ABSTRACT

Broken rail is the most common cause for the mainline freight-train derailment in the United States. Rail defect detection and removal is critical for the risk reduction due to broken-rail-caused derailments. The current practice is to periodically inspect rails using non-destructive technologies, particularly ultrasonic inspection. Determining and prioritizing the frequency of rail defect inspection is an important decision in broken rail risk management. A generalized, risk-based mixed integer nonlinear programming (MINLP) model is developed which can optimize segment-specific rail defect inspection frequency to minimize route broken rail risk, especially under limited inspection resources. A numerical example to optimize the inspection frequency is used to illustrate the application of the model. The result analysis states that the optimization approach can lead to a risk reduction of broken rail compared to an empirical heuristic that all segments on the same route are tested at an equal frequency. The optimization algorithm is being implemented into a computer-aided decision support tool called “Rail Risk Optimizer” that can automatically recommend an optimal segment-specific ultrasonic rail defect inspection frequency given risk factors such as rail age and traffic density. The research methodology and the practice-ready optimization tool can aid the railroad industry to mitigate broken rail risk in a cost-efficient manner.

INTRODUCTION

Railway transportation plays an important role in the U.S. economy. While society derives substantial benefits from rail transportation, there are accompanying risks hidden behind that must be managed and reduced. Train accidents cause damage to infrastructure and rolling stock, disrupt services, and may cause casualties and harm the environment [1-3]. Although the U.S. freight-train accident rate has declined by over 80 percent since 1980 [2], accidents still present a major safety concern. For instance, all types of train accidents resulted in about 360 million dollars’ worth of reported damage costs and 853 casualties in 2015 [4].

Derailment is the most frequent freight-train accident on mainlines in the U.S., accounting for approximately 72% of all types of accidents [1]. In light of continual growth in nationwide rail traffic, a further reduction in derailment risk is a high priority for both the railroad industry and the Federal Railroad Administration (FRA) [5]. There are different accident causes that can result in train derailment. The FRA categorizes over 400 accident causes into five major groups: track, equipment, human factors, signal, and miscellaneous. Each of these broad groupings contains respective subgroups with dozens of detailed causes. It is important to note that each accident cause has corresponding accident frequency [6]. Among all freight-train derailment causes, broken rails or welds are the most frequent (Figure 1a), making broken rail the most frequent freight train accident of the four major causes that can result in train derailment [7]. Train derailment causes, broken rails or welds are the most frequent (Figure 1a), making broken rail the most frequent freight train accident of the four major causes that can result in train derailment [7]. Among all freight-train derailment causes, broken rails or welds are the most frequent (Figure 1a), making broken rail the most frequent freight train accident of the four major causes that can result in train derailment [7].
LITERATURE REVIEW AND OBJECTIVES OF THIS STUDY

Literature Review

This section reviews relevant studies in the areas of the occurrence of broken rails, broken rail prediction, and inspection frequency scheduling.

1) Occurrence of broken rails. Several types of defects might occur to the rail, such as longitudinal defects, transverse defects, base defects, and others [18]. Transverse defects related to metal fatigue are one of the most common severe defects leading to rail service failures and train derailments [15, 19]. Studies have found that rail design, rolling stock characteristics, inspection, and maintenance schedules all affect broken rail risk. The mechanism of rail crack formation and growth through theoretical modeling and laboratory testing has been extensively studied in the literature [11, 20-23]. Orringer et al. proposed a fracture mechanics model to estimate the defect growth in rail and predict the accumulative tonnage to grow a defect from a detectable size to a critical size [11]. Orringer et al. described a guide to calculate the inspection frequencies based on the observed defects or broken rail number between rail tests with the assumption of the detection probability model [12].

2) Broken rail prediction. In addition to the engineering analysis, the prior effort also predicted broken rail occurrence using statistical approaches. For example, Shry and Ben-Akiva [24] developed both a survival function and a hazard function to predict the rail condition. Dick [25] evaluated the factors affecting broken rail service failures and derailments using a multivariate analysis of predictor variables. Dick et al. [26] developed a broken rail prediction model to estimate broken rail risk given rail age, rail weight, degree of curvature, speed, and several other factors. Sourget and Riollet [27] developed two models for prediction of broken rails: logistic regression and decision trees. A general model to estimate the total number of broken rail between two successive rail defect inspections is developed by the Volpe National Transportation Systems Center [8, 15]. This model combines the defect formation, growth, detection process which are proposed by Orringer et al. [11,12]. Using the outputs of the Volpe model, Liu et al. [15] developed an exponential model to correlate broken rail rate with inspection frequency, and found that the higher the inspection frequency, the lower the broken rail risk, given all else being equal. In the United Kingdom, Zhao et al. [16, 17] also found that an exponential function can approximately describe the relationship between annual number of broken rails per track-mile and inspection frequency.
3) **Inspection frequency scheduling.** Liu et al. [28] developed an optimal condition-based rail inspection frequency that incorporates the effect of the seasonal variation. Liu et al. [15] proposed an analytical model to optimize the inspection interval based on the risk category (low risk, medium risk, and high risk). Orringer et al. [29] studied a delayed action concept for prioritizing immediate repair of critical defects (those with a large surface area), while delaying repair of non-critical defects for a defined grace period.

**Knowledge Gaps**

While the knowledge of broken rail risk management continues to grow, there are several areas for further research. On the subject of rail defect inspection, the prior research either assumed that all segments are inspected at an equal frequency (empirical heuristic) [12, 30] or the segments whose broken rail risks are within the same risk category are inspected at an equal frequency (group-based inspection strategy) [7, 31]. To our knowledge, there is no published study that explicitly models segment-specific inspection frequency. This presents the most general and complex inspection scenario. For example, consider a 10-segment route. Each segment might be inspected at a minimum of twice per year to a maximum of seven times per year (6 possible frequencies on each segment, ranging from two to seven). Using combinatorial mathematics, we can develop a total of $6^{10}$ (60 million scenarios) across all 10 segments. All the existing and emerging inspection schedules are within these scenarios. Given a large number of possible inspection schedules, which one would lead to the lowest number of broken rails under resource constraints? For instance, instead of inspecting all segments at four times per year, can we prioritize more inspections on certain segments that can minimize the route risk, without requiring additional miles of inspection? This research is developed to address these questions.

**Research Objectives and Scope**

To ensure a deep understanding of rail inspection issues within the content limit, this research focuses its effort on fatigue-related rail defects, including detail fractures, transverse defects, and vertical split head defects. Other types of rail defects or track geometry defects are beyond the scope of this paper, but shall be addressed in a separate detailed study. Also, this paper focuses on defective rails for freight railroads, without considering passenger or transit rails. This research aims to address the following objectives:

- Develop a generalized risk-based optimization methodology that can prioritize the allocation of inspection resources to different track segments with heterogeneous risks.
- Implement the methodology into a computer-aided decision support tool for automatically computing location-specific broken rail risk and recommending optimal inspection schedules.
- Provide new insights regarding the effects of rail age on rail inspection scheduling.

**METHODOLOGY**

The risk-based rail defect inspection frequency optimization methodology comprises of two modules, 1) the estimation of broken rails by inspection frequency and 2) the prioritization of segment-specific inspection frequency given the total miles to inspect.

**Estimation of Number of Broken Rails by Inspection Frequency**

A number of factors can affect the rate of broken rails, such as temperature differential, rail age, traffic density, curvature, roadbed condition, axle load, vehicle dynamics, rail wear, and others [24-27]. The Volpe National Transportation Systems Center has developed an engineering model that incorporates rail defect formation, growth, and detection processes [15]. According to the Volpe model, a rail defect is assumed to form at an increasing rate as the rail ages due to the accumulation of tonnage. The model for the rate of defect formation is derived based on a Weibull distribution. The Weibull distribution model was calibrated based on observations of defect occurrence at the Facility for Accelerated Service Testing (FAST) at the Transportation Test Center in Pueblo, Colorado and on several segments of revenue track studied by the Association of American Railroads. After a defect was formed, its size progression was calibrated from the original detail fracture growth test conducted at FAST and has been further verified and validated by tests conducted through a joint international research effort supported by the Union of International Railways/World Executive Council [32, 33]. Temperature differential, axle load, track modulus, rail wear, and other factors were found to affect defect size growth. The probability of detecting a rail defect depends on the equipment used and the size of the defect. Although larger defects are more likely to be detected, they still can be missed during the inspection process [34]. The Volpe model focuses on rail fatigue defects, such as detail fractures, transverse/compound fissures, and vertical split head defects. Note that the Volpe model was developed in the 1990s based on the rail infrastructure conditions that correspond with that time period. We are unaware of recent updates to this model. This may introduce some level of uncertainty when applying this model to predict rail defects under today’s infrastructure conditions. A more detailed description of the Volpe model has been provided in Orringer [12], and thus not duplicated herein.

The Volpe model is presented as follows:

$$y_{(i-1,i)} = R \times \frac{e^{\frac{-N_{i+1}}{\beta}} - e^{\frac{-N_{i+1}+X_i}{\beta}}}{1+\lambda(X_i - \mu)} \times \lambda(X_i - \mu)$$

(1)
Where:
\[ y_{i-1,i} = \text{number of broken rails per track-mile between the (i-1)th and ith inspection} \]
\[ R = 39\text{-foot rail segments per track-mile, 273} \]
\[ X_i = \text{interval (million gross tons (MGT)) between the (i-1)th and ith inspection} \]
\[ \alpha = \text{Weibull shape factor, 3.1 [35]} \]
\[ \beta = \text{Weibull scale factor, 2150 [35]} \]
\[ \lambda = \text{slope of the number of rail breaks per detected rail defect (S/D) versus inspection interval curve, 0.014 [12]} \]
\[ \mu = \text{minimum rail inspection frequency, 10 MGT [12]} \]
\[ N_i = \text{rail age (cumulative tonnage on the rail) at the ith inspection, } N_i = N_{i-1} + X_i \]

The parameters in Equation (1) are based on published statistics in the literature. As stated earlier, broken rail occurrence is subject to many engineering factors. In the absence of detailed information for all these factors, this paper uses the two focused factors, rail age and traffic density, in the Volpe model. The methodology can be adapted to other factors or an updated version of the Volpe model in future research. Figure 2 calculates annual number of broken rails per track-mile given different rail ages (when traffic density is 80 MGT per year) using the Volpe model. We assume that the inspection intervals between each two successive inspections are identical. Each data point represents the estimated number of broken rails given the number of ultrasonic rail defect inspections per year using the Volpe model. Through a nonlinear regression, the number of broken rails per track-mile can be estimated by an exponential function of annual inspection frequencies. For example, at 1000 MGT rail age and 80 MGT annual traffic density, the relationship between the annual number of broken rails and inspection frequency is fitted as an exponential function \( y = 5.7579 \exp(-0.525x) \), where \( y \) is the annual number of broken rails per track-mile, and \( x \) is the annual inspection frequency. The coefficient of determination \( R^2 \) is more than 0.98, indicating a reasonable goodness of fit.

Similarly, an exponential regression model is found to adequately fit the data given other rail ages and traffic densities. A general exponential model is presented as follows:
\[ y = a \times \exp(bx) \]  (2)
where:
\[ y = \text{the total number of broken rail per track-mile; } \]
\[ x = \text{annual rail inspection frequency on the track segment; } \]
\[ a \text{ and } b \text{ are the model parameters which are dependent on the rail age and annual traffic density, and other factors. These model parameters can be obtained through linear regression method. } \]

For example, If rail age is 1000 MGT and annual traffic density is 80 MGT (\( a = 0.7579; b = -0.525 \)) on a segment which is inspected four times annually (\( x=4 \)), the approximate number of broken rails per track-mile is \( 0.7579 \exp(-0.525 \times 4) = 0.093 \). If this segment is 10-miles long, the annual broken rail frequency on this segment would be \( 0.093 \times 10 = 0.93 \).

**Rail Inspection Frequency Optimization**

Suppose that there are \( n \) track segments on the whole rail track with different rail ages and annual traffic densities. Define \( N \equiv \{1, 2, \ldots, n\} \) as the set of track segments. The available inspection resource is \( B \). In this paper, the inspection resource is specified in terms of the total miles inspected. Each track segment can have its own inspection frequency (denoted as \( x_i \)). The theoretical premise is that inspecting high-risk track segments more frequently may lead to a minimization of total route risk, with the equal or even fewer miles inspected. This premise was shown to be valid in Liu and Dick [31] who classify segment-specific hazardous materials transportation risks into three categories (low, medium, and high), and required all segments within each category to have the same inspection frequency. This paper significantly advances the study by Liu and Dick [31] by relaxing the requirement for group-specific inspection scheduling to segment-specific inspection scheduling. In essence, the work developed in Liu and Dick [31] is viewed as a special and simplified case of the generalized methodology developed here.

Because there are numerous possible inspection schedules to choose from, the enumeration approach is extremely time-consuming and cumbersome in identifying the optimal solution. Therefore, this research uses a mathematical optimization technique, which is devised to achieve a predetermined objective under constraints. In the context of rail inspection frequency, a generalized model is presented below:
\[
\text{MINIMIZE } \sum_{i \in N} y_i \times l_i \quad (3)
\]
where
\[ y_i = a_i \times \exp(b_i x_i) \]
Subject to
\[
\sum_{i \in N} x_i \times l_i \leq B \quad (4)
\]
where
\[ y_i = \text{total number of broken rails per year} \]
\[ x_i = \text{annual rail inspection frequency, non-negative integer} \]
\[ l_i = \text{the mileage of segment i} \]
\[ B = \text{the available inspection resource; other parameters are previously defined} \]

The objective function is to minimize the total number of broken rails on the whole route given the total miles inspected. The constraint (Equation 4) imposes the inspection resource availability restriction. The model parameters \( a \) and \( b \) can be determined based on the rail age and annual traffic density of each segment. Given the total miles inspected per year, the optimization model can be used to determine the minimum number of broken rails with an optimal allocation scheduling of inspection frequencies to each segment.

The optimization model represents a mixed integer nonlinear programming problem (MINLP). As one of the most challenging optimization models, the computational complexity of this problem increases dramatically when the number of track segments increases. For example, consider a route comprising of 100 segments. Each segment can be inspected at a frequency ranging from 2 to 7 inspections per year. Therefore, there are a total of \( 6^{100} \) (about 3700 trillion) possible scenarios of rail inspection schedules for this route. Therefore, the enumeration method or most of the off-the-shelf software cannot solve such a complex model. In an extensive review of the operations research literature, the authors find that an advanced algorithm called the Outer Approximation (OA) [36-40] can be used to solve this type of complex optimization problem. The mathematical details of the OA algorithm can be found in [37].

Computer-Aided Decision Support Tool

A computer-aided decision support tool is being developed at Rutgers University to implement the complex model formulation and solution algorithms into a practice-ready tool that can automatically generate various rail defect inspection schedules, estimate their corresponding broken rail risks, and identify optimal scheduling, given any level of resource availability. The decision support tool contains four major modules as illustrated in Figure 3.

Input Module: The user provides track-segment information such as rail age and annual traffic density. Also, the user specifies the maximum resources available (i.e. total miles inspected) and other constraints.

Calibration Module: Based on segment-specific information, the calculation module firstly estimate the total number of broken rails per track-mile given the annual inspection frequency on each segment, using the Volpe model presented in subsection “Estimation of Number of Broken Rails by Inspection Frequency”. Based on the Volpe model output, a nonlinear regression model will be developed to approximate the number of broken rails by annual inspection frequency.

Optimization Module: After an exponential regression function is calibrated to describe the relationship between the total number of broken rail per track-mile and the annual inspection frequency, a nonlinear integer programming model is developed to determine segment-specific inspection frequency. The objective is to minimize the total number of broken rails.

Output Module: The outputs include the minimum route risk that can be achieved by using the optimal inspection scheduling, and the corresponding segment-specific inspection frequency are recommended.

![Figure 3. Schematic of segment-specific inspection frequency optimization model](image)

NUMERICAL EXAMPLE

To illustrate the model application, several hypothetical scenarios are performed for a track with a total length 100 miles and 10 track segments. Each track segment has segment-specific length and rail age. Through the scenario analysis, this study analyzes the impact of different track characteristics on rail inspection frequency optimization.

Optimal Inspection Frequency under Resource Constraint

Suppose that the whole track has the uniform annual traffic density of 80 MGT. Table 1 presents segment-specific rail ages and lengths.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail Age (MGT)</td>
<td>700</td>
<td>500</td>
<td>300</td>
<td>400</td>
<td>300</td>
</tr>
<tr>
<td>Rail Length (Mile)</td>
<td>11</td>
<td>12</td>
<td>16</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Segment</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
</tr>
<tr>
<td>Rail Age (MGT)</td>
<td>200</td>
<td>800</td>
<td>400</td>
<td>900</td>
<td>600</td>
</tr>
<tr>
<td>Rail Length (Mile)</td>
<td>15</td>
<td>5</td>
<td>13</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>

Each track segment can have an annual inspection frequency of 2, 3, 4, 5, 6, or 7. In total, there are \( 6^{10} \) (600 million) possible combinations of rail inspection frequency schedules on this route. The estimated numbers of broken rails caused by total mileage inspected for possible rail inspection schedules were quantified and plotted. With the same number of miles to be inspected for each inspection schedule, some inspection schedules resulted in lower broken-rail risk than others. The inspection schedules resulting in the lowest level of
the number of broken rail given a total mileage to inspect were denoted as “optimal” schedules. Other alternative inspection schedules were called “non-optimal” schedules. These “optimal” schedules constitute a Pareto frontier (Figure 4). The Pareto frontier demonstrates the optimal inspection scheduling given limited inspection resources. For example, given a total of 400 miles inspected, the optimal inspection frequency for each segment on this route is (6, 4, 3, 4, 3, 2, 6, 4, 7, 5), with the minimal broken rail as 14. One of alternative inspection schedules denoted by gray dots is (2, 2, 6, 3, 7, 7, 5, 2, 2, 2), with the total number of broken rails as 41. Therefore, the alternative inspection schedules might cause more broken rails given the same inspection resources than the optimal schedule. This means that the first segment should be inspected 6 times per year, and the last segment should be inspected 2 times per year. The minimum number of broken rails on this route is 14 per year.

**Figure 4. Pareto optimization of broken-rail risk in terms of total miles to inspect**

The analysis shows that given total miles inspected, there exists an optimal inspection frequency schedule that can lead to a minimum broken rail risk. On the Pareto-frontier, the more miles to inspect, the lower broken rail risk under the optimal schedule. The Pareto-frontier demonstrates the necessity of the inspection scheduling optimization. The next subsection will illustrate the superiority of the optimization-based rail defect inspection frequency compared with the empirical approach.

**Comparison of Optimization-Based Rail Defect Inspection Frequency Versus the Empirical Approach**

We assume that all the 10 segments on this route have 80 MGT annual traffic density. See Table 1 for segment-specific rail age and length information. For a given number of total miles inspected, the optimal inspection schedules could result in minimum expected broken rail risk. We consider four total miles for inspection 300, 400, 500, 600, respectively. For example, “300” means that the railroad has resources to inspect a total of 300 miles per years on this route. We compare two inspection frequency schedule approaches. The first approach is an empirical heuristic that all the segments are inspected at an equal frequency. This approach is treated as a baseline in the analysis and the predicted number of broken rails is called “base risk”. An alternative approach is to optimize segment-specific inspection frequency using the MINLP model described in this paper. Railroads often use a road–rail vehicle (aka. hi-rail vehicle) that can operate both on railway tracks and on conventional roadways to inspect rail defects. This type of inspection method allows for different inspection frequencies on different track segments. Skipping the inspection of certain lower-risk segments might enable more frequent inspections of higher-risk track segments, thus maximizing the magnitude of risk reduction. The minimum number of broken rails optimized by this model is regarded as “optimization risk”. Figure 5 illustrates the inspection frequency scheduling for each track segment and the risk reduction of broken rail risk between the two approaches.

**Figure 5. Schematic illustration of ultrasonic rail inspection frequency on each track-segment**

**Remark:**

\[
Risk\ reduction = \frac{(Base\ risk - Optimization\ risk)}{Base\ risk} \times 100\% 
\]

There are risk reductions between the total numbers of broken rails obtained by the risk-based infrequency optimization model and the method in which all segments are inspected with the equal frequencies. For example, given 300 inspection mileages, if an empirical schedule calls for all track segments to be inspected three times per year; this schedule could be denoted as (3, 3, 3, 3, 3, 3, 3, 3, 3). However, the optimal inspection schedule could be (4, 3, 2, 3, 2, 2, 5, 3, 5, 4). Compared with the empirical schedule (with an inspection of all track segments three times per year), the optimal inspection schedule would reduce the broken rail number by 22.7%. With 300 inspection miles per year, the annual inspection frequency for segment G and I increase to 5 from the average inspection frequency 3, while the annual inspection for segment C, E and F decreases to 2. Other segments with larger rail ages than segment B also have an increase to some extent.

As for the empirical schedule, the inspected miles allocated on the segments with rail age 500 MGT or more account for 42 percent of the total amount of inspection resources. However, for the optimal schedule, the percentage has increased to 55 percent. This example indicates that more frequent inspection of higher-risk track segments (higher rail ages) would achieve the minimum broken rail risk.
RESEARCH CONTRIBUTION

• This research develops a new methodology to optimize the inspection frequency schedules for each segment while keeping in mind the problem of inspection resource allocation. The scenario simulation results show that effective scheduling of rail defect inspections could reduce the risk of broken rails in a cost-efficient manner.

• Using the methodology developed, a decision support tool entitled the “Rail-Risk Optimizer” is produced. This decision support tool can practically be used to prioritize resource allocation while improving the safety effectiveness of ultrasonic rail defect inspections at the same time. This would be done through determining more efficient inspection schedules in which higher-risk segments might be inspected more frequently than lower-risk segments.

• Since optimal scheduling is practically feasible given that many railroads use bimodal road-rail inspection vehicles for the detection of broken rails, this practice-ready optimization tool can lead to systematic improvements in track maintenance and inspection strategies.

CONCLUSION

This research developed a risk-based segment-specific inspection frequency model to optimize the inspection schedules for each track segment given a certain amount of inspection resources available. The model was used on scenario simulations to demonstrate the safety effectiveness of optimizing rail inspection frequency schedules. The analysis showed that prioritizing more inspections on certain higher-risk segments will minimize the total route risk with minimal additional inspection resources. Also, rail age was found to be an important influencing factor. If the percentage differences of rail ages among track segments of the same route are small, inspecting all segments with equal frequencies will lead to a near-optimum inspection schedule. If this is not the case, there could be a reduction in broken rail risk if the optimization approach is used. The research here holistically provides a generalized methodology to quantify segment-specific broken rail risk in an effort to aid decision makers in arranging their most proper rail defect inspection frequency schedules.

LIMITATIONS OF CURRENT RESEARCH & FUTURE RESEARCH

This paper focused on how to determine the segment-specific inspection frequency under a certain amount of inspection resources, thus minimizing broken rail risk. The current practice allows for adjacent segments to be inspected with different frequencies which might destroy the continuousness of the inspection. If the inspection vehicles (high-speed rail vehicles) get on and off the track too frequently, it will cause practical inconvenience. In future research, a “stretch constraint” should be added into the optimization model in order to impose a restriction on the minimum mileage within which the inspection frequency is homogenous.

This paper adopted the Volpe model developed in 1990, without update more than 25 years. It might bring about some uncertainty for predicting rail defects. In the future, an update for the Volpe model is needed based on more recent rail defect data. Additionally, future research can adapt this methodology to account for other types of rail defects. To understand the sensitivity of optimal inspection scheduling to a variety of track characteristics (segment length, traffic density and so on) better, more analyses involving varying scenarios will be developed.

ACKNOWLEDGMENTS

The first author was funded by the Department of Civil and Environmental Engineering (CEE) through a research fellowship at Rutgers University. The second author was partially funded by the Department of Civil and Environmental Engineering (CEE), Center for Advanced Infrastructure and Transportation (CAIT), and the Federal Railroad Administration. The authors thank Mr. Daniel Rodriguez for his assistance. However, the authors are solely responsible for all views and analyses in this paper.

REFERENCES


