Abstract
Research and studies have shown that the risk of broken rail derailments is directly related to the rate of rail defect development and the associated relationship between service defects and detected defects. This paper presents a study on the relationship between rail defects and broken rail derailments, based on defect data from several US railroads. The paper focuses on the relationship between service defects and derailments and examines the effect on improved inspection, either through improved inspection technology or improved inspection scheduling, on reducing rail service defects and associated derailments. The paper also examines the levels of broken rail risk that have been found on freight and passenger systems in North America and Europe. Guidelines for the range of broken rail risk that has been found effective are presented, with a focus
on risk values as a function of traffic type (e.g. passenger, hazmat, conventional freight), speed and other appropriate defining parameters.

**Introduction**

Broken rail derailments remain one of the largest track related derailment causes with an average of 280+ derailments a year over the last 8 years and a total average annual FRA reported cost of over $50 Million. Noting that actual railroad cost is of the order of twice the FRA reported cost; the annual cost to the US railroad industry of broken rail derailments is over $100 Million. Figure 1 presents the annual number of rail caused derailments and associated annual FRA reported broken rail derailment costs for the period 1997 through 2004. As can be seen from this figure, for the US railroad industry as a whole, the number and cost of broken rail derailments has remained constant (and in fact has shown a small increase) over this eight-year time period.

Clearly control of broken rail derailments is a serious concern. Research studies encompassing over two decades have shown a relationship between rail defect occurrence and broken rail derailments [1,2]. This general relationship, which was developed in the mid to late 1980s, is illustrated in Figure 2 and Table 1. This data shows that in the 1980s, a broken rail derailment occurred every 770 defects for a derailment per rail defect rate of 0.0013. This rate was more severe for contact stress related defects such as detail fractures from shells, vertical split heads, etc. where, a broken rail derailment occurred every 526 defects for a derailment per defect rate of 0.0019 [2].

Examination of the breakdown between detected defects (those defects located by non-destructive inspection processes such as ultrasonic testing) and service defects (those
defects located by visual inspection, observation of breaks, etc.) shows that, for this time period, approximately 75% of the total defects were detected defects, while almost 25% were found by other means, i.e. service defects. However, as seen in Figure 3, there is a defined correlation between service defects and derailments, with a broken rail derailment occurring every 200 service defects for a derailment per service defect rate of 0.005. For contact stress related defects, a broken rail derailment occurred every 133 service defects for a derailment per defect rate of 0.0075 [2].

Subsequent research which focused on control of rail service defects by improved inspection efficiencies, showed the potential for identifying and controlling the risk of broken rail derailments by reducing the percentage of service defects and the associated service defect rate [3, 4].

**Recent Experience**

Recent studies on the relationship between derailments and rail defects have confirmed the behavior identified during the 1980s. Examination of data from one major US Class 1 railroad for the period 1995 though 2003 shows again that there is a relationship between total defects and derailments as illustrated in Figure 4. When similar rail testing approaches were compared, the derailment rate was found to be one broken rail derailments for every 826 defects (based on all mainline defects), which corresponded to a derailment per defect rate of 0.0012, which was virtually the same rate as was found in the mid-1980s.

In order to examine the relationship between derailments and rail service defects, which is a function of rail detection effectiveness, it is necessary to remove those defects that are not reliably found by conventional rail testing technologies (e.g. ultrasonic
testing). Table 2 shows a list of “detectable” rail defects, and identifies those defects that are not reliably found through conventional detection means. Note: Those defects are those identified by and used in ZETA-TECH’s RailTest broken rail risk assessment methodology [4, 5, 6]. Figure 5, shows this relationship between derailments and total detectable (RailTest) defects.

The actual percentage of service defects to total defects (the service to total ratio) will vary significantly based on whether “the detectable only” defects shown in Table 2 are considered or whether all defects are considered, even after accounting for the differences in the total derailment count. Thus, while the service to total ratio for all defects remains at the same 25% level identified in the 1980 study, if only detectable defects are considered (and associated derailments linked to detectable defects) this ratio drops to the 15% level, i.e. 85% of the detectable defects were found by detector cars, with 15% found by visual inspection.

As noted previously, the total defects include both detected defects and undetected service defects. Examination of the relationship between derailments and detected defects, as presented in Figure 6, shows a virtually flat line, which suggests that there is little or no significant relationship, which would be the case if the derailment-defect relationship were being driven by the service defects rather than the detected defects. Figures 7 and 8 confirm this, showing that there is in fact a stronger relationship between derailments and service defects than with detected defects. This supports the results of the earlier 1980s studies as well as the follow on service defect driven risk based studies noted above, that the rate of broken rail derailments is directly associated with the number (and rate of) undetected or service defects.
Increased testing has a direct effect on this behavior, as shown in Figure 9. As can be seen in this figure, the service defect rate decreases directly with increased testing. Thus, as expected, the number of derailments likewise decreases with increased testing, as shown in Figure 10.

However, increased testing is expensive and should be judiciously applied. Simply increasing testing “across the board” is not the best approach. Rather, assessment of the rail condition, and the “risk” of broken rails (service defects) and associated broken rail derailments, offers the most efficient and cost effective approach to increasing rail testing. One such approach, referred to as the “Self-Adaptive Scheduling of Rail Tests” was developed by Volpe National Transportation Systems Center of the US Department of Transportation [2]. This approach, which was modified and adapted by ZETA-TECH into it’s RailTest risk based ultrasonic test scheduling model has been applied extensively in the United States and overseas in order to increase the percentage of detected defects and reduce the percentage of broken rails (and thus the associated risk of broken rail derailments).

The objective of this analysis methodology is to schedule ultrasonic testing (UT) so that a defined level of risk (of rail failure or breakage) is held constant, even as rail ages. Risk is defined as the annual rate of service defect occurrence, specifically the number of service defects (rail breaks) per mile per year, which as shown above has a direct relationship to the rate of broken rail derailment occurrence.

The risk analysis methodology addresses several key factors that affect the occurrence of rail breaks, to include defect initiation and growth and the reliability of the non-destructive detection technique used. It then analyzes the probability that an internal
rail defect will “escape” detection and grow to critical size before the next scheduled rail test. It then adjusts the ultrasonic (or other) test schedule to minimize the risk of “missing” such an internal defect, which otherwise is undetectable by visual and manual inspection.

A key parameter in this scheduling process is the risk itself, defined as noted above, as the number of service defects per mile per year. Analysis of defect records on thousand of track segments have shown that many locations of higher than acceptable risk do occur, in some cases significantly higher than “average” and that these locations have a higher probability of failure than do the other lower risk locations.

The track segments that exceed these risk levels are deemed high risk areas and must be addressed by improved UT testing, either through the use of better equipment or through improved test scheduling.

This risk based scheduling methodology has been successfully implemented on several major rail systems in the US and Europe. The results have been dramatic and consistent. A multi-year application of RailTest on BNSF show a well documented reduction in both service defects and in rail related derailments. This is illustrated in Figures 11 though 13, which show the rail testing performance and associated derailment rate for a nine-year period from 1985 through 2003. Noting that the application of risk based scheduling (RailTest) occurred in late 1997 (1998 was the first full year of application), it can be seen that there was a well defined reduction in the percentage of total defects that went undetected (i.e. were found as service defects), shown in Figure 11, as Service/Total ratio with a corresponding reduction in FRA reported broken rail derailments. Likewise, the service defect rate behavior over that same time period shown
in Figures 12 and 13, shows a similar reduction corresponding to a reduction in broken rail risk (service defects per mile per year) and in actual broken rail derailments. Again, note the correlation between derailment rate and service defects.

A recent application on another major US Class 1 railroad shows similar behavior. This is illustrated in Figures 14 and 15, where, following an application of risk based scheduling (RailTest) in 2003, there was a well defined reduction in percentage of service defects (Service/Total ratio) and service defect rate, with a corresponding reduction in broken rail derailments.

Table 3 presents a consolidated set of broken rail and broken rail derailment data using both conventional UT scheduling techniques (“pre-RailTest”) and risk based scheduling (“post-RailTest”). As can be seen from this data, both the total number of derailments and the derailment rate as a function of “detectable” defects decreases by about 1/3 (32 to 33%). Similarly, the number of service defects per mile decreases by over 40% and the ratio of service defects to total defects (percentage of defects non-found by UT testing) decreases by 30% (with a corresponding increase in percentage of defects found by UT detection). However, the ratio of derailments to service defects remains virtually constant, of the order of 0.008 or one derailment every 125 defects, thus again supporting the results shown earlier, that there is a strong relationship between service defects such as broken rails and broken rail derailments. Furthermore, this ratio of derailments to service defects based on current data (Table 3) is similar to the ratio of derailments to service defects observed in the in the mid-1980’s. In particular, the contact stress associated defects (Table 1) correspond significantly to the RailTest defects defined in Table 2.
Broken Rail Risk Criterion
Noting the data presented above taken from several different time periods over a range of US Class 1 railroads, there appears to be a clear and well-defined relationship between the occurrence of broken rails and associated service defects and the occurrence of broken rail derailments. This broken rail “risk” has been defined [3, 4, 5, 6] as service defects (including rail breaks) per mile per year.

Using this concept and the range of data that has been collected, both in the US and overseas [5, 6], it is further possible to introduce the concept of “maximum allowable risk”, that is the maximum allowable rate of broken rails that can be accepted on a given line segment. Clearly this maximum level of allowable risk will vary with such key factors as type of traffic, and in particular the presence of passenger trains or hazardous materials (hazmat traffic).

Analysis of North American rail defect data indicated that the current North American industry average is approximately 0.1 broken rail per mile per year. (Table 3, shows it to be 0.12 service defects/mile/year without using risk based test scheduling and 0.07 service defects per mile if risk based scheduling is used across a broad range of track and traffic conditions. (On the BNSF, it is in the range of 0.06 as shown in Figure 12.)

One US freight railroad [4] defined risk on a segment by segment basis based on specific track and traffic conditions, starting with a system wide maximum allowable level risk of 0.09 service defects/mile/year (“base risk”) and then reducing the base risk for each of the following characteristics by the amount shown.

Passenger Carrying Line* -0.02
Single Track             -0.01
Dark Territory            -0.02
Key Route (railroad defined)  -0.01

*Limited to a maximum of several passenger trains per day

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 can be built up for each segment and defined on a per segment basis.

In contrast to freight lines with little or no passenger traffic, for high-speed (125 mph) passenger lines, risk factors of between 0.001 and 0.01 have been used, based on actual service break history and desired level of risk. In Europe, risk on passenger-carrying lines has been defined as 0.001 for high-speed lines, 0.01 for moderate speed passenger lines, and 0.03 for low-speed lines.

Based on North American and European experience [5,6], a preliminary set of guidelines for setting risk factors is presented as follows:

- 0.09 to 0.10 General freight route (no passenger or hazmat)
- 0.07 to 0.08 Key freight route
- 0.06 to 0.07 Hazmat route (no passenger traffic)
- 0.04 to 0.06 Freight with limited passenger traffic
- 0.01 to 0.03 Low-speed passenger route (less than 90 mph)
- 0.005 to 0.01 Moderate-speed Passenger (90 to 125 mph)
- 0.001 High-speed passenger (125 mph and higher)
The actual values to be used can vary from segment to segment, with a test segment being defined as the approximate distance that can be tested in one day (generally of the order of 80 to 160 km).

**Summary**
Research studies over the last several decades have shown a correlation between rail defects and broken rail derailments, a major class of track caused derailments, and one of the most expensive. Noting that defects are divided into two primary classes, based on how they are found, specifically detected defects and undetected or service defects, the same research studies have shown a real correlation of broken rail derailments with service defects, with approximately 1 derailment per 125 service defects for main line track under modern heavy axle loading (based on those defects that are most readily detectable by conventional ultrasonic testing). Thus, the higher the percentage of detected defects, and thus the lower the percentage of service defects, the lower the risk of derailments. Service defect percentages (service defects/total defects) of 10% or lower offer lower risk of broken rail derailments than service defect percentage of 15% and higher.

Recent use of risk based test scheduling techniques has been found to reduce the risk of derailments, using a service defect based definition of risk, as service defects per mile per year. These techniques, which have been based on over a decade’s worth of rail integrity research, are used to determine optimum rail test intervals, per segments of track, based on a defined level of risk for that track segment. They have been shown to reduce the rate of broken rails (service defects) and associated broken rail derailments by 30% or more.
The use of service defect rate based risk definition also allows for the defining of maximum allowable risk for different classes of freight and passenger track. Thus, while average level of risk for US freight railroads is of the order of 0.1 service defects/mile/year, more appropriate levels of risk are of the order of 0.06 to 0.08 for freight only routes, 0.04 to 0.06 for freight lines with limited passenger traffic, and 0.01 to 0.001 for passenger dominated lines, with the highest speed lines requiring the lowest maximum allowable level of risk.

**Acknowledgements:**

The authors would like to thank Burlington Northern Santa Fe and in particular BNSF: John Gillette, Tom Wright and Rick Boals for their support and assistance.

**References:**


TABLE 1: Rail defect statistics for selected roads (1988)

(grouped by failure mode)

<table>
<thead>
<tr>
<th></th>
<th>Flex Stress</th>
<th>Contact Stress</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Derailments</td>
<td>115</td>
<td>140</td>
<td>255</td>
</tr>
<tr>
<td>Number of Service Failures</td>
<td>32,441</td>
<td>18,554</td>
<td>50,995</td>
</tr>
<tr>
<td>Number of Detected Defects</td>
<td>86,763</td>
<td>56,362</td>
<td>143,125</td>
</tr>
<tr>
<td>Total Number of Defects</td>
<td>119,319</td>
<td>75,056</td>
<td>194,375</td>
</tr>
<tr>
<td>Derailments/Service Defects</td>
<td>0.0035</td>
<td>0.0075</td>
<td>0.0050</td>
</tr>
<tr>
<td>Service/Detected Defects</td>
<td>0.364</td>
<td>0.329</td>
<td>0.350</td>
</tr>
<tr>
<td>Derailments/Total Defects</td>
<td>0.0010</td>
<td>0.0019</td>
<td>0.0013</td>
</tr>
<tr>
<td>Service/Total Defects</td>
<td>0.272</td>
<td>0.247</td>
<td>0.262</td>
</tr>
<tr>
<td>Detected/Total Defects</td>
<td>0.727</td>
<td>0.751</td>
<td>0.736</td>
</tr>
</tbody>
</table>
TABLE 2: “Detectable “ fatigue-related defect types

<table>
<thead>
<tr>
<th>Reliably Detected*</th>
<th>Not Reliably Detected</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 Break 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>

* Defects included in RailTest broken rail risk analysis
TABLE 3: Rail defect* statistics for selected roads

<table>
<thead>
<tr>
<th>Mainline Tracks Only</th>
<th>Pre-RailTest</th>
<th>Post RailTest</th>
<th>Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Derailments</td>
<td>19.4</td>
<td>13.4</td>
<td>32%</td>
</tr>
<tr>
<td>Derailments/mile</td>
<td>0.00085</td>
<td>0.00057</td>
<td>33%</td>
</tr>
<tr>
<td>Service defects/mile/year</td>
<td>0.12</td>
<td>0.071</td>
<td>41%</td>
</tr>
<tr>
<td>Detected defects/mile/year</td>
<td>0.66</td>
<td>0.61</td>
<td>8%</td>
</tr>
<tr>
<td>Total defects/mile/year</td>
<td>0.78</td>
<td>0.68</td>
<td>13%</td>
</tr>
<tr>
<td>Derailments/Service Defects</td>
<td>0.0084</td>
<td>0.0079</td>
<td>6%</td>
</tr>
<tr>
<td>Service/Detected Defects</td>
<td>0.176</td>
<td>0.117</td>
<td>32%</td>
</tr>
<tr>
<td>Derailments/Total Defects</td>
<td>.00121</td>
<td>.00083</td>
<td>33%</td>
</tr>
<tr>
<td>Service/Total Defects</td>
<td>0.15</td>
<td>0.105</td>
<td>30%</td>
</tr>
<tr>
<td>Detected/Total Defects</td>
<td>0.85</td>
<td>0.895</td>
<td></td>
</tr>
</tbody>
</table>

* RailTest defects only
FIGURE 1: Rail Caused Derailments 1997 - 2004

FIGURE 2: Accidents vs. Defects/mile (1988)
FIGURE 3: Derailments vs. Service Defect Rate (1988)


FIGURE 9: Service Defect Rate vs. Test Miles (normalized)

FIGURE 10: Derailments vs. Test Miles (normalized)

FIGURE 14: Derailments vs. Service/Total Defect Ratio US Class 1 RR (2002-2004)

FIGURE 15: Derailments vs. Service Defect Rate, US Class 1 RR (2003-2004)
Tables and Figures

TABLE 1: Rail defect statistics for selected roads (1988) (grouped by failure mode)

TABLE 2: “Detectable“ fatigue-related defect types

TABLE 3: Rail defect* statistics for selected roads

FIGURE 1: Rail Caused Derailments 1997 - 2004

FIGURE 2: Accidents vs. Defects/mile (1988)

FIGURE 3: Derailments vs. Service Defect Rate (1988)


FIGURE 9: Service Defect Rate vs. Test Miles (normalized)

FIGURE 10: Derailments vs. Test Miles (normalized)


FIGURE 14: Derailments vs. Service/Total Defect Ratio US Class 1 RR (2002-2004)

FIGURE 15: Derailments vs. Service Defect Rate, US Class 1 RR (2003-2004)