \frac{C_F}{T}\right)\left(1 - p_d\right) \tag{36}$$

Since all of the wheels in the critical feature combination are inspected daily with drive-by inspection, C_{I_d} is multiplied by *N*.

The above equation may be simplified by assuming that all of the cracks in the wheels are detectable since, in the reliability calculations in Chapter IV, all of the wheels are assumed to contain initial cracks of size 3.175 mm. It is also assumed that ultrasonic testing has a 100% probability of detection, p_d, for cracks that are 3.175 mm or larger. Based on these assumptions, the terms containing $(1 - p_d)$ are reduced to zero. Summing the remaining terms produces the final objective function

$$C_T = \left(C_{I_d} \cdot N\right) + \left(p_f \frac{C_R}{T}\right) + \left(p_f \frac{C_F}{T}\right) + \left(\frac{C_{I_{UT}} \cdot N}{T}\right) \tag{37}$$

The cost of repair included in the objective function, C_R, depends on the wheel design, the wheel size, and if the wheel was turned. Based on the wheels' critical features, found in the cluster analysis, the cost to replace a wheel with that specific feature combination is considered a constant.

The cost of a failed wheel could range from simply the cost to replace the wheel, which would be considered a replacement cost, to a $100,000,000 derailment. Because of the unpredictable damage a derailment will cause, the failure cost, C_F, is simply considered the average cost per derailment.

The cost of ultrasonic inspections, $C_{I_{UT}}$, is assumed to be the value of the average cost of ultrasonically inspecting wheels on a rip track. Specifically, this includes the

costs of the ultrasonic testing equipment, actually testing the wheels, as well as the costs associated with lost service time because of testing (called "setting the car out" in industry terms).

The constraints are formulated as follows. The optimized inspection interval, T, must be greater than or equal to a minimum inspection interval, T_{min}, as determined by the decision makers. Therefore the constraint is written as

$$T \geq T_{\min} \tag{38}$$

where T_{min} is equal to one day for this application. It is possible that the failure cost could outweigh the inspection costs so greatly that it would be more economical to rigorously inspect (ultrasonic test) all of the wheels with a critical feature combination every day than it would be to not rigorously inspect them and risk a possible derailment. However, it is not practically feasible to take a car out of service everyday to inspect its wheels. Therefore, the implemented inspection interval will need to be equal to or more than the time corresponding to the practically feasible inspection interval, as determined by the decision makers.

It is also required to keep the safety (or reliability) above a minimum acceptable level, R_{min}. The expected time-dependent reliability, $E[R(T)]$, was found during the reliability analysis as discussed in Chapter IV. Therefore the following constraint is

$$E\left[R(T)\right] \geq R_{\min} \tag{39}$$

As mentioned previously, the minimum allowable reliability level for the critical feature combinations, R_{min}, is chosen to be 0.8333 for this application.

Therefore, from the above discussion, the optimization problem for the inspection scheduling of railroad wheels was formulated as follows

$$\begin{aligned} \min \quad & C_T = \left(C_{I_d} \cdot N\right) + \left(p_f \frac{C_R}{T}\right) + \left(p_f \frac{C_F}{T}\right) + \left(\frac{C_{I_{UT}} \cdot N}{T}\right) \\ \text{s.t.} \quad & T \geq T_{\min} \\ & E\left[R(T)\right] \geq R_{\min} \end{aligned} \tag{40}$$

where,

C_F = cost of failure
C_{I_d} = cost of drive-by inspection
$C_{I_{UT}}$ = cost of ultra sonic inspection
C_R = cost of repairs
C_T = total cost per day
$E[R(T)]$ = the expected time-dependent reliability
N = number of wheels in the population with that critical feature combination
p_f = calculated probability of failure; which is dependent on T
R_{min} = the minimum acceptable reliability level
T = time of ultrasonic inspection
T_{min} = minimum inspection interval

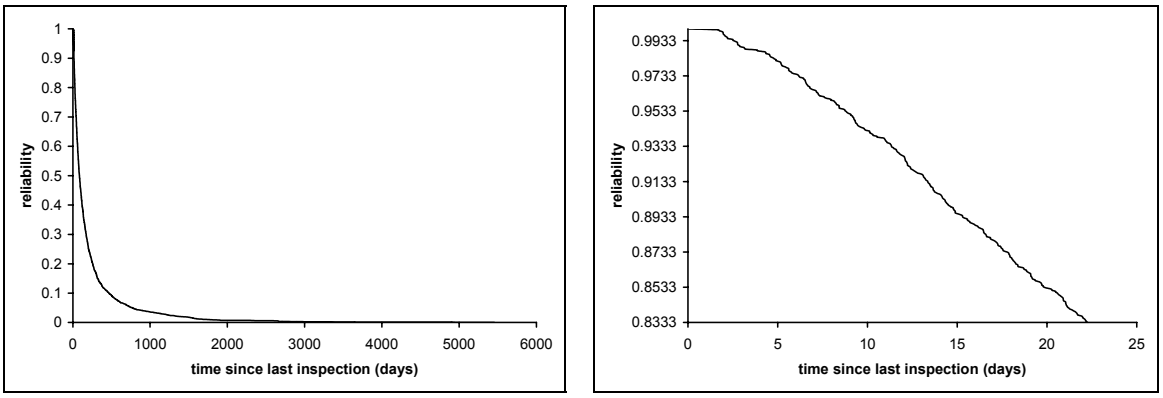
Equation 40 is a constrained nonlinear optimization problem, and can be solved using many different algorithms; the authors used Matlab's fmincon function (MathWorks, 2001), which employs a Sequential Quadratic Programming (SQP) method, for this purpose.

Discussion of the Railroad Wheel Inspection Optimization Illustrative Example Results

For the illustrative example in Chapter III, four critical clusters were found which contain twelve critical feature combinations (Table 7) in which inspections should be focused.

The reliability methodology proposed in Chapter IV was used to calculate the

reliability of wheels with each critical feature combination. The reliability curve of critical feature combination 2 is shown in Figure 46, as an example.



a) Reliability of Total Wheel Life b) Zoomed Reliability Curve of Figure 46(a)

Figure 46. Reliability of critical feature combination 2 from Table 7

An equation is fit to the above reliability curve so that the reliability, as a function of time, can be implemented in the optimization problem. The probability of failure equation is simply

$$p_f = 1 - Reliability = 1 - E[R(T)] \tag{41}$$

where T, in days, is the inspection interval being optimized. Both the reliability and probability of failure are functions of time. Each critical feature combination has its own reliability and probability of failure equations. Therefore, for this illustrative example there are twelve reliability and twelve probability of failure equations corresponding to their specific critical feature combination.

For this application the goal is to maintain the reliability of the critical feature combinations at or above 0.8333. Each critical feature combination will have a different reliability, and thus will have a different interval before required inspection (this time

corresponds to the reliability value of 0.8333).

Once the critical wheels are identified and their reliabilities are computed, a simple way to determine inspection interval is to set a reliability target (e.g., $R \geq 0.8333$ as in Figure 46(b)), and compute the corresponding inspection interval (e.g., 23 days in Figure 46(b), for critical feature combination 2). Each critical feature combination will need to be considered individually in order to get the inspection time for that specific critical feature combination. In order to maintain the reliability of the critical feature combination above the required safety level all of the wheels with that critical feature combination will need to be inspected at or before the optimized inspection time of that critical feature combination.

However, a more comprehensive approach is to consider various costs such as inspection cost, failure cost, etc. in deciding the inspection schedule. The optimization procedure developed above considers both reliability and cost. The results of the illustrative example's inspection scheduling optimization are discussed below.

The solution for the inspection interval is highly dependent on the number of wheels within the population that have a particular critical feature combination. If there were only 5,000 wheels within the population that have that specific critical feature combination then they could be inspected more often (which would keep the reliability higher than the minimum allowable reliability level) than if there were 10,000 wheels with that critical feature combination. Table 9 shows the number of wheels in the entire North American wheel population (including all railroads) with the given critical feature combinations (the actual numbers used in the illustrative example have been disguised).

Table 9. Number of wheels in the population for the individual critical feature combinations in the illustrative example

Critical Feature Combination	Approximate Number of Wheels in North America with given Feature Combination
1	678
2	678
3	339
4	452
5	452
6	339
7	565
8	226
9	339
10	904
11	452
12	226

The reliability is time dependent, which means that it changes over time, however, because of the rigorous inspection method it is assumed that the reliability of the wheels is reset to the initial reliability level (which assumes the wheels initially contain a 3.175 mm crack) after each ultrasonic inspection. This assumption may or may not be valid for different applications. Because the reliability is assumed to be reset after each ultrasonic inspection, the optimized inspection interval remains constant. Table 10 presents the optimized inspection intervals for the twelve critical clusters found in this example.

The more often wheels with a specific feature combination are inspected the higher the reliability those wheels will maintain. However, those inspections will also result in higher costs since they are performed more often than inspecting them at the minimum reliability level. A balance between cost and increased reliability above the set minimum reliability level can be made based on the overall objective of the industry.

Table 11 shows a cost/reliability comparison for the critical feature combinations discussed in this illustrative example.

Table 10. Optimized inspection intervals for individual critical feature combinations

Critical Feature Combination	Optimized Inspection Time (days between ultrasonic inspections)
1	9
2	23
3	1
4	5
5	8
6	9
7	4
8	8
9	4
10	1
11	25
12	1

For this illustrative example, the cost per year for the ultrasonic inspections was reduced by approximately $271,116 ($303,896 - $32,780) per year when the wheels were inspected at the times corresponding to their minimum required reliability levels, as opposed to inspecting them at their optimized inspection times. However, for some industries it may be beneficial to spend the additional money to maintain the increased reliability associated with inspecting the wheels at their optimized inspection times. It is also possible to inspect some critical feature combinations at the times corresponding to their minimum reliability levels (increase the inspection interval to decrease cost) while inspecting the other critical feature combinations at their optimal inspection intervals. The purpose of this would be to find the optimal balance between cost and reliability for

the industry.

Table 11. Total ultrasonic inspection costs for the additional inspections of the critical feature combinations

Critical Feature Combination	Optimized Inspection Interval (days)	Reliability of wheels with given Critical Feature Combination at Optimal Inspection Time	Cost per Year to Inspect all wheels with given Critical Feature Combination at Optimal Inspection Interval (dollars)	Inspection Interval at minimum Reliability level (days)	Cost per Year to Inspect all wheels with given Critical Feature Combination at minimum reliability level (dollars)
1	9	0.9545	11274	22	4612
2	23	0.8333	4411	23	4411
3	1	0.9999	50731	35	1449
4	5	0.9935	13528	37	1828
5	8	0.9830	8455	39	1734
6	9	0.9570	5637	26	1951
7	4	0.9875	21138	21	4026
8	8	0.9690	4228	27	1253
9	4	0.9895	12683	22	2306
10	1	0.9999	135284	26	5203
11	25	0.8333	2706	25	2706
12	1	0.9999	33821	26	1301
		Total cost per year to inspect all wheels with critical feature combinations at optimal inspection time	303896	Total cost per year to inspect all wheels with critical feature combinations at inspection time corresponding to minimum reliability level	32780

For this illustrative example the cost per year to inspect all of the wheels with all twelve critical feature combinations at their optimized inspection intervals is $303,896. As mentioned in Chapter III, if inspection resources were focused on these critical feature combinations then wheel failures could possibly decrease by approximately 52 percent. However, only 83.33 percent of the 52 percent of critical wheel failures will be prevented

since the probability of failure is being maintained at or above the 0.8333 level and not completely eliminated. The costs associated with wheel failures vary each year, however, the average failure cost per year for a major railroad company is approximately $6,000,000. This means that if 83.33 percent of the 52 percent of wheel failures were prevented then the railroad company would save approximately $2,600,000 on wheel failure costs. However, because of the additional inspection costs the actual savings are approximately $2,296,104 per year. This translates in to an approximate return on investment (ROI) of 755.6% when inspecting the wheels at their optimized inspection intervals. However, because of the practical infeasibility of inspecting the wheels as often as the optimized inspection intervals require, it is more realistic to inspect the wheels at the times corresponding to their minimum reliability level. Doing this would save the railroad company approximately $2,567,220 per year, resulting in a return on investment of 7831.7%. In addition to the monetary savings, the critical feature combinations would be more reliable than without the additional ultrasonic inspections, which also means increased safety of the railroad industry.

It is important to note that some features change over time based on the age and condition of the wheel; for this illustrative example these features include the rim thickness and flange thickness. Inspection intervals for wheels will be determined by the rim thickness and flange thickness at the current inspection time. Table 12 explains this scenario. For example, critical feature combinations 7 and 8 are exactly the same except that their feature values for rim thickness are different, and the corresponding inspection intervals are 21 days and 27 days respectively. If a specific wheel has critical feature combination 8 then it would be scheduled to be ultrasonically inspected in 27 days. If, at

that next ultrasonic inspection, the rim thickness decreased because of age and wear to have feature values corresponding to critical feature combination 7, then its next ultrasonic inspection would be scheduled for 21 days later.

Table 12. Optimized inspection scheduling table for critical feature combinations in illustrative example

Critical Feature Combination	Rim Thickness	Flange Thickness	Wheel Diameter	Wheel Manufacturer	Heat Treatment	Car Type	Wheel Type	Plate Design	Wheel Design	Inspection Interval at minimum reliability level (days)
1	insignf.	1	33	insignf.	insignf.	insignf.	2W	SP	J33HA	22
2	insignf.	2	33	insignf.	insignf.	insignf.	2W	SP	J33HA	23
3	19	2	38	TD	S	G	2W	CP	JK38SG	35
4	20	2	38	TD	S	G	2W	CP	JK38SG	37
5	20	2	38	TD	V	G	2W	CP	JK38SG	39
6	19	2	38	insignf.	S	BD	2W	CP	D38PR	26
7	18	2	36	insignf.	S	BD	2W	CP	T36XX	21
8	20	2	36	insignf.	S	BD	2W	CP	T36XX	27
9	19	2	36	insignf.	S	BD	2W	CP	Y36FE	22
10	insignf.	2	36	BA	S	insignf.	2W	CP	XY33RS	26
11	insignf.	2	36	TD	S	insignf.	2W	CP	XY33RS	25
12	insignf.	2	36	AC	S	insignf.	2W	CP	XY33RS	26

In summary, a wheel's reliability does change over time based on its current features. This means that a specific wheel can change critical feature combinations and will consequently change its inspection interval. When applying this method to large populations, there will be a much larger number of critical feature combinations which means once a wheel becomes a critical wheel it will always be associated with a critical feature combination regardless of its changing features. This is because the (numerical) feature values for the changing features, i.e. rim thickness and flange thickness, in the critical feature combinations will be ranges as opposed to a single value (this can be seen in the full analysis results in Chapter VI). Therefore, all of the possible values will generally be accounted for, so once a wheel becomes critical it will remain critical.

The total number of wheels associated with a specific critical feature combination

will also change because of the wheels' changing feature values. But, it is assumed that because the wheels age and wear at approximately the same rate that each critical feature combination will always have approximately the same number of wheels associated with it. Thus, the overall ultrasonic inspection costs will remain relatively constant over time; maintaining the cost savings for the railroad.

It is an option to re-analyze the data and perform a new cluster analysis once additional data is collected in order to "start over" and find new critical feature combinations. However, it is assumed that since the critical feature combinations have been identified, that they are the wheels that will continue to fail; it seems illogical that wheels with feature values that don't currently result in failures will start to fail simply because the current critical feature combination wheel failures will be prevented by the ultrasonic inspections.

However, as another option to increase the proposed methodology's effectiveness it is possible to maintain the critical feature combinations, but to update their inspection intervals based on new failure data. It is inevitable that some wheels with a critical feature combination will fail despite being inspected at their inspection interval. The reliability analysis, as developed in Chapter IV, assumes a 3.175 mm initial defect in the wheels. Also, ultrasonic inspection will detect cracks that are 3.175 mm in size or greater. This means that if a wheel was ultrasonically tested and failed before its next scheduled inspection, then the inspection interval should perhaps be tightened.

This can be done by using a Bayesian updating technique to update the reliability estimate, found from the methodology proposed in Chapter IV, with the new failure data. The updated reliability estimate can then be used to update the minimum allowable

inspection interval, as well as the probability of failure, p_f, used in the optimization problem. Once the updated equations are used in the optimization problem, a new optimized inspection interval can be calculated. Figure 47 shows the use of Bayesian updating for this purpose.

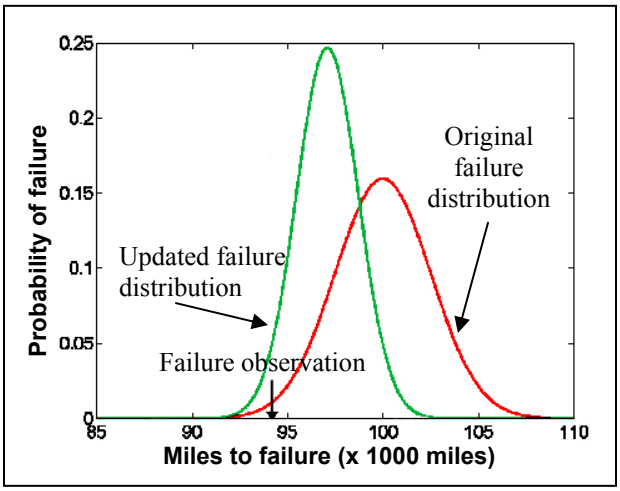


Figure 47. Updating the probability of failure using Bayesian methods

Conclusion

The concepts of clustering analysis (to identify critical samples), reliability analysis (to calculate the expected life of the critical samples), and reliability-based inspection optimization are combined into an overall methodology to optimize the inspection schedule of a large population. The proposed methodology will minimize the number of components inspected while at the same time maintain or exceed the desired reliability of the population.

To help expedite the inspection process an optimized inspection scheduling table (such as Table 12) should be created to list the critical feature combinations along with

their respective inspection intervals. The methodology is able to account for features that change over time as well as static features that remain the same. This overall methodology is suitable for any industry in which inspection scheduling optimization of large populations is needed.

The methodology developed in Chapters III, IV, and V is applied to a large population of railroad wheels in the next chapter.

CHAPTER VI

ANALYSIS OF LARGE RAILROAD WHEEL POPULATION DATA

The overall methodology developed in the previous chapters is applied to a large railroad wheel population data set in this chapter. A description of the full data set is provided in the following section, and then the results of the full analysis are presented and discussed.

Description of Data set

As with the illustrative example, the goal of the full analysis was to identify the critical wheels out of the entire population in which inspections should be focused. All of the samples contained a failed/safe feature value which was used to identify the critical feature combinations, i.e., feature combinations with the highest percentages of failed wheels.

The modified clustering methodology (discussed at the end of Chapter III) was used to identify the critical feature combinations. Initially the failed samples were clustered in order to find all of the *feature combinations*. Once all of the feature combinations were found then they were queried against all of the safe samples in order to determine the percentage of failed/safe samples for each feature combination. These percentages were then used to identify the *critical feature combinations* in which inspections should be focused.

The wheels were clustered based on 12 features: six numeric features and six nominal features (see Table 13). This full analysis included 112,218 wheels that had not yet failed (denoted as safe) and 1,078 wheels that had failed (denoted as failed), (i.e., total population = 113,296). The safe samples that were used in the cluster analysis were data from Union Pacific Railroad's database. The failed samples used data from the WABL_MD115 database and included both Why Made Codes 68 (vertical split rim failures) and 71 (shattered rim failures) because of data recording inconsistencies amongst railroads (i.e., it is not uncommon that a shattered rim failure is mistakenly recorded as a vertical split rim failure). The WABL_MD115 database contains confirmed failed wheels for all of the North American railroads.

Table 13. Wheel Clustering Features

Clustering Feature	Feature Type	Number of Possible Feature Values
manufactured year	Numeric	100
rim thickness	Numeric	41 (1/16" intervals)
flange thickness	Numeric	32 (1/16" intervals)
wheel diameter	Numeric	4
age (days)	Numeric	continuous values from 1 – 36122
gross ton mileage	Numeric	continuous values from 4470 – 843631210
heat treatment	Nominal	3
plate design	Nominal	2
car type the wheel was on	Nominal	17
wheel manufacturer	Nominal	155
cast/forged	Nominal	3
wheel design	Nominal	27

Many of these features have been discussed in Chapter III. The manufactured year is a numeric feature that is incremented in years. The dates range from 1907 to 2006 for this full analysis. The rim thickness and flange thickness are both numeric features which are measured following the procedures discussed in Chapter III. The rim thickness ranged from 0/16" (0) to 40/16" (40) in this analysis and the flange thickness ranged from 0/16" (0) to 31/16" (31). The wheels in this analysis had diameters of 28, 33, 36, and 38 inches in size.

The age and gross ton mileage are numeric features which have a wide range of values. The age of the wheel was calculated and recorded in days for this analysis. It is important to point out that the age changed depending on the wheel failure/recording date, whereas the wheel manufactured year stayed the same regardless of when the wheel failed or the data was recorded. The gross ton mileage was calculated by multiplying the total miles the wheel had traveled by the maximum weight, in tons, the wheel was certified to carry (including car and cargo weight).

The heat treatment was a nominal feature that had 3 possible values in this analysis: B, C, U. The plate design feature designated either a straight plate or a curved plate design (nominal feature with 2 possible feature values). The feature indicating the car type the wheel was on was also a nominal feature which had 18 possible feature values.

The wheel manufacturer was considered a nominal feature which included both national and international train wheel manufacturers. This feature indicated which plant of its manufacturer that it was manufactured in; 155 possible feature values were included in the analysis.

The cast/forged feature indicated if the wheel was cast or forged when it was manufactured. Wheels manufactured by Griffin were assigned one cast designation, while wheels manufactured by Southern were assigned a different cast designation. The wheels manufactured by all of the other manufacturers were assigned a forged designation. This resulted in 3 possible feature values for this feature.

The final feature included in this analysis, was the wheel design. The wheel design is a nominal feature which signifies the design of the wheel. There are many wheel designs which vary according to manufacturer and wheel diameter. In this analysis 27 wheel designs were included.

Thus, the twelve features included in this analysis were: manufactured year, rim thickness, flange thickness, wheel diameter, age, gross ton mileage, heat treatment, plate design, car type the wheel was on, wheel manufacturer, cast/forged, and wheel design. The results of the full analysis are presented and discussed in the following section.

Results of the Full Railroad Wheel Data Analysis

All of the failed wheels were clustered using SBAC with a threshold value t = 0.25. From the cluster analysis 32 clusters were found for the failed wheels. The 32 clusters were then analyzed further to find 363 feature combinations that could identify specific groups of failed wheels. All of the feature combinations were then queried against the entire population of failed and safe samples in order to determine the percentage of failed and safe wheels that each feature combination would identify out of the entire population. This analysis resulted in 214 critical feature combinations; a critical feature combination was determined to be a feature combination that contained

equal or more failed samples than safe samples when the entire population of both failed and safe wheels was queried (i.e., a feature combination that included four (out of the total 1,078 failed samples) failed samples was considered critical when four (out of the total 112,218 safe samples) or less safe samples was also included in the queried feature combination). The critical feature combinations in which inspections should be focused are presented in Table 14 in the order of significance based on the percentage and number of failed wheels in the critical feature combination. In Table 14, the manufacturer feature is masked for proprietary reasons. It is important to note that the number of failed wheels and the percentage of failed wheels in the critical feature combination are given in Table 14 only to show how the critical feature combinations were ranked; they are not used to identify the wheels which need inspection.

The criteria for each feature in Table 14 are discussed below:

1. A critical feature combination that included samples with a range of 10 or more years for the manufactured year was assigned a feature value of *insign* (i.e., insignificant) for the manufactured year feature.
2. A critical feature combination that included samples with a range of 10 or more (i.e., 10/16 inch or more) for the rim thickness was assigned a feature value of *insign* for the rim thickness feature.
3. For the flange thickness a range was only included when the lower value of the range was greater than zero (i.e., 3 or 5), otherwise it was assigned *insign* for its feature value.
4. A diameter was listed in Table 14 when all of the samples in the critical feature combination had the same diameter; if all of the samples didn't

have the same diameter then *insign* was assigned for the diameter feature value.

5. If the range of samples in the critical feature combination had age values of more than 365 days then it was assigned *insign* for the age feature value. If the range was within 365 days then each end of the range was rounded to the nearest 25 days (rounded range could exceed 365 days).

6. The gross ton mileage must have had all samples within 10% of each other (i.e., range's upper bound must be equal or less than 110% of range's lower bound). If the range was within 10% for all of the samples in the critical feature combination then each end of the range was rounded to the nearest 1000000 miles (rounded range could exceed a difference greater than 10% in feature values).

7. The heat treatment feature values that were given in Table 14 required that all of the samples for that critical feature combination have the same feature values (*B*, *C*, or *U*). A feature value of *HT* was assigned for critical feature combinations that contained samples with heat treatment feature values of both *B* and *C*, which simply means that the wheel was heat treated in one way or another. If all of the samples didn't have the same heat treatment, or a combination of *B* and *C*, then *insign* was assigned for that feature value.

8. If all of the samples in the critical feature combination didn't have the same plate design feature value, then *insign* was assigned for the plate design feature value for that critical feature combination.

9. If all of the samples in the critical feature combination didn't have the same car type feature value, then *insign* was assigned for the car type feature value for that critical feature combination.

10. If all of the samples in the critical feature combination didn't have the same manufacturer feature value, then *insign* was assigned for the manufacturer feature value for that critical feature combination.

11. All of the samples in a critical feature combination had to have the same feature value to be listed in Table 14 for the cast/forged feature value (*C1*, *C2*, or *F*). If the samples in the critical feature combination contained both cast designations (*C1* and *C2*) then *Cast* was assigned as the feature value. If all of the samples in the critical feature combination were not the same, or some combination of the two cast designations, then *insign* was assigned for the cast/forged feature value for that critical feature combination.

12. If all of the samples in the critical feature combination did not have the same wheel design feature value, then *insign* was assigned for the wheel design feature value for that critical feature combination.

Table 14. Critical feature combinations in which inspections should be focused

critical feature combination number	manufactured year	rim thickness	flange thickness	diameter	age (days)	gross ton mileage	heat treatment	plate design	car type	manufacturer	cast/forged	wheel design	number of failed wheels	percentage of failed wheels in the critical feature combination
1	insign	14-19	insign	36	insign	insign	C	CP	C	P	C1	CJ36	32	100
2	1993-2000	17-32	insign	36	insign	insign	C	CP	C	P	C1	CJ36	21	100
3	1990-1999	12 to 22	insign	33	insign	insign	C	CP	S	E	C2	CJ33	13	100
4	1991-1999	15-19	insign	33	insign	insign	C	CP	S	Q	C1	CJ33	12	100
5	1990-1995	15-20	0 to 2	38	insign	insign	C	CP	S	B	F	B38	11	100
6	1993-1998	19	insign	36	insign	insign	C	CP	J	insign	insign	insign	10	100
7	insign	14-16	insign	36	insign	insign	C	CP	C	P	C1	CJ36	8	100
8	2000-2004	19-24	0	33	insign	insign	C	CP	S	B	F	J33	8	100
9	1981-1988	20-26	insign	36	insign	insign	C	CP	C	P	C1	CJ36	8	100
10	1968-1972	16-24	insign	36	insign	insign	C	SP	T	B	F	H36	8	100
11	1998	20-22	0	33	insign	insign	C	CP	Q	T	F	J33	7	100
12	1981-1982	20-22	insign	36	insign	insign	U	CP	C	G	C1	CH36	7	100
13	1987-1997	13-20	0	33	insign	insign	C	CP	Q	Q	C1	CJ33	7	100
14	1979-1985	19-25	insign	36	insign	insign	U	CP	C	E	C2	CH36	7	100
15	1990-1999	15-22	insign	36	insign	insign	C	CP	J	P	C1	CJ36	7	100
16	1994-1998	23-25	0	36	insign	insign	C	CP	J	insign	F	insign	7	100
17	insign	22	0	36	insign	insign	C	CP	C	P	C1	CJ36	6	100
18	1990-1996	15-20	0	33	insign	insign	C	CP	V	Q	C1	CJ33	6	100
19	1991-1998	18	insign	36	insign	insign	C	CP	C	P	C1	CJ36	6	100
20	1967-1973	19-24	insign	36	insign	40000000-46000000	C	SP	T	insign	F	H36	6	100
21	1970	14-21	insign	36	insign	insign	C	SP	insign	Y	F	H36	6	100
22	insign	17	insign	36	insign	insign	C	CP	C	P	C1	CJ36	5	100
23	1991-1997	20	insign	36	insign	insign	C	CP	C	P	C1	CJ36	5	100
24	1993-1994	14-16	0	33	insign	insign	C	CP	Q	B	F	J33	5	100
25	1989-1996	19	0	36	insign	insign	C	CP	C	P	C1	CJ36	5	100
26	1991-1995	19	insign	36	insign	insign	C	CP	J	E	C2	CH36	5	100
27	1995-1997	18-24	0	36	insign	insign	C	CP	C	V	C1	CH36	5	100
28	1974	21-23	insign	36	insign	62000000-64000000	U	SP	H	D	F	H36	5	100
29	1973-1977	29-35	insign	36	insign	insign	C	SP	H	Y	F	J36	5	100
30	1971-1976	26-32	insign	36	insign	insign	C	SP	H	B	F	J36	5	100
31	1989-1998	17	insign	36	insign	insign	C	CP	J	V	C1	CH36	5	100
32	insign	16-18	0	36	insign	insign	C	CP	E	P	C1	CJ36	5	100
33	1969-1970	16-20	insign	36	insign	92000000-99000000	C	SP	H	B	F	H36	5	100
34	1992-1999	19-20	0 to 1	38	insign	insign	HT	CP	S	insign	Cast	insign	5	100
35	1991-1994	14	0	33	insign	insign	C	CP	insign	B	F	J33	5	100
36	1991-1996	17-21	insign	36	insign	insign	C	CP	C	P	C1	CH36	4	100
37	1990-1996	24	0	36	insign	insign	C	CP	C	E	C2	CH36	4	100
38	insign	21	0	36	insign	insign	C	CP	C	P	C1	CJ36	4	100
39	1988-1994	28-32	insign	36	insign	insign	C	CP	C	P	C1	CJ36	4	100
40	1980-1986	21-26	insign	36	insign	insign	U	CP	C	B	F	H36	4	100
41	1974-1979	21-26	insign	36	insign	insign	U	SP	C	insign	F	insign	4	100
42	1984-1994	12 to 14	0	33	insign	insign	C	CP	S	Q	C1	CJ33	4	100
43	1975-1979	19-24	insign	36	insign	insign	U	CP	C	insign	F	H36	4	100
44	1991-1997	19	insign	36	insign	insign	C	CP	J	G	C1	CH36	4	100
45	1995-1998	26	0	36	insign	insign	C	CP	J	E	C2	CH36	4	100
46	1996-2001	28-33	0	36	insign	insign	C	CP	J	P	C1	CJ36	4	100
47	1994-1995	21-26	0	36	insign	insign	C	CP	J	C	F	H36	4	100
48	1995-1998	21-24	insign	36	insign	insign	C	CP	J	T	F	H36	4	100
49	1989-1994	13-16	0	33	insign	insign	C	CP	P	E	C2	CJ33	4	100
50	1990-1999	19-20	0	33	insign	insign	C	CP	V	B	F	J33	4	100
51	1968-1969	16-19	2 to 4	36	insign	264000000-294000000	U	SP	J	R	F	J36	4	100
52	1990-1993	20	insign	36	insign	insign	C	CP	J	V	C1	CH36	4	100
53	1993-1995	19	insign	36	insign	insign	C	CP	K	E	C2	CH36	4	100
54	1996-2000	16-18	insign	38	insign	insign	C	CP	S	E	C2	CB38	4	100
55	1993-2001	22-24	insign	36	insign	insign	C	CP	Q	P	C1	CJ36	4	100
56	1981-1984	14-19	insign	36	insign	insign	U	CP	J	E	C2	CJ36	4	100
57	1967-1968	18-25	insign	36	insign	insign	U	SP	H	L	F	J36	4	100
58	1978-1988	14-17	insign	33	insign	insign	U	CP	A	Q	C1	CJ33	4	100
59	1971-1977	19-28	insign	36	insign	insign	C	SP	insign	Y	F	J36	4	100
60	1965-1973	19	insign	36	insign	insign	U	SP	insign	insign	F	J36	4	100
61	insign	17-18	insign	33	insign	insign	C	SP	insign	insign	F	J33	4	100
62	1973-1975	insign	0	36	insign	43000000-46000000	U	SP	C	E	C2	CJ36	3	100
63	1968-1970	17-20	insign	36	insign	insign	C	SP	C	Y	F	H36	3	100
64	1991-1994	24	0	36	insign	insign	C	CP	C	P	C1	CJ36	3	100
65	1988-1995	23	insign	36	insign	insign	C	CP	C	P	C1	CJ36	3	100
66	1996-2003	25-26	insign	36	insign	insign	C	CP	C	P	C1	CH36	3	100
67	1980-1987	25-28	insign	36	insign	insign	U	SP	C	insign	F	H36	3	100
68	1980	20-23	0	36	insign	insign	U	SP	C	insign	F	H36	3	100
69	1992-1993	19	insign	36	insign	insign	C	CP	J	V	C1	CH36	3	100
70	1979	18-20	2 to 3	36	insign	insign	U	SP	C	K	F	H36	3	100
71	1971-1977	18-21	insign	36	insign	insign	U	SP	C	E	C2	CH36	3	100
72	1968-1972	18-20	insign	36	insign	insign	U	SP	C	B	F	H36	3	100
73	1968-1976	26-31	insign	36	insign	insign	C	SP	J	B	F	J36	3	100
74	1976	19-24	insign	36	insign	insign	U	SP	T	L	F	H36	3	100
75	1982-1992	16	0	36	insign	insign	C	CP	J	G	C1	CH36	3	100
76	1980-1984	16	insign	36	insign	insign	U	CP	J	E	C2	CH36	3	100
77	1991-1999	20	0	36	insign	insign	C	CP	J	Q	C1	CH36	3	100
78	1989-1990	insign	0	36	insign	52000000-58000000	C	CP	J	insign	insign	insign	3	100
79	1988-1996	18	0	36	insign	insign	C	CP	J	V	C1	CH36	3	100
80	1995	18	insign	36	insign	insign	C	CP	J	V	C1	CH36	3	100
81	1981-1987	19-21	0	36	insign	41000000-42000000	U	CP	H	E	C2	CH36	3	100
82	1992-1998	21-22	0	33	insign	insign	C	CP	A	P	C1	CJ33	3	100
83	1993-1995	11 to 20	0	33	insign	insign	C	CP	V	E	C2	CJ33	3	100
84	1988-1993	17-22	0	36	insign	insign	C	CP	S	Q	C1	CH36	3	100
85	1994-1995	16	insign	36	insign	insign	C	CP	C	E	C2	CH36	3	100
86	1994-1996	18-23	5 to 9	36	insign	insign	C	CP	P	P	C1	CJ36	3	100
87	1975-1977	18-21	insign	33	insign	54000000-58000000	C	SP	A	B	F	J33	3	100
88	1967-1968	30-32	insign	36	insign	79000000-84000000	U	SP	H	R	F	J36	3	100
89	1993-1995	17-19	insign	36	insign	insign	C	CP	K	V	C1	CH36	3	100
90	1973-1975	22-24	insign	36	insign	insign	U	SP	E	D	F	H36	3	100
91	1967-1970	17-19	insign	36	insign	insign	C	SP	H	Y	F	H36	3	100
92	1988-1989	12 to 21	insign	33	insign	insign	U	CP	C	A	F	J33	3	100
93	1971-1973	insign	0	36	insign	insign	U	SP	insign	E	C2	CJ36	3	100
94	1994-1995	23-25	insign	36	insign	insign	C	CP	insign	D	F	H36	3	100

Table 14 -- continued

critical feature combination number	manufactured year	rim thickness	flange thickness	diameter	age (days)	gross ton mileage	heat treatment	plate design	car type	manufacturer	cast/forged	wheel design	number of failed wheels	percentage of failed wheels in the critical feature combination
95	1978-1979	18-23	insign	36	insign	insign	B	SP	insign	L	F	H36	3	100
96	1986-1988	0	18-20	33	insign	insign	U	SP	insign	insign	F	J33	3	100
97	1991-1999	<=15	0	36	insign	insign	C	CP	C	V	C1	CH36	2	100
98	1981-1984	17-24	insign	36	insign	43000000-46000000	C	CP	C	E	C2	CH36	2	100
99	1987-1994	17	0	36	insign	20000000-22000000	insign	CP	C	B	F	insign	2	100
100	1982	22-23	0	36	insign	35000000-36000000	C	CP	C	G	C1	CH36	2	100
101	1995-1996	28	0	36	insign	insign	C	CP	C	E	C2	CH36	2	100
102	1970-1976	21	0	36	insign	insign	U	SP	C	Y	F	H36	2	100
103	1974-1976	21-24	0	36	insign	insign	C	SP	C	B	F	H36	2	100
104	1971-1973	28-32	insign	36	insign	49000000-50000000	C	SP	C	E	C2	CJ36	2	100
105	1995	25-26	insign	36	insign	insign	C	CP	C	C	F	H36	2	100
106	1987	21-23	insign	36	insign	insign	U	CP	C	Q	C1	CH36	2	100
107	1972-1982	21-24	insign	36	insign	insign	U	CP	C	V	C1	CH36	2	100
108	1994-1995	17-25	0	36	insign	26000000-28000000	C	CP	A	E	C2	CH36	2	100
109	1994	20-21	0	33	insign	63000000-64000000	C	CP	Q	T	F	J33	2	100
110	1988-1994	20	insign	36	insign	insign	C	CP	C	V	C1	CH36	2	100
111	1979-1980	20-22	insign	36	insign	35000000-36000000	U	CP	C	Z	F	CJ36	2	100
112	1987-1996	19	0	36	insign	insign	C	CP	C	V	C1	CH36	2	100
113	1986-1989	19-21	0	36	insign	28000000-32000000	U	CP	C	M	C1	CH36	2	100
114	1986-1995	19-21	0	36	insign	19000000-20000000	insign	CP	C	E	C2	CH36	2	100
115	1995-1999	15	0	36	insign	insign	C	CP	J	G	C1	CH36	2	100
116	1989-1995	19	insign	36	insign	insign	C	CP	J	E	C2	CJ36	2	100
117	1975-1976	19-22	insign	36	insign	insign	C	SP	J	B	F	J36	2	100
118	1972-1979	18-22	0	36	insign	insign	B	SP	C	E	C2	CJ36	2	100
119	1973-1977	18-24	0	36	insign	insign	C	SP	C	Y	F	J36	2	100
120	1999	21	0	36	insign	insign	C	CP	J	E	C2	CH36	2	100
121	1995	24	0	36	insign	insign	C	CP	J	E	C2	CH36	2	100
122	1999	26-28	0	36	insign	42000000-43000000	C	CP	J	F	F	H36	2	100
123	1995-1997	27	0	36	insign	insign	C	CP	J	E	C2	CH36	2	100
124	1986	23-25	2 to 4	36	insign	86000000-87000000	U	CP	J	E	C2	CH36	2	100
125	1969-1970	28-30	2 to 4	36	insign	241000000-266000000	B	SP	J	Y	F	J36	2	100
126	1970-1975	29-33	0	36	insign	77000000-82000000	C	SP	H	Y	F	H36	2	100
127	1993-1995	21	insign	36	insign	19000000-20000000	C	CP	H	E	C2	CH36	2	100
128	1995	23-24	insign	36	insign	22000000-24000000	C	CP	H	E	C2	CH36	2	100
129	1971-1975	28-30	0	36	insign	insign	C	CP	H	Y	F	J36	2	100
130	1979-1980	21-26	insign	36	insign	insign	U	SP	H	E	C2	CH36	2	100
131	1977-1978	19-24	3 to 5	36	insign	30000000-34000000	U	SP	T	H	F	H36	2	100
132	1974-1975	19-21	insign	36	insign	35000000-38000000	B	SP	T	R	F	H36	2	100
133	1979-1988	17	0	36	insign	153000000-157000000	C	CP	J	P	C1	CJ36	2	100
134	1982-1984	16-19	0	36	insign	insign	U	CP	H	E	C2	CJ36	2	100
135	1997-1998	16	insign	36	insign	insign	C	CP	J	V	C1	CH36	2	100
136	1995-1997	16-18	insign	36	insign	insign	C	CP	J	G	C1	CH36	2	100
137	2000-2003	22-23	insign	38	insign	insign	C	CP	S	G	C1	CB38	2	100
138	1993-1994	24-29	0	38	insign	insign	C	insign	S	B	F	insign	2	100
139	1968	20-24	2 to 4	36	insign	252000000-271000000	U	SP	J	L	F	J36	2	100
140	1998	19-24	0	36	insign	insign	C	CP	K	T	F	H36	2	100
141	1989-1990	insign	0	36	insign	insign	C	CP	J	V	C1	CH36	2	100
142	1990	insign	0	36	insign	126000000-138000000	C	CP	J	insign	Cast	CH36	2	100
143	1985-1986	18-21	insign	36	insign	40000000-42000000	C	CP	H	E	C2	CH36	2	100
144	1993-1995	27	0	36	insign	19000000-21000000	C	CP	A	E	C2	CH36	2	100
145	1997-1998	18	0	36	insign	insign	C	CP	J	V	C1	insign	2	100
146	1993-1996	18-19	0	36	insign	71000000-75000000	C	CP	J	P	C1	insign	2	100
147	1996-2000	18-25	insign	36	insign	insign	C	CP	J	P	C1	CH36	2	100
148	1968-1977	18-21	insign	36	insign	insign	U	SP	J	L	F	H36	2	100
149	insign	18-22	4 to 8	36	insign	insign	U	SP	J	B	F	J36	2	100
150	1992-1996	19-20	0	38	insign	insign	C	CP	S	V	C1	CB38	2	100
151	1994-1999	16	insign	36	insign	insign	C	CP	C	V	C1	CH36	2	100
152	1972-1974	23-28	0	36	insign	23000000-26000000	U	SP	E	L	F	insign	2	100
153	1985	19	0	36	insign	41000000-42000000	C	CP	F	P	C1	CH36	2	100
154	1991	21	0	36	insign	insign	C	CP	K	V	C1	CH36	2	100
155	1975-1976	28-32	insign	36	insign	117000000-120000000	C	SP	K	B	F	J36	2	100
156	1979	21-23	0	36	insign	insign	U	SP	T	E	C2	CH36	2	100
157	1973	24	insign	36	insign	insign	U	SP	T	Y	F	H36	2	100
158	1970-1974	23-24	insign	36	insign	52000000-55000000	U	SP	T	B	F	H36	2	100
159	1981-1984	24-26	insign	36	insign	25000000-28000000	U	CP	T	Q	C1	insign	2	100
160	1995	16-19	0	33	insign	17000000-20000000	C	CP	A	Q	C1	CJ33	2	100
161	1979	16-19	insign	33	insign	52000000-55000000	U	CP	A	E	C2	CJ33	2	100
162	1996-1998	20	0	33	insign	insign	C	CP	V	P	C1	CJ33	2	100
163	1982-1988	17	0	36	insign	insign	C	CP	F	P	C1	CJ36	2	100
164	1985-1992	20-24	0	36	insign	43000000-46000000	C	CP	M	P	C1	CJ36	2	100
165	1990- 1991	16	0	36	insign	32000000-36000000	C	CP	F	E	C2	insign	2	100
166	1994-1997	18-19	4	36	insign	insign	C	CP	K	G	C1	CH36	2	100
167	1981-1985	18	0	insign	insign	insign	U	CP	K	Q	C1	insign	2	100
168	1997-2005	21-26	0	33	insign	insign	C	CP	S	M	C1	CJ33	2	100
169	1987-1992	18-19	0	36	insign	insign	C	CP	S	P	C1	CJ36	2	100
170	1969	16-17	0	36	insign	insign	C	insign	K	S	F	H36	2	100
171	1983-1988	16-18	insign	36	insign	insign	U	CP	K	E	C2	CH36	2	100
172	1994-2000	19	insign	33	insign	68000000-70000000	C	CP	Q	B	F	J33	2	100
173	1990-1994	12 to 16	0	36	insign	insign	C	CP	E	Q	C1	CH36	2	100
174	1980-1983	insign	0	36	insign	insign	C	CP	E	E	C2	insign	2	100
175	1986-1989	14-21	insign	36	insign	insign	U	CP	E	Q	C1	CH36	2	100
176	1985-1989	16	0	33	insign	insign	C	CP	P	insign	F	J33	2	100
177	insign	16-23	2	36	insign	insign	C	CP	M	Q	C1	CH36	2	100
178	1985-1988	14-21	insign	36	insign	insign	C	CP	A	P	C1	CJ36	2	100
179	1975-1976	insign	insign	36	insign	68000000-71000000	C	SP	H	E	C2	CJ36	2	100
180	1988-1992	insign	insign	36	insign	18000000-20000000	C	CP	F	B	F	H36	2	100
181	1981-1987	24-25	0	36	insign	insign	C	insign	P	insign	F	J36	2	100
182	1988-1991	insign	0	36	insign	insign	C	CP	insign	N	F	CJ36	2	100
183	1971-1976	insign	insign	36	insign	73000000-76000000	C	CP	insign	B	F	J36	2	100
184	1975-1985	13	0	33	insign	insign	U	SP	insign	insign	F	J33	2	100
185	1979-1989	12 to 13	0	33	insign	insign	insign	CP	insign	P	C1	CJ33	2	100
186	1991-1994	14-15	2 to 5	33	insign	insign	C	CP	insign	insign	F	J33	2	100
187	1994-1998	28	1	36	insign	insign	C	CP	insign	E	C2	CH36	2	100
188	1986	29-31	0	36	insign	insign	U	CP	insign	E	C2	CJ36	2	100

Table 14 -- continued

critical feature combination number	manufactured year	rim thickness	flange thickness	diameter	age (days)	gross ton mileage	heat treatment	plate design	car type	manufacturer	cast/forged	wheel design	number of failed wheels	percentage of failed wheels in the critical feature combination
189	1978	27-28	2	36	insign	insign	U	CP	insign	X	C1	insign	2	100
190	1972	27-31	insign	36	insign	insign	U	SP	insign	Y	F	J36	2	100
191	1968-1971	21-27	insign	36	insign	insign	C	SP	insign	S	F	J36	2	100
192	1970-1971	19	insign	36	insign	insign	U	SP	insign	L	F	H36	2	100
193	1968-1977	18-20	insign	36	insign	83000000-88000000	C	SP	insign	L	F	H36	2	100
194	1967-1970	18-20	0	36	insign	insign	U	insign	insign	Y	F	H36	2	100
195	1970-1974	18-22	insign	36	insign	insign	U	SP	insign	I	F	H36	2	100
196	1972-1977	19-21	0	33	insign	insign	C	SP	insign	Y	F	J33	2	100
197	1976-1982	16	insign	36	insign	insign	C	CP	insign	J	C1	CJ36	2	100
198	1975-1976	32	insign	36	insign	insign	U	SP	insign	Y	F	insign	2	100
199	1975-1976	3 to 5	30	36	insign	insign	C	SP	insign	Y	F	J36	2	100
200	1968-1971	0	19-21	36	insign	insign	C	SP	insign	insign	F	H36	2	100
201	1979-1980	0	22-25	36	insign	insign	U	insign	insign	insign	F	H36	2	100
202	insign	21-30	insign	36	insign	insign	C	CP	C	Q	C1	CJ36	24	92.30769231
203	1960-1970	22-30	1 to 5	36	insign	insign	U	SP	insign	insign	F	J36	7	87.5
204	1995-2001	22-28	insign	38	insign	insign	C	CP	S	E	C2	CB38	6	85.71428571
205	1968-1972	17-18	insign	36	insign	insign	C	SP	insign	insign	F	insign	6	85.71428571
206	1993-1999	12 to 26	0	36	insign	insign	C	CP	C	B	F	H36	5	83.33333333
207	1993-2000	17-27	0-4	36	insign	insign	C	CP	A	P	C1	CH36	5	83.33333333
208	insign	17-28	0-7	36	insign	insign	C	CP	A	P	C1	CJ36	16	76.19047619
209	1991-1995	21-31	0	36	insign	insign	C	CP	A	E	C2	CH36	6	75
210	1975	16-24	insign	36	insign	insign	C	CP	insign	P	C1	CJ36	3	75
211	insign	insign	insign	36	insign	insign	C	CP	C	insign	F	H36	10	71.42857143
212	1988-1996	15	0	36	insign	insign	C	CP	insign	E	C2	CH36	6	66.66666667
213	1965-1975	22-30	insign	36	insign	insign	U	SP	insign	R	F	J36	6	54.54545455
214	insign	12 to 17	0	36	insign	insign	C	CP	insign	P	C1	CJ36	24	52.17391304

For the full analysis, as in the illustrative example of Chapter V, the optimum inspection schedule is highly dependent on the number of wheels within the population that have a particular critical feature combination. Table 15 shows the number of wheels in the wheel population with the specific critical feature combination corresponding to the critical feature combination number in Table 14.

The reliability analysis and optimization of inspection intervals follow the same procedures discussed in Chapters IV and V. For this application, as discussed in the illustrative example, the goal was to keep the reliability of wheels in each critical feature combination at or above the minimum allowable reliability level of 0.8333.

Table 15. Number of wheels in North American population for each critical feature combination

critical feature combination number	approximate number of wheels in North America with given critical feature combination	critical feature combination number	approximate number of wheels in North America with given critical feature combination	critical feature combination number	approximate number of wheels in North America with given critical feature combination
1	3615	73	339	145	226
2	2373	74	339	146	226
3	1469	75	339	147	226
4	1356	76	339	148	226
5	1243	77	339	149	226
6	1130	78	339	150	226
7	904	79	339	151	226
8	904	80	339	152	226
9	904	81	339	153	226
10	904	82	339	154	226
11	791	83	339	155	226
12	791	84	339	156	226
13	791	85	339	157	226
14	791	86	339	158	226
15	791	87	339	159	226
16	791	88	339	160	226
17	678	89	339	161	226
18	678	90	339	162	226
19	678	91	339	163	226
20	678	92	339	164	226
21	678	93	339	165	226
22	565	94	339	166	226
23	565	95	339	167	226
24	565	96	339	168	226
25	565	97	226	169	226
26	565	98	226	170	226
27	565	99	226	171	226
28	565	100	226	172	226
29	565	101	226	173	226
30	565	102	226	174	226
31	565	103	226	175	226
32	565	104	226	176	226
33	565	105	226	177	226
34	565	106	226	178	226
35	565	107	226	179	226
36	452	108	226	180	226
37	452	109	226	181	226
38	452	110	226	182	226
39	452	111	226	183	226
40	452	112	226	184	226
41	452	113	226	185	226
42	452	114	226	186	226
43	452	115	226	187	226
44	452	116	226	188	226
45	452	117	226	189	226
46	452	118	226	190	226
47	452	119	226	191	226
48	452	120	226	192	226
49	452	121	226	193	226
50	452	122	226	194	226
51	452	123	226	195	226
52	452	124	226	196	226
53	452	125	226	197	226
54	452	126	226	198	226
55	452	127	226	199	226
56	452	128	226	200	226
57	452	129	226	201	226
58	452	130	226	202	2937
59	452	131	226	203	904
60	452	132	226	204	791
61	452	133	226	205	791
62	339	134	226	206	678
63	339	135	226	207	678
64	339	136	226	208	2373
65	339	137	226	209	904
66	339	138	226	210	452
67	339	139	226	211	1582
68	339	140	226	212	1017
69	339	141	226	213	1243
70	339	142	226	214	5197
71	339	143	226		
72	339	144	226		

Table 16. Inspection interval and cost per year for each critical feature combination

critical feature combination number	inspection time at minimum reliability level for given critical feature combination (days)	cost per year to inspect all wheels at minimum reliability level for given critical feature combination (dollars)	critical feature combination number	inspection time at minimum reliability level for given critical feature combination (days)	cost per year to inspect all wheels at minimum reliability level for given critical feature combination (dollars)	critical feature combination number	inspection time at minimum reliability level for given critical feature combination (days)	cost per year to inspect all wheels at minimum reliability level for given critical feature combination (dollars)
1	57	9455	73	20	2600	145	21	1609
2	97	3655	74	127	400	146	21	1609
3	17	12853	75	19	2704	147	21	1609
4	25	8264	76	19	2704	148	21	1609
5	27	6879	77	21	2394	149	21	1609
6	22	7766	78	11	4493	150	35	970
7	57	2364	79	21	2414	151	98	345
8	31	4423	80	21	2414	152	204	166
9	105	1283	81	64	791	153	125	271
10	123	1095	82	110	461	154	40	843
11	69	1711	83	60	843	155	40	843
12	105	1122	84	32	1563	156	131	258
13	37	3194	85	98	518	157	131	258
14	104	1133	86	39	1306	158	131	258
15	16	7601	87	106	477	159	131	258
16	20	6066	88	68	747	160	98	346
17	101	1005	89	39	1317	161	98	346
18	83	1221	90	204	248	162	112	302
19	111	916	91	60	841	163	110	308
20	127	799	92	53	952	164	82	411
21	25	4018	93	25	2009	165	98	344
22	97	870	94	46	1108	166	41	834
23	105	802	95	42	1210	167	41	834
24	37	2282	96	discard	NA	168	31	1108
25	104	809	97	57	591	169	31	1090
26	22	3883	98	97	348	170	35	969
27	111	763	99	97	348	171	35	969
28	60	1399	100	101	335	172	68	495
29	60	1399	101	101	335	173	113	298
30	60	1399	102	101	335	174	113	298
31	20	4311	103	101	335	175	113	298
32	202	418	104	101	335	176	34	994
33	58	1451	105	101	335	177	70	485
34	35	2424	106	101	335	178	59	576
35	22	3794	107	101	335	179	37	924
36	97	696	108	106	319	180	61	558
37	101	670	109	69	489	181	41	821
38	101	670	110	105	321	182	25	1339
39	101	670	111	105	321	183	25	1339
40	101	670	112	104	324	184	22	1518
41	101	670	113	104	324	185	22	1518
42	17	3955	114	104	324	186	22	1518
43	104	648	115	16	2172	187	46	739
44	22	3107	116	22	1553	188	46	739
45	20	3466	117	22	1553	189	46	739
46	20	3466	118	111	305	190	46	739
47	20	3466	119	111	305	191	46	739
48	20	3466	120	20	1733	192	47	720
49	21	3211	121	20	1733	193	42	807
50	114	596	122	20	1733	194	42	807
51	19	3605	123	20	1733	195	42	807
52	21	3192	124	20	1733	196	41	817
53	40	1689	125	20	1733	197	41	834
54	31	2189	126	60	559	198	47	727
55	69	979	127	60	559	199	discard	NA
56	11	5990	128	60	559	200	discard	NA
57	64	1062	129	60	559	201	discard	NA
58	60	1136	130	60	559	202	101	4357
59	47	1440	131	127	266	203	46	2955
60	47	1440	132	127	266	204	37	3164
61	40	1693	133	20	1724	205	42	2792
62	57	886	134	58	580	206	57	1773
63	97	522	135	19	1802	207	106	958
64	101	503	136	19	1802	208	106	3351
65	101	503	137	37	904	209	114	1181
66	101	503	138	37	904	210	41	1668
67	101	503	139	21	1596	211	57	4136
68	105	481	140	40	845	212	37	4067
69	22	2330	141	11	2995	213	46	4063
70	111	458	142	11	2995	214	25	30801
71	111	458	143	64	531			
72	111	458	144	114	295			

The optimized inspection times and the cost per year to inspect all of the wheels in the critical feature combinations at their optimized inspection times are presented in Table 16. As discussed in Chapter V, even though it is financially feasible, it is practically infeasible to inspect the railroad wheels as often as the optimized inspection intervals concluded. Therefore, the inspection times corresponding to the minimum allowable reliability level for each critical feature combination are given in Table 16. It is important to note that in Table 16, when the inspection interval for a critical feature combination is given as *discard* it means the wheels with that critical feature combination should be removed from service instead of scheduled for their next inspection.

It is important to once again note that some features change over time based on the age and condition of the wheel; for the full analysis these features include the rim thickness and flange thickness. The next inspection interval for wheels will be determined by the rim thickness and flange thickness at the wheels' current inspection time. Also, the railroads will need to inspect all of the wheels that have a specific critical feature combination. If only a portion of the wheels with that specific critical feature combination are inspected then the reliability level of that critical feature combination will fall.

For this full analysis of railroad wheels, the cost per year to inspect all of the wheels with a critical feature combination at their minimum reliability inspection intervals is $359,604. If inspection resources were focused on these 214 critical feature combinations then wheel failures could possibly decrease by approximately 75.7 percent. It is important to note that only 0.72 percent of the samples had a feature combination that was considered critical. This means that if inspection resources were focused on this

critical 0.72 percent of the population then wheel failures could possibly decrease by approximately 63 percent (83.33 percent of the 75.7 percent of wheels identified by the critical feature combinations). Once again, the average failure cost per year for a major railroad company from derailments caused by shattered rim failures is approximately $6,000,000; this means that if 83.33 percent of the 75.7 percent of wheel failures were prevented then the railroad company would save approximately $3,784,849 on wheel failure costs. However, because of the additional inspection costs the actual savings are approximately $3,425,245 per year. This translates to an approximate return on investment (ROI) of 953% when inspecting the wheels at their minimum allowable reliability inspection intervals. In addition to the monetary savings, the critical feature combinations would be more reliable than without the additional ultrasonic inspections, which also means increased safety of the railroad industry.

CHAPTER VII

CONCLUSIONS AND FUTURE WORK

The methodology developed in this dissertation for optimizing the inspection schedules of large populations, has been shown to produce promising results, as demonstrated through applying the methodology to a railroad wheel population. The proposed methodology combines clustering, reliability, and optimization techniques, and is transferable to many industries in which there are large populations of samples that need to be inspected. Not only are there monetary benefits for applying this methodology, but the reliability of the critical samples within the optimized population is also maintained at or above the desired level. In addition, the proposed methodology is able to handle features that remain constant, as well as features that change over time. The optimized inspection intervals can also be optimized further using Bayesian updating to update the reliability estimates based on inspection data.

Currently the proposed methodology optimizes the inspection intervals of the critical feature combinations individually without simultaneously considering the inspection costs of the other critical feature combinations. This means that all of the critical feature combinations are optimized separately and the total cost of inspections is then calculated at the end of the analysis. A methodology that includes a total inspection cost constraint would optimize the total cost of all of the critical feature combinations at once.

The proposed methodology also requires that all of the samples contained within the critical feature combinations be inspected at their calculated inspection intervals in order to maintain the critical feature combinations' desired reliability. However, some industries may not be able to, either for financial or logistical reasons, inspect all of the samples within a critical feature combinations, especially if the sample size is very large. A methodology for selecting a reduced number of samples for inspection within a critical feature combination that could simultaneously satisfy the overall reliability requirement for that critical feature combination needs to be explored in the future.

A methodology that includes the population reliability in the inspection scheduling optimization algorithm could account for changes in the population's reliability when selected critical feature combinations' reliability changes. Being able to relate the reliability of a particular feature combination to the population reliability would add further versatility to the proposed methodology.

For the railroad wheel application specifically, currently only shattered rim failures have been included in the analysis. Therefore, the methodology needs to be extended to other failure modes, such as vertical split rims. In order to increase the precision of the calculations the railroad industry also needs to increase the accuracy of their data collection, e.g., confusing failure types when recording data.

Another important issue to consider in the future is cross-feature interaction and feature correlation. Additional modifications to the methodology can also be made based on individual requirements of each application.

REFERENCES

[1] Ansys, "ANSYS Theory Reference, release 7.0", ANSYS, Inc., 2003.

[2] Association of American Railroads (AAR), "Field Manual of the AAR Interchange Rules", Rule 41, 2005.

[3] Association of American Railroads (AAR), "Manual of Standards and Recommended Practices: Section G-Wheels and Axles", Issue of 1998.

[4] Acar, E., Haftka, R.T., "Reliability based aircraft structural design optimization with uncertainty about probability distributions", *Proceedings of the 6th World Congress on Structural and Multidisciplinary Optimization*, Rio de Janeiro, 2005.

[5] Bartley, G.W., "A practical view of wheel tread shelling", *Proceedings of the 9th International Wheelset Congress*, Montreal, 1988.

[6] Berge, S., "Shattered rim fracture research", *Proceedings of the 2000 Brenco Rail Conference*, LaQuinta, California, 2000.

[7] Cheesman, P., Stutz, J., "Bayesian classification (AUTOCLASS): theory and results", *Advances in Knowledge Discovery and Data Mining*, pp. 153-180, 1995.

[8] Chiu, T., Fang, D., Chen, J., Wang, Y., "A robust and scalable clustering algorithm for mixed type attributes in large database environment", In: *Proceedings of 2001 International Conference On Knowledge Discovery and Data Mining*, pp. 263-268, 2001.

[9] Chung, H.Y., Manuel, L., Frank, K.H., "Optimal inspection scheduling with alternative fatigue reliability formulations for steel bridges", *Proceedings of the International Conference on Applications of Statistics and Probability in Civil Engineering*, San Francisco, 2003.

[10] Dieter, G.E., Mechanical Metallurgy, 3rd Edition, McGraw-Hill, New York, 1986.

[11] Ekberg, A., Kabo, E., Andersson, H., "An engineering model for prediction of rolling contact fatigue of railway wheels", *Fatigue & Fracture of Engineering Materials and Structures*, Vol. 25, pp. 899-909, 2002.

[12] Ekberg, A., Marais, J., "Effects of imperfections on fatigue initiation in railway wheels", *IMechE Journal of Rail and Rapid Transit*, Vol. 214, pp. 45-54, 1999.

[13] Ekberg, A., Bjarnehed, H., Lundén, R., "A fatigue life model with application to wheel/rail damage", *Fatigue & Fracture of Engineering Materials and Structures*, Vol. 18 (No. 10), pp. 1189-1199, 1995.

[14] Ekberg, A., Bjarnehed, H., "Rolling contact fatigue of wheel/rail systems – a literature survey", 1995, Retrieved from Chalmers University of Technology, Division of Solid Mechanics Web site: http://www.am.chalmers.se/~anek/research/report_F182.pdf

[15] Ester, M., Kriegel, H.P., Sander, J., Xu, X., "A density-based algorithm for discovering clusters in large spatial databases", In: *Proceedings of the 1996 International Conference of Knowledge Discovery and Data Mining*, Portland, pp. 226-231, 1996.

[16] Estes, A., Frangopol, D., "An optimized lifetime reliability-based inspection program for deteriorating structures", *Proceedings of the 8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability*, University of Notre Dame, 2000.

[17] Faccioli, R., Ferrettii, C., Piva, R., Copello, S., "System fatigue updating for offshore structures", *Proceedings of the 14th International Conference on Offshore Mechanics and Arctic Engineering*, Copenhagen, 1995.

[18] Fatemi, A., Socie, D.F., "Critical plane approach to multiaxial fatigue damage including out-of-phase loading", *Fatigue and Fracture of Engineering Materials & Structures*, Vol. 11 (No. 3), pp. 149-165, 1998.

[19] Fisher, D.H., "Knowledge acquisition via incremental conceptual clustering", *Machine Learning*, Vol. 2, pp. 139-172, 1987.

[20] Fisher, R.A., Statistical Methods for Research Workers, 13th Edition, Oliver and Boyd, Edinburgh and London, 1963.

[21] Galliera, G., "Fatigue behaviour of railway wheels affected by sub-surface defects in the tread. Control methods and manufacturing process", In: *Proceedings of the 11thInternational Wheelset Congress*, Paris, pp. 69-76, 1995.

[22] Ganti, V., Gehrke, J., Ramakrishnan, R., "CACTUS - Clustering Categorical Data Using Summaries", In: *Proceedings of the 1999 International Conference of Knowledge Discovery and Data Mining*, San Diego, pp. 73-83, 1999.

[23] Garbatov, Y., Soares, C.G., "Cost and reliability based strategies for fatigue maintenance planning of floating structures", *Reliability Engineering & System Safety*, Vol. 73 (Issue 3), pp. 293-301, 2001.

[24] Giammarise, A.W., Gilmore, R.S., "Wheel quality: a North American locomotive builder's perspective", 2001, Retrieved from the GE Research & Development Center (CRD140).

[25] Gibson, D., Kleiberg, J., Raghavan, P., "Clustering categorical data: an approach based on dynamic systems", In: *Proceedings of the 1998 International Conference On Very Large Databases*, New York, pp. 311-323, 1998.

[26] Gimenez, J.G., Sobejano, H., "Theoretical approach to the crack growth and fracture of wheels", In: *Proceedings of the 11th International Wheelset Congress*, Paris, pp. 15-20, 1995.

[27] Gluck, M., Corter, J., "Information, uncertainty, and the utility of categories", In: *Proceedings of the 7th Annual Conference of the Cognitive Science Society*, Irvine, CA, pp. 283-287, 1985.

[28] Goodall, D.W., "A new similarity index based on probability", *Biometrics*, Vol. 22, pp. 882-907, 1966.

[29] Gordon, J., Perlman, A.B., "Estimation of residual stresses in railroad commuter car wheels following manufacture", In: *Proceedings of the 1998 International Mechanical Engineering Congress and Exhibition*, ASME RTD Vol. 15, pp. 13-18, 1998.

[30] Guha, S., Rastogi, R., Shim, K., "ROCK: A robust clustering algorithm for categorical attributes", In: *Proceedings of the 1999 International Conference of Data Engineering*, Sydney, Australia, pp. 512-521, 1999.

[31] Guha, S., Rastogi, R., Shim, K., "CURE: A clustering algorithm for large databases", In: *Proceedings of the ACM SIGMOD International Conference of Management of Data*, Seattle, Washington, pp. 73-84, 1998.

[32] Guins, T., Personal e-mail communication, September 5, 2006.

[33] Haldar, A., Mahadevan, S., Probability, Reliability, and Statistical Methods in Engineering Design, John Wiley and Sons, 2000.

[34] Han, E.-H., Karypis, G., Kumar, V., Mobasher, B., "Clustering based on association rule hypergraphs", In: *Proceedings of the 1997 SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discover*, Tucson, Arizona, 1997.

[35] Hanson, S.J., Bauer, M., "Conceptual clustering, categorization, and polymorphy", *Machine Learning*, Vol. 3, pp. 343-372, 1989.

[36] He, Z., Xu, X., Deng, S., "Squeezer: An efficient algorithm for clustering categorical data", *Journal of Computer Science and Technology*, Vol. 17 (No. 5), pp. 611-625, 2002.

[37] Hirakawa, K., Toyama, K., Yamamoto, M., "Fundamental study of shelling in railway wheel materials", *Proceedings of the 8th International Wheelset Congress*, Madrid, 1985.

[38] Huang, Z., "Extensions to the k-means algorithm for clustering large data sets with categorical values", *Data Mining and Knowledge Discovery*, Vol. 2, pp. 283-304, 1998.

[39] Huang, Z., "A fast clustering algorithm to cluster very large categorical data sets in data mining", In: *Proceedings of the SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discover*, Tech. Report 97-07, 1997.

[40] Jain, A.K., Dubes, R.C., Algorithms for Clustering Data, Prentice Hall, Englewood Cliffs, 1988.

[41] Karypis, G., Han, E.-H., Kumar, V., "CHAMELEON: A hierarchical clustering algorithm using dynamic modeling", *IEEE Computer*, Vol. 32 (No. 8), pp. 68-75, 1999.

[42] Lancaster, H.O., "The combining of probabilities arising from data in discrete distributions", *Biometrika*, Vol. 36, pp. 370-382, 1949.

[43] Li, C., Biswas, G., "Unsupervised Learning with Mixed Numeric and Nominal Data", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 14 (No. 4), pp. 673-690, 2002.

[44] Liu, Y., Liu, L., Stratman, B., Mahadevan, S., "Multiaxial fatigue reliability analysis of railroad wheels", *Reliability Engineering and System Safety*, 2007 (accepted December 11, 2006).

[45] Liu, Y., "Stochastic multiaxial fatigue and fracture modeling", Ph.D. Dissertation, Vanderbilt University, Nashville, TN, March 2006.

[46] Liu, Y., Stratman, B., Mahadevan, S., "Fatigue crack initiation life prediction of railroad wheels", *International Journal of Fatigue*, Vol. 28 (Issue 7), pp. 747-756, 2006.

[47] Liu, Y., Mahadevan, S., "Threshold intensity factor and crack growth rate prediction under mixed-mode loading", *Engineering Fracture Mechanics*, Vol. 74, pp. 332-345, 2007.

[48] Liu, Y., Mahadevan, S., "Multiaxial high-cycle fatigue criterion and life prediction for metals", *International Journal of Fatigue*, Vol. 7 (Issue 7), pp. 790-800, 2005.

[49] Lonsdale, C., Pilch, J., Dedmon, S., "Stress effects of wheel impact loads", *Proceedings of the 14th International Wheelset Congress*, Orlando, FL, pp. 6-12, 2004.

[50] Lundén, R., "Cracks in railway wheels under rolling contact load", In: *Proceedings of the 10th International Wheelset Congress*, Sydney, pp. 163-168, 1992.

[51] MacQueen, J.B., "Some methods for classification and analysis of multivariate observations", In: *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, CA, pp. 281-297, 1967.

[52] Marais, J.J., "Wheel failures on heavy haul freight wheels due to subsurface effects", In: *Proceedings of the 12th International Wheelset Congress*, Qingdao, China, pp. 306-314, 1998.

[53] Marais, J.J., Pistorius, P.G.H., "Terminal fatigue of tires on urban transport service", *Proceedings of the 4th International Conference of Contact Mechanics and Wear of Rail/Wheel Systems (preliminary proceedings)*, Vancouver, 1994.

[54] MathWorks, *MATLAB Release 12*, 2001.

[55] McKusick, K., Thompson, K., "COBWEB/3: A portable implementation", Retrieved from the NASA Ames Research Center, Technical Report FIA-90-6-18-2, 1990.

[56] Moyar, G.J., Stone, D.H., "An analysis of the thermal contributions to railway wheel shelling", *Wear*, Vol. 144, pp. 117-138, 1991.

[57] Murikami, Y., Usuki, H., "Quantitative evaluation of effects of non-metallic inclusions of fatigue strength of high strength steels. II: Fatigue limit evaluation based on statistics for extreme values of inclusion size, *International Journal of Fatigue*, pp. 299-307, 1989.

[58] Mutton, P.J., Epp, C.J., Dudek, J., "Rolling contact fatigue in railway wheels under high axle loads", *Wear*, Vol. 144, pp. 139-152, 1991.

[59] Nagode, M., Fajdiga, M., "On a new method for prediction of the scatter of loading spectra", *International Journal of Fatigue*, Vol. 20 (Issue 4), pp. 271-277, 1998.

[60] Nanopoulos, A., Theodoridis, Y., Manolopoulos, Y., "C2P: Clustering based on closest pairs", In: *Proceedings of the 27th International Conference On Very Large Databases*, Rome, Italy, 2001.

[61] Norton, R., Machine Design - An Integrated Approach, 2nd Edition, Prentice Hall, Upper Saddle River, NJ, 2000.

[62] Onoufriou, T., Fowler, D., Smith, J.K., "Reliability-based optimized inspection planning", *Proceedings of 7th International Conference on Behavior of Offshore Structures*, Boston, 1994.

[63] Rail, "Derailment Report. Report Number: R94W0019 (January 23, 1994)", Retrieved from the Web site:
http://www.tsb.gc.ca/en/reports/rail/1994/r94w0019/r94w0019.asp

[64] Rail, "Derailment Report. Report Number: R00Q0023 (May 22, 2000)", Retrieved from the Web site:
http://www.bst.gc.ca/en/reports/rail/2000/r00q0023/r00q0023.asp

[65] Reich, Y., "Building and improving design systems: A machine learning approach", Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA, 1991.

[66] Reich, Y., Fenves, S., "The formation and use of abstract concepts in design", In: Fisher, D., Pazzani, M., Langley, P. (Eds.), Concept Formation: Knowledge and Experience in Unsupervised Learning, Morgan Kaufmann Publishers, Inc., San Francisco, pp. 323-353, 1991.

[67] "Rules Respecting Track Safety (December 18, 1997)", Retrieved from the Web site:
http://www.railcan.ca/documents/rules/en/TC_E04_2_Eng.pdf

[68] Sheikholeslami, G., Chatterjee, S., Zhang, A., "WaveCluster: A multi-resolution clustering approach for very large spatial databases", In: *Proceedings of the 1998 International Conference On Very Large Databases*, New York, NY, pp. 428-439, 1998.

[69] Shetty, N.K., Gierlinski, J.T., Smith, J.K., Stahl, B., "Structural system reliability considerations in fatigue inspection planning", *Proceedings of the 8th International Conference on the Behavior of Offshore Structures*, Delft, 1997.

[70] Stone, D.H., Majumder, G., Bowaj, V.S., "Shattered rim wheel defects and the effect of lateral loads and brake heating on their growth", *Proceedings of the ASME International Mechanical Engineering Congress & Exposition*, New Orleans, LA, 2002.

[71] Stone, D.H., Kalay, S.F., Lonsdale, C.P., "Effect of wheel impact loading on shattered rims", *Proceedings of the 8th International Wheelset Congress*, Rome, 2001.

[72] Stone, D.H., Dahlman, G.E., "The effect of discontinuity size on the initiation of shattered rim defects", *Rail Transportation*, Vol. 19, pp. 7-14, 2000.

[73] Stone, D.H., "Wheel shattered rims - an interpretive review", *IMechE Seminar Publication*, London, England, pp. 75-84, 2000.

[74] Stone, D.H., Moyar, G.J., Guins, T.S., "An interpretive review of railway wheel spalling and shelling", *Rail Transportation*, Vol. 5, pp. 97-103, 1992.

[75] Stone, D.H., Moyar, G.J., "Wheel shelling and spalling - an interpretive review", *Rail Transportation*, pp.19-31, 1989.

[76] Snyder, T., Beck, R., Personal conversations, July 2005.

[77] Stratman, B., Liu, Y., Mahadevan, S., "Structural health monitoring of railroad wheels using wheel impact load detectors", *Journal of Failure Analysis and Prevention*, accepted June 3, 2007.

[78] Tournay, H.M., Mulder, J.M., "The transition from the wear to the stress regime", *Wear*, Vol. 191, pp.107-112, 1996.

[79] Union Pacific Railroad, Information collected from Union Pacific Railroad's laboratory database, Retrieved by Brant Stratman on May 10, 2004 thru August 13, 2004, Omaha, NE, 2004.

[80] Wang, K., Xu, C., Liu, B., "Clustering transactions using large items", In: *Proceedings of the 8th ACM International Conference on Information and Knowledge Management*, Kansas City, MI, pp. 483-490, 1999.

[81] Xing, L., Yingzhi, Z., Jizhong, L., "Study on rim fatigue crack and prevention", In: *Proceedings of the 12th International Wheelset Congress*, Quingdao, China, pp. 53-58, 1998.

[82] Zhang, R., Mahadevan, S., "Reliability-based reassessment of corrosion fatigue life", *Structural Safety*, Vol. 23 (Issue 1), pp. 77-91, 2001.

[83] Zhang, T., Ramakishnan, R., Livny, M., "BIRTH: An efficient data clustering method for very large databases", In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Tucson, AR, pp. 103-114, 1996.

[84] Zhang, Y., Fu, A.W., Cai, C.H., Heng, P.-A., "Clustering categorical data", *Proceedings of the 2000 IEEE International Conference on Data Engineering*, San Diego, CA, 2000.