Risk Based Ultrasonic Rail Test Scheduling
on Burlington Northern Santa Fe

Joseph W. Palese, MCE, P.E.
Zeta-Tech Associates, Inc.
Cherry Hill, NJ USA

Thomas W. Wright
Burlington Northern Santa Fe
Fort Worth, TX USA
ABSTRACT

The Federal Railroad Administration requires railroads operating Class 3 track and above to perform a continuous search for internal rail defects (Track Safety Standards, Part 213,). These regulations provide a minimum guideline for test frequency scheduling, and many railroads test more frequently to minimize the probability of service defects (broken rails) occurring, thus minimizing the risk and cost of derailment.

Rail caused derailments are a major class of derailments costing, on average, in excess of $200,000, some exceeding $1,000,000. Efforts have been made to minimize service failures and risk of derailments, through more often and more effective ultrasonic testing.

In 1997 Burlington Northern Santa Fe instituted a risk-based system for scheduling ultrasonic rail inspections for the 1998 test program. A risk factor is developed (allowable service defects per mile per year) based on the characteristics of the track, including passenger carrying, hazardous materials, dark territory, etc for segments of railroad to be analyzed. The methodology considers historic defect populations (service and detected), number of tests previously performed, and amounts of traffic on the segment. The rate at which rail fatigue defects form and grow to service failures, and the historical characteristics of the study segment provide the basis for the risk-based system, and a frequency of required testing for the next year is determined.
The following paper presents an introduction to the risk-based methodology and the results of its practical application on the Burlington Northern Santa Fe. The success of the application is shown in measured reductions in service failures and mainline derailments.

Key Words: Risk-Based, Service Failure, Derailment, Ultrasonic Inspection, Test Frequency
INTRODUCTION

Rail caused derailments are a major class of derailments in North America with more than 550 rail related derailments occurring on Class I railroads over the past three years, with an average reported cost of over $200,000 [1]. Further, the average cost of a derailment on mainline track for class I railroads was $440,000, for the past three years [2], with a total of over 220 derailments. These costs, which are the reported damage to the track and equipment, do not include costs associated with loss of lading and damage and delay costs due to lost track availability. These additional costs can double the actual derailment cost to the railroad. Ultrasonic rail testing is the primary method used to locate rail defects before they manifest themselves into an accident.

Under the current FRA regulations, railroads operating Class 3 track and above are required to perform a continuous search for internal rail defects [3]. The majority of the railroads test their track more frequently than prescribed by the FRA. This inspection is provided primarily through ultrasonic testing (while some railroads supplement this with magnetic testing), and can be performed on-track at speeds of up to 20 MPH. Given the ever increasing restrictions on track time availability, it is becoming considerably more important to optimize scheduling requirements based on this limited availability of track time, and a frequency of testing consistent with maintaining a high level of safety.

As rail accumulates tonnage, it tends to develop more internal fatigue defects, based on various factors such as metallurgy of the rail, axle loading, support conditions, etc. As defects occur more frequently, it becomes important to test more frequently in order to insure that internal
defects can be located and replaced before they have the opportunity to propagate to failure, and possibly result in a derailment.

Many railroads currently use rules of thumb for scheduling ultrasonic testing, which include age of rail in cumulative MGT, annual traffic density, fixed scheduling frequencies, class of track, type of traffic, defect counts, etc. These rules of thumb lend themselves to the implementation of a scheduling methodology that can be based on measurable performance parameters related to a defined level of risk. In this manner, schedules can be determined based on past performance. Such a methodology was developed by US Department of Transportation Center Volpe National Transportation Systems Center [4], and further enhanced by ZETA-TECH Associates, Inc. for use on the BNSF.

This risk-based analysis methodology considers the performance of the rail as it pertains to historical defect occurrence, rail age and usage, and allowable risk for the track. Such a methodology had not been traditionally used, however, the methodology presented here has been successfully applied to the Burlington Northern Santa Fe (BNSF) since 1998 for its mainline trackage.

Taking the theory to a practical application resulted in an effort for defining data flow as well as keeping the practical and logistical nature of testing in mind. In order to achieve this, a computer model (RailTest) was developed for analyzing large amounts of data effectively and efficiently.
The application of this methodology has lead to real measurable benefits for the BNSF, with a significant reduction in service failures and service to detected defect ratio, and a corresponding reduction in rail related mainline derailments.

The following paper discusses the risk-based methodology for determining test frequencies, the development of RailTest, and the system-wide application of this model to BNSF.
DERAILMENT HISTORY

The Federal Railroad Administration Office of Safety makes available FRA reportable accidents in several databases via their web site [2]. These databases were downloaded and used to evaluate the trend in mainline rail defect related derailments for Class I railroads. The accident databases were downloaded for 1997 – 1999. This analysis was performed primarily to evaluate the effects of applying the risk-based methodology on BNSF, which was instituted for the 1998 rail-testing program.

Considering only those derailments that occurred on mainline track as a result of a rail defect showed that, as a whole, Figure 1 shows that the industry has been experienced an overall 4% increase in derailments per billion gross ton mile (BGTM) from 1997 to 1999. In addition, the average cost of a rail related main line derailment has increased significantly (43%).

Figure 2 shows the distribution of derailments by rail defect type along with the average cost of the derailments for that defect type. This figure clearly shows that the most predominant cause of rail related derailments, is the transverse class of defects. The average costs vary from $200,000 to $1,400,000 depending on defect type, with an overall average of the order of $400,000.

The overall significance of Figures 1 and 2 indicate the need for improved inspection practice such that these class of derailments (and significant costs) can be minimized. This is the goal of the risk-based scheduling methodology.
RISK-BASED THEORY

The risk-based test frequency scheduling methodology is used to schedule ultrasonic testing such that a defined level of risk (failure) is held constant, even as rail ages, thus reducing the level of service failures and risk of derailment to an acceptable level. The methodology evaluates risk as allowable service defects per mile per year. Considering this, there are three primary phenomena that effect the occurrence of a service defect:

1) Defect Initiation: how frequent do defects initiate?
2) Defect Growth: how quickly does a defect grow from initiation to failure size?
3) Detection Reliability: how probable is it that a defect of certain size will be found?

The details of much of the early developmental work is discussed in [4] and has been presented here in a modified and summarized form [5]. The basics of the methodology follow and the reader is referred to references [3] and [4] for further details.

Defect Initiation

As tonnage accumulates on a stretch of track, defects will form and grow. The number of defects that can be expected to enter the population within a given year (or at some time interval after the rail is installed, say between tests) can be fairly well predicted using the standard Weibull cumulative probability distribution [6]. Figure 3 shows the linear representation of the Weibull equation (on a log-log scale) for standard freight railroad parameters, and shows the increased probability of a defect occurring in a rail as the age (in cumulative MGT) of the rail increases.
Defect Growth

Once a defect is formed, it will continue to grow in size with passing traffic. The characteristics of this growth are defined by classical fracture mechanics, and represent decades of research. The key in risk based rail testing is to find the defect between the time it grows to detectable size (minimum detection threshold) and the time when actual failure is eminent (maximum allowable defect size). This interval is of the order of 10 to 50 MGT, depending on a number of factors including curvature, rail section, track modulus, vehicle dynamics, wheel/rail contact, axle load, temperature differential from neutral as well as other residual stress components, and location of the defect.

Figure 4 shows a defect growth model considering a defect that is first detectable at a size of approximately 10% HA (rail head area). This model shows that a defect size of 80% HA (where failure is eminent) will be reached in 40 MGT from first detectable size.

The 40 MGT interval is defined as the safe defect growth life, or the maximum amount of traffic that will propagate a defect to critical size, and thus that should be permitted over the defect between tests.
Detection Reliability

Since the ability to reliably detect a defect varies with its size, particularly for small defects, there is a “probability” of finding any given defect based on its size. Figure 5 shows a standard detection performance model, along with a model that takes into account the effects of improved inspection equipment (more reliable). This figure clearly shows that as defects grow in size they are much more likely to be found during an ultrasonic inspection.

It can also be seen from this figure that an improved reliability equation offers a higher probability of finding defects of smaller sizes. In fact the overall detection reliability for the improved equipment curve is approximately 90.7% as compared to 81.5% for the base detection reliability curve.

Net Detection Efficiency

Knowing that defects will initiate and propagate at a given rate, and that not every defect will be found during a given test, it can be seen that during the next test there is a greater likelihood of finding a missed defect (as long as it has not propagated to failure). In other words, the more often you test, the more likely you are to find a defect at some point in its growth cycle. This allows for the definition of the net detection efficiency, or the weighted average probability of finding a given defect over its growth cycle during multiple tests. The net detection efficiency leads to the development of a scheduling guideline.
**Scheduling Guideline**

Based on the above equations, the scheduling guideline can be determined by defining a set of easily and normally measured parameters, specifically the ratio of service defects (failures) to detected defects (located by a test car prior to failure). This ratio can be related to the net detection efficiency as a function of the service to detected defect ratio \((S/D)\) as follows:

\[
\frac{S}{D} = 0.014(\Delta N - 10)
\]

where \(\Delta N\) is the test frequency in MGT. The derivation of this equation is covered in [4]. The above approximation is determined graphically based on several iterations of test frequency for standard freight railroad parameters, the Weibull equation, the defect growth model and the detection reliability model.

An idealized scheduling curve for freight traffic can be determined as shown in Figure 6, for a risk factor of 0.1 service failures per mile between tests.

This figure provides an “idealized” scheduling interval based on cumulative tonnage and expected defect occurrence. Since this is an idealized scheduling curve it does not allow for actual experience, *i.e.* actual failure and inspection history. Rather it provides a “first-cut” scheduling interval. Note that the curve is cut off at 40 MGT, as this is the defined maximum level of tonnage (conservative) for a defect to grow to failure. The trend towards decreased test interval with rail age is clearly seen as traffic accumulates.
The above equation can be used to develop a master scheduling curve for any defined level of risk. Considering the equations to be a function of total defect rate, in defects per mile per test, Figure 6 can be transformed into a master scheduling curve as shown in Figure 7.

Figure 7 represents a master scheduling curve based on total actual defect rate, as opposed to predicted number of defects. This master scheduling curve can be used to identify the next test interval given the number of defects per mile from the most recent test. As an example (for standard freight traffic and a risk factor of 0.1 service failures per mile per test, used to generate Figures 6 and 7), if 1.0 defect per mile (service and detected) occurred since the previous test, then the next test interval would be set at as 19 MGT.

However, this master scheduling curve does not take into account the effects of defect occurrence fluctuations that happen in practice. These fluctuations often manifest themselves in the service defect behavior or alternatively detected defect variations (e.g. winter versus summer). In order to handle these fluctuations, an adjustment procedure can be implemented.

Considering the service defect rate from the most recent test results in a reduced interval as follows:

\[ \Delta N^* = \Delta N - \frac{S\Delta N - 0.1}{0.014} \]

where \( \Delta N^* \) is the reduced interval when the service defect rate (\( S\Delta N \) service defects/mile from the most recent test) is greater than the risk (0.1 service defects/mile/test in this equation). This
reduced interval can be used directly, or to avoid “over-control” of the testing, the average of this value with the master scheduling curve value can be used as the test interval.

The equations presented above provide a mechanism for scheduling the next test given actual defect occurrence from the previous test. Annualizing information (risk, defects, test frequencies, etc.) provides a mechanism for determining annual or semi-annual test frequencies for territories for which multiple tests occur, which is the case on most mainline tracks. This also allows for the incorporation of seasonal and other time related effects.

**RAILTEST**

In order to determine test frequency requirements for large amounts of trackage and data, a user friendly computer model *RailTest* was developed to analyze large amounts of data effectively and efficiently. The model takes as input the following data on a per segment basis:

- Location
- Length
- Last Test Interval
- Risk Factor (Segment Specific)
- Annual Tonnage

The model, working together with a relational database of the key input information, locates defects on the segment to be analyzed, discriminating on defect type as necessary (Note that in scheduling rail tests, certain classes of defects may be excluded from the defect counts, e.g. engine burn fractures), and performs the comprehensive scheduling analysis for that segment.
Great effort was taken to perform several levels of data error checking and provide output that can be used in a practical manner. In addition, parametric analysis capabilities were built in to provide the user with the ability to analyze track segments with distinctly different operating characteristics in an appropriate manner.

Figure 8 shows an example of the RailTest model with an example analysis.

The screenshot shown in Figure 8 is an example of the RailTest calculator portion of the model that allows for easy parametric analysis of a single segment. As an example, for a 60 mile segment with 50 MGT per year that was tested twice last year and had 19 service failures and 20 detected defects, a revised test interval of 20 MGT (2.5 tests per year) results.

The effects of changing the primary input variables that effect test frequency requirements are shown in Figure 9, particularly risk, annual MGT and service defects. These figures are generated following the above example, varying only the key parameter under consideration. As can be seen from Figure 9a, as the allowable risk decreases (reduced acceptable probability of a derailment), the number of test required increases as expected. Likewise, with increasing annual MGT, (see Figure 9b) the required number of test also increases. Finally, as the number of service defects increases, (see Figure 9c) the number of annual tests required increases as expected.
Note that the model handles annualized data, including risk, defect rates, tests, etc. The production part of the model allows for several input files and generates an output file of track segments with predicted test frequencies on a segment by segment basis. An example of such an output is shown below in Table 1.

**APPLICATION ON BNSF**

The Burlington Northern Santa Fe operates 22 divisions with over 25,500 mile of track, and over 469 Billion Gross Ton Miles of traffic. Figure 10 shows the BNSF system, which illustrates the daunting task of scheduling ultrasonic testing.

The first step in applying the model on BNSF was to develop the track (inspection) segments to be analyzed. This process had to take into account such things as average age of rail, type and amount of traffic, railroad operations, and test car logistical constraints. Several algorithms were developed to identify logical breakpoints for segments, which were then evaluated and modified by the BNSF. The result was a definition of hundreds of segments averaging 70 miles in length.

An annual risk factor was developed for each segment based on the type of track and traffic of that segment. A value normally used by the industry for risk has been 0.1 allowable service failures per mile per year [4]. The base risk factor was defined as 0.09 allowable service failures per mile per year. This factor was then reduced on a segment by segment basis, based on the conditions described below, allowing for an annual risk factor to be applied to the segment for the year.
Thus, a double track, signaled line that carries passengers and is defined as a key route would have a risk factor = 0.09 – 0.02 –0.01 = 0.06 allowable service failures per mile per year. In this manner risk factors were built up for each segment and were defined for the entire segment.

In addition, the defect types to be analyzed as part of the analysis were defined, and the RailTest model structured to allow for definition of the appropriate defect types for analysis (normally associated with fatigue). Table 2 presents the makeup of defects analyzed on BNSF.

Using full system data from BNSF, the RailTest model was run and a segment by segment report of test requirements was developed (similar to that in Table 1). Analyses were run for annual and semi-annual time frames to allow for scheduling during seasonal fluctuations, which may result in significantly different test frequencies, depending on the location of the segment.

The RailTest results were then evaluated by the BNSF and compared to previous frequencies. A final test frequency was decided and test car schedules were made up taking into account practical and logistical constraints.
RESULTS AND BENEFITS

With the vast amounts of track to be scheduled for ultrasonic testing, the task of determining appropriate frequencies and scheduling test cars is enormous. Traditionally this task has been quite subjective and based on various rules of thumb. With the introduction of the RailTest model, an objective scientific method was established for scheduling ultrasonic tests on an allowable risk basis. This increased the level of comfort of the managers of rail testing due to the scientific nature of the analysis.

In addition, the use of such an automated tool allowed for the analysis of large amounts of data and provided test requirements for thousands of miles of railroad in hours. The ever changing conditions of the railroad, including rail age and operation, were handled easily, and the consequences of these changes were evaluated quickly.

Multiple analyses could be performed each year to account for seasonality issues, so that test frequencies could be modified quickly and easily. This allowed for an optimization of test resources and allowed for the placement of test cars at the right locations at the right times.

With the implementation of this improved test-scheduling program, BNSF was able to realize several benefits in both reduced service defect occurrence and reduced derailments.

The first benefit, reduction in defect rates, is illustrated in Figure 11, on a relative basis. This figure shows that since 1997, one year before the system was introduced, both the service defect rate (or risk), and the service to detected defect ratio have decreased significantly, of the order of
28%. This shows that while the age of the rail is getting older on average (some new rail is being installed) and more defects are occurring, more are being found as opposed to being allowed to propagate to failure. In addition, BNSF has been able to realize their defined level of risk through the implementation of this program.

In addition to the significant reductions in defect rates, Figure 12 shows a reduction in rail caused derailments since the program was implemented was realized. In fact, a 33% reduction in derailments was achieved since 1997. This is in light of the fact that a slight increase occurred from 1998 to 1999. This is made even more significant given the fact that the rest of the industry (Figure 12) as a whole has been experiencing a steady increase in rail caused derailments (16%). Note that this is an average of all of the remaining Class I railroads, and does not represent any given railroad.
CONCLUSIONS

The ever-increasing need for safety demanded improvements in rail testing performance and the development of an improved method of ultrasonic test scheduling. The result was the development of an easy to use computer model (*RailTest*) for determining test scheduling requirements, based on a defined level of risk. Risk for this model is annualized and is defined as the allowable service failures per mile per year for a given track segment.

This methodology was successfully implemented on the Burlington Northern Santa Fe for scheduling rail tests on their system. A segment by segment schedule of test frequency requirements was developed for the entire system, which was then used to plan the test program and assign equipment accordingly.

The resulting methodology and model allowed for defining test frequency requirements easily and objectively in an ever changing railroad environment of more trains, more tons, and less track time. With an aging infrastructure, such objective and engineering based tools are becoming increasingly important.

In the case of BNSF, the methodology was applied such that they were able to achieve a significant reduction in service defects and service to detected ratio, which means that they are finding more defects before they become failures. In addition, a significant reduction in mainline rail caused derailments was observed.
REFERENCES


Figure 1. Mainline Rail Related Derailments.
Figure 2. Rail Defect Related Derailments.
Figure 3. Weibull Defect Probability.

Weibull Plot

Cumulative MGT

Probability of a Defect

$\alpha = 3.0$, $\beta = 2000$
Figure 4. Defect Growth Model.
Figure 5. Detection Reliability Models.
Figure 6. Idealized Scheduling Curve.
Figure 7. Master Scheduling Curve.
Figure 8. *RailTest.*
Figure 9. Sensitivity of Test Requirements to Primary Variables.

(a) Risk.

(b) Annual MGT.

(c) Service Defects.
Table 1. Sample Output.

<table>
<thead>
<tr>
<th>ID</th>
<th>Div</th>
<th>Sub Div</th>
<th>Line Seg</th>
<th>Sta From</th>
<th>MP From</th>
<th>Sta To</th>
<th>MP To</th>
<th>Trk Len</th>
<th>Ann MGT</th>
<th>Tot Def</th>
<th>Tot Rate</th>
<th>Ser Def</th>
<th>Ser Rate</th>
<th>Risk Days</th>
<th>Last Days</th>
<th>Next Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>j</td>
<td>Sta a</td>
<td>100.00</td>
<td>Sta b</td>
<td>188.64</td>
<td>S</td>
<td>88.6</td>
<td>71.7</td>
<td>92</td>
<td>2.48</td>
<td>2</td>
<td>0.05</td>
<td>0.06</td>
<td>65.3</td>
<td>63.0</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>k</td>
<td>Sta c</td>
<td>200.00</td>
<td>Sta d</td>
<td>203.08</td>
<td>1</td>
<td>3.1</td>
<td>38.0</td>
<td>9</td>
<td>6.97</td>
<td>0</td>
<td>0.00</td>
<td>0.09</td>
<td>65.3</td>
<td>85.0</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>l</td>
<td>Sta e</td>
<td>110.00</td>
<td>Sta f</td>
<td>215.80</td>
<td>S</td>
<td>105.8</td>
<td>56.3</td>
<td>39</td>
<td>0.88</td>
<td>7</td>
<td>0.16</td>
<td>0.08</td>
<td>45.0</td>
<td>41.0</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>m</td>
<td>Sta g</td>
<td>120.00</td>
<td>Sta h</td>
<td>206.80</td>
<td>S</td>
<td>86.8</td>
<td>57.0</td>
<td>42</td>
<td>1.15</td>
<td>2</td>
<td>0.05</td>
<td>0.08</td>
<td>47.0</td>
<td>73.0</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>n</td>
<td>Sta i</td>
<td>0.00</td>
<td>Sta j</td>
<td>42.70</td>
<td>S</td>
<td>42.7</td>
<td>20.9</td>
<td>14</td>
<td>0.78</td>
<td>3</td>
<td>0.17</td>
<td>0.08</td>
<td>74.3</td>
<td>52.0</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>o</td>
<td>Sta 1</td>
<td>81.00</td>
<td>Sta 2</td>
<td>111.90</td>
<td>S</td>
<td>30.9</td>
<td>16.1</td>
<td>7</td>
<td>0.54</td>
<td>2</td>
<td>0.15</td>
<td>0.08</td>
<td>108.5</td>
<td>73.0</td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>p</td>
<td>Sta 3</td>
<td>21.00</td>
<td>Sta 4</td>
<td>97.94</td>
<td>S</td>
<td>76.9</td>
<td>19.2</td>
<td>13</td>
<td>0.40</td>
<td>2</td>
<td>0.06</td>
<td>0.06</td>
<td>85.0</td>
<td>84.0</td>
</tr>
<tr>
<td>8</td>
<td>C</td>
<td>q</td>
<td>Sta B</td>
<td>5.00</td>
<td>Sta C</td>
<td>26.00</td>
<td>S</td>
<td>21.0</td>
<td>49.1</td>
<td>10</td>
<td>1.14</td>
<td>2</td>
<td>0.23</td>
<td>0.07</td>
<td>40.0</td>
<td>31.0</td>
</tr>
<tr>
<td>9</td>
<td>C</td>
<td>r</td>
<td>Sta D</td>
<td>0.00</td>
<td>Sta E</td>
<td>31.10</td>
<td>1</td>
<td>31.1</td>
<td>40.1</td>
<td>3</td>
<td>0.23</td>
<td>3</td>
<td>0.23</td>
<td>0.08</td>
<td>40.3</td>
<td>29.0</td>
</tr>
<tr>
<td>10</td>
<td>C</td>
<td>r</td>
<td>Sta D</td>
<td>0.00</td>
<td>Sta E</td>
<td>31.10</td>
<td>2</td>
<td>31.1</td>
<td>74.9</td>
<td>4</td>
<td>0.31</td>
<td>0</td>
<td>0.00</td>
<td>0.08</td>
<td>40.5</td>
<td>106.0</td>
</tr>
</tbody>
</table>
Figure 10. BNSF System. [7]
Table 2. Defect Types Analyzed.

<table>
<thead>
<tr>
<th>Included</th>
<th>Not Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detail Fracture</td>
<td>Engine Burn Fracture</td>
</tr>
<tr>
<td>Transverse Defect</td>
<td>Crushed Head</td>
</tr>
<tr>
<td>Horizontal Split Head</td>
<td>Broken Base</td>
</tr>
<tr>
<td>Vertical Split Head</td>
<td>Rail End Weld Fracture</td>
</tr>
<tr>
<td>Piped Rail</td>
<td>Welded Engine Burn Fracture</td>
</tr>
<tr>
<td>Bolt Hole Break</td>
<td></td>
</tr>
<tr>
<td>Head and Web Separation</td>
<td></td>
</tr>
<tr>
<td>Weld</td>
<td></td>
</tr>
<tr>
<td>Plain Break</td>
<td></td>
</tr>
</tbody>
</table>
Figure 11. Changes in Defect Rates.
Figure 12. Mainline Rail Caused derailments.
LIST OF TABLES and FIGURES

1. FIGURE 1 Mainline Rail Related Derailments - ref on pg 7, para 2
2. FIGURE 2 Rail Defect Related Derailments - ref on pg 7, para 3
3. FIGURE 3 Weibull Defect Probability - ref on pg 8, last para
4. FIGURE 4 Defect Growth Model - ref on pg 9, para 2
5. FIGURE 5 Detection Reliability Models - ref on pg 10, para 1
6. FIGURE 6 Idealized Scheduling Curve - ref on pg 11, para 2
7. FIGURE 7 Master Scheduling Curve - ref on pg 12, para 1
8. FIGURE 8 Rail Test - ref on pg 14, para 4
9. FIGURE 9 Sensitivity of Test Requirements to Primary Variables
   
   Figure 9a  RISK - ref on pg 14, para 4
   Figure 9b  Annual MGT - ref on pg 14, para 4
   Figure 9c  Service Defects - ref on pg 14, para 4
10. TABLE 1 Sample Output - ref on pg 15, para 1
11. FIGURE 10 BNSF System - ref on pg 15, last para
12. TABLE 2 Defect Types Analyzed - ref on pg 16, para 2
13. FIGURE 11 Changes in Defect Rates - ref on pg 17, last para
14. FIGURE 12 Mainline Rail Caused Derailments - ref on pg 18, para 2