Railway Track Geometry Defect Modeling: Deterioration, Derailment Risk and Optimal Repair

Qing He*, Hongfei Li, Debarun Bhattacharjya, Dhaivat P. Parikh, and Arun Hampapur

Qing He, Hongfei Li, Debarun Bhattacharjya
{qhe, liho, debarunb}@us.ibm.com
IBM T J Watson Research Center[1]
1101 Kitchawan Rd/ Route 134
Yorktown Heights, NY 10598

Dhaivat P. Parikh
dhaivat.parikh@us.ibm.com
IBM Global Business Service[2]
1177 S Belt Line Rd
Coppell, TX 75019-4642 USA

Arun Hampapur
arunh@us.ibm.com
IBM T J Watson Research Center[1]
1101 Kitchawan Rd/ Route 134
Yorktown Heights, NY 10598

Revised on October 28 2012

Figures + Table = 10

Total words: 4987 + 10*250 = 7487

* Corresponding author, currently affiliated with Department of Civil Engineering and Department of Industrial Engineering, University at Buffalo, The State University of New York, Email: qinghe@buffalo.edu, Telephone: 716-645-3470

Abstract

Analyzing track geometry defects is critical for safe and effective railway transportation. Repairing the right number and type of track geo-defects can appropriately reduce the probability of derailments. Additionally, prioritized track geometry repair work reduces dynamic vehicle and track interaction, thus reducing the stress state of the railroad. In this paper, we propose an analytical framework for making optimal geo-defect repair decisions by minimizing total expected costs, which include potential derailment costs and repair costs. Our major contribution lies in formulating and integrating the following three data-driven models: 1). A track deterioration model to study the degradation of Class II geo-defects; 2). A survival model to assess the derailment risk as a function of the track condition; 3). An optimization model under uncertainty for track repair decisions. In real-world examples, compared with heuristic strategies in practice, our proposed models can reduce 20% of the total composite cost on average, and potentially even more for long track sections.

1. Introduction

In the United States, rail is a crucial mode of transportation. According to the National Transportation Statistics report from the Bureau of Transportation Statistics [1], 42.7% of the United States freight revenue ton-miles were carried by railroad; this represents the largest portion of the inter-city freight market. Proper maintenance of the existing lines through repair and renewal is critical to railroad operation and safety. In 2008, Class I railroads, defined as line haul freight railroads with operating revenues of $398.7 million or more [1], spent $7.52 billion on track maintenance [2].

Track maintenance activities can be categorized into two main groups: preventive maintenance and corrective maintenance [3]. Preventive maintenance is pre-planned and carried out to avoid future defects, whereas corrective maintenance repairs existing defects in the infrastructure. Most of the literature in this area describes preventive or planned maintenance [4–11] due to its large scale of operation and high complexity, whereas very few studies have addressed the problem of corrective or unplanned maintenance [12-13], referred as *track repair* in this paper. Track repair is usually performed by the local track master in the network, and it is typically conducted on demand. Although corrective maintenance occurs on a relatively small scale as compared to preventive maintenance, it can be crucial to repair severe track defects because they may lead to catastrophic train derailments, the consequences of which can include death, injury, costs and the loss of public confidence.

Track defects have become the leading cause of train accidents in the United States since 2009. 658 of 1,890 (34.8%) train accidents were caused by track defects in 2009, incurring a $108.7 million loss [14].

According to the existing literature on track degradation [15], these defects may be categorized into one of two groups: track structural defects and track geometry defects. Track structural defects are generated from the structural conditions of the track, which include the condition of the rail, sleeper, fastening systems, subgrade and drainage systems. On the other hand, track geometry defects (referred to as *geo-defects* in the remainder of this paper) indicate severe ill-conditioned geometry parameters such as profile, alignment, gauge, cant and twist [16], shown in Figure 1. The top 12 major geo-defect types pertaining to our analysis are described in Table 1.

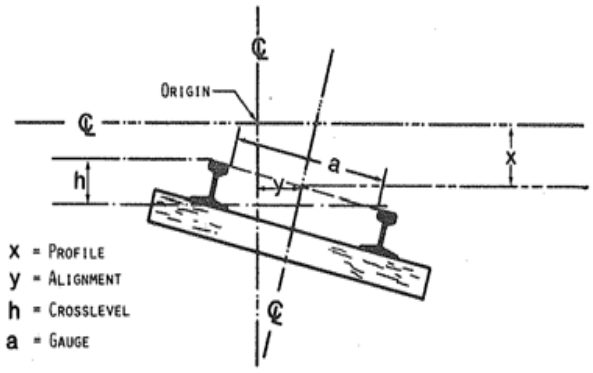


Figure 1 Track geometry parameters [19]

Previous studies on track deterioration [15–18] divide track segments into several shorter sections for analyzing summary statistics of raw geometry measurements. The overall statistics provide a measure of segment quality, called Track Quality Indices (TQIs). TQIs have been widely used for preventive maintenance scheduling [1-2], since they provide a high level assessment of railway track performance. However, TQIs only provide an aggregate level picture and they cannot identify individual severe geo-defects for track repair. According to the US Federal Railroad Administration (FRA) track safety standards, individual defects whose amplitudes exceed a certain tolerance level must be treated. Traditionally, geometry cars generally classify each defect by its severity as either Class I or Class II. Class I defects are those in violation of the FRA track safety standards, and railroads must fix these defects within a certain time period after their discovery or else they risk being fined. Class II defects are those whose amplitudes are currently below FRA limits, and they may or may not meet the particular railroad's own standards for repair. According to current practice, railroads fix Class I geo-defects immediately after inspection and they examine the Class II defects, repairing them based on their field experience. Hence, in order to make track repair decisions, it is necessary for existing railroads to address the following three questions: 1) how Class II geo-defects deteriorate into Class I defects; 2) how Class II geo-defects affect derailment risk; 3) how to prioritize and repair Class II geo-defects within a limited budget.

Table 1 Geo-defect summary (in alphabetical order)

Defect Type	Description
ALIGN	ALIGN is the average of the left and right a certain chord alignment.
CANT	Rail cant (angle) measure the amount of vertical deviation between two flat rails from their designed value. (1 degree = 1/8" for all Rail Weights, approximation)
DIP	DIP is the largest change in elevation of the centerline of the track within a certain distance moving window. Dip may represent either a depression or a hump in the track and approximates the profile of the centerline of the track.
GAGE_C	Gage Change is the difference in two gage readings up to a certain distance.
GAGE_TGHT	GAGE_TGHT measures how much tighter from standard gage (56-1/2").
GAGE_W1	Gage is the distance between right and left rail measured 5/8" below the railhead. GAGE_WIDE measures how much wider from standard gage (56-1/2"). The amplitude of GAGE_WIDE plus 56-1/2" is equal to the actual track gage reading.
GAGE_W2	Same as GAGE_W1 for concrete
HARM_X	Harmonic cross-level defect is two cross-level deviations a certain distance apart in a curve.
OVERELEV	Over-elevation occurs when there is an excessive amount of elevation in a curve (overbalance) based on the degree of curvature and the board track speed.
REV_X	Reverse cross-level occurs when the right rail is low in a left-hand curve or the left rail is low in a right-hand curve.
SUPER_X	Super cross-level is cross-level, elevation or super-elevation measured at a single point in a curve.
SURF	Uniformity of rail surface measured in short distances along the tread of the rails. Rail surface is measured over a 62-foot chord, the same chord length as the FRA specification
TWIST	Twist is the difference between two cross-level measurements a certain distance apart.
WARP	Warp is the difference between two cross-level or elevation measurements up to a certain distance apart.
WEAR	The Automated Rail Weight Identification System (ARWIS) identifies the rail weight while the car is testing and measures the amount of head loss. The system measures for vertical head wear (VHW) and gage face wear (GFW) per rail.
XLEVEL	Cross-level is the difference in elevation between the top surfaces of the rails at a single point in a tangent track segment.

The main objective of this study is to propose a framework for making optimal track geo-defect repair decisions, in order to appropriately reduce the probability of a derailment as well as its associated costs. Additionally, prioritized track geometry maintenance reduces dynamic vehicle and track interaction, thus reducing the stress state of the railroad. Our proposed analytical framework for geo-defect repair minimizes total expected costs, which include potential derailment costs and repair costs. Our major contribution is in formulating and integrating the following three models:

- A track deterioration model to study the degradation of Class II geo-defects' amplitudes

- A survival model to assess the derailment risk as a function of the track condition
- An optimization model to optimize track repair decisions

2. Data Summary and Pre-processing

The field datasets from an existing railroad include 3-year traffic data, derailment data, and geo-defect data from January 2009 to December 2011. Since main line tracks carry most of the traffic, and derailments associated with these tracks usually cost much more than other track types, we focus our analysis on about 2000 miles of main line tracks in this study. In total, there are 4,000 Class I defects and 27,000 Class II defects. The entire dataset was processed along both spatial and temporal dimensions:

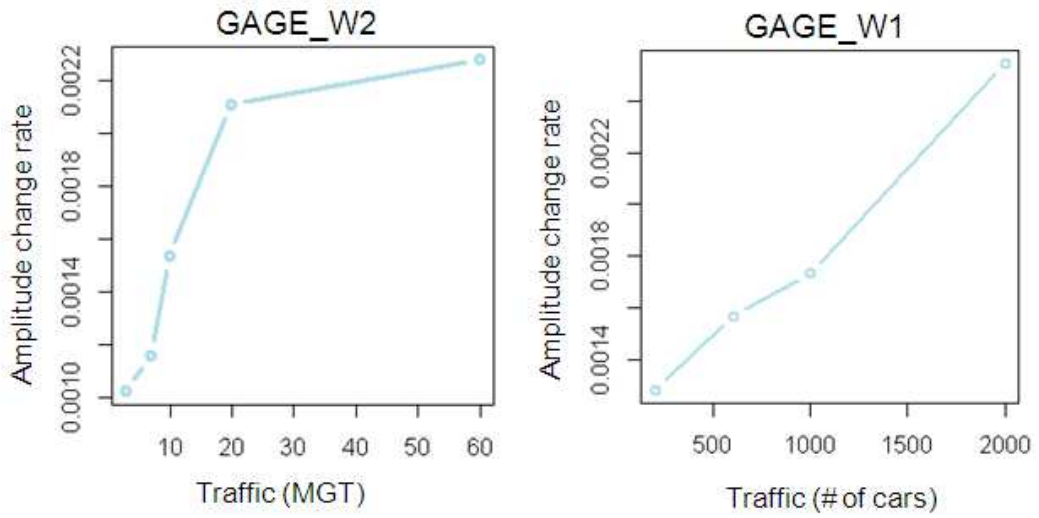
- Spatially, the rail network was originally defined by line segments (they usually connect two cities), track numbers (0-8 for main line tracks) and mile post locations. Constructed in such a fashion, the rail lines range from a few miles to hundreds of miles. To generate consistent spatial units and accommodate different modeling purposes, we divide the main line network further into two different levels of smaller segments, called *lots* and *sections*. At the finer level of granularity, each lot is 0.02 mile (about 100ft) in length, used for track deterioration analysis. At a higher level, a continuous track segment is divided into 2 mile long sections, used for track derailment risk as well as geo-defect repair modeling.
- Temporally, regular track geometry inspection is performed 3 to 6 times per year according to characteristics of each track segment. Geo-defects are reported and updated after each inspection run. When they occur in the same inspection run window, different types of geo-defects are aggregated to the level of an inspection run.

3. The Track Deterioration Model

We develop a rail track deterioration model to represent the causes and consequences of track deterioration. The model takes various factors into account, including the current track conditions and traffic information, and it has the capability to predict future track conditions. Track deterioration is captured by studying geo-defect amplitude changes, measured at each geometry inspection.

The statistical model constructs the relationships between the effective parameters and the track deterioration rate, while incorporating the uncertainty caused by the unknown factors and measurement noise. By developing the statistical model, we are able to predict the deterioration of each geo-defect and the risk of a Class II geo-defect becoming Class I in the future.

To model track deterioration, we track the evolution of track defects. However, due to the lack of geo-defect indices, it is not possible to track any particular geo-defect over time. We handle this situation by tracking the condition of small track segments, where each segment contains very few geo-defects for each inspection run. As the first step, we divide the tracks into non-overlapping lots of equal length 0.02 miles, i.e., 105.6 feet. 90% of geo-defect lengths are shorter than 100 feet and about 50% geo-defect lengths are about 30~40 feet. Then we aggregate the defects by inspection run for each defect type. We take the 90 percentile of the amplitude to represent the track segment condition for the inspection run under consideration.



(a) (b)

Figure 2 Geo-defect amplitude change rate (per day) for different geo-defect types and different factors (a) type GAGE_W2 and traffic in MGT; (b) type GAGE_W1 and traffic in number of cars

Data analysis suggests fitting different models for different defect types, since the model parameters have varying effects on deterioration rate for each defect type (see Figure 2). We assume that geo-defects get worse over time, i.e., defect amplitudes increase when there is no maintenance work. For each defect type, let $y_k(t)$ denote the aggregated geo-defect amplitude (the 90 percentile of the defect amplitudes) of the track lot k at inspection time t. The deterioration rate or the amplitude change rate over time Δt can be represented by $(y_k(t+\Delta t)-y_k(t))/\Delta t$. We model the deterioration rate (only for a single defect type) as follows:

$$\log\left(\frac{y_k(t+\Delta t)-y_k(t)}{\Delta t y_k(t)}\right)=\alpha_0+\alpha_1 X_{1k}(t)+\ldots+\alpha_p X_{pk}(t)+\varepsilon_k(t) \qquad \forall k=1\ldots N \tag{1a}$$

where N is the total number of track lots. $X_{pk}(t)$ are the *p*th external factor or predictor for *k*th track lot at inspection time *t*. Based on our exploratory data analysis, we choose to use an exponential

relationship between the external factors or predictors $X_{1k}(t),\ldots,X_{pk}(t)$ and the deterioration rate in our model, similar to the model suggested in [15]. We assume the deterioration rate is linearly related with the current track condition. The random error, $\varepsilon_k(t)$, is assumed normally distributed with mean equal to 0 and standard deviation σ^2.

The factors considered in the model include monthly traffic MGT ($X_{1k}(t)$), monthly total number of cars ($X_{2k}(t)$), monthly total number of trains ($X_{3k}(t)$),Number of inspection runs in sequence since last observed Class I geo-defect($X_{4k}(t)$), and traffic average speed in mph ($X_{5k}(t)$). Model fitting shows that factors have different impacts on deterioration rates for each defect type. The coefficients, $\alpha_0,\alpha_1,\ldots,\alpha_p$, are listed in Table 2, where α_0 is the intercept of the model, and α_i represents the coefficient for *i*th *X* factor.

Table 2 Estimation results of track deterioration model

Defect Type	α_0	α_1 - Traffic (MGT)	α_2 - Traffic (# of cars)	α_3 - Traffic (# of trains)	α_4 - Sequence #	α_5 - Traffic speed
ALIGN	-7.71E+00	--	--	--	--	--
CANT	-7.66E+00	-7.01E-02	6.05E-06	6.52E-04	6.71E-02	--
DIP	-7.58E+00	7.21E-02	--	--	--	--
GAGE_C	-8.53E+00	8.62E-02	--	--	1.13E-01	--
GAGE_W1	-7.42E+00	3.58E-02	4.64E-06	--	7.02E-02	-5.08E-01
GAGE_W2	-8.08E+00	1.90E-02	--	2.05E-04	7.98E-02	--
HARM_X	-7.44E+00	--	--	--	--	--
OVERELEV	-7.58E+00	2.45E-01	--	--	6.99E-02	
REV_X	-7.40E+00	--	--	--	--	--
SUPER_X	-8.97E+00	--	--	--	--	--
SURF	-6.99E+00	2.00E-01	--	-1.33E-03	4.36E-02	--
WEAR	-8.22E+00	2.95E-02	--	4.73E-04	7.49E-02	--
XLEVEL	-7.66E+00	--	2.64E-06	3.23E-04	9.18E-02	--

To compute the probability of a Class II defect becoming Class I in the future, we predict the defect amplitude for the next inspection run, as shown as Figure 3. Based on information about real-world inspection run intervals, we choose Δt as 90 days. Assume the threshold for of a Class II defect becoming a Class I for a certain defect type is *r*. By assuming both *r* and $y_k(t)$ are positive, we define

$$h_k(t) = \log\left(\frac{r - y_k(t)}{\Delta t y_k(t)}\right),$$

as the logarithm transformation of deterioration rate threshold of exiting amplitude $y_k(t)$. According to the assumption of linear regression model, the dependent variable

$$z = \log\left(\frac{y_k(t+\Delta t) - y_k(t)}{\Delta t y_k(t)}\right)$$

is normally distributed. Then the risk, $p_k(t)$, of a Class II geo-defect at time t on track segment k becoming Class I in Δt is

$$p_k(t) = \int_{h_k(t)}^{\infty} z dz \tag{1b}$$

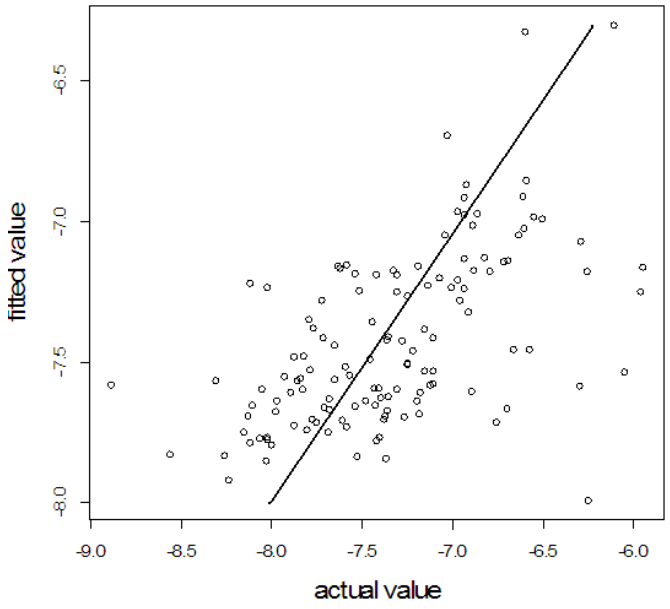


Figure 3 Predicted vs. actual geo-defect amplitude

4. The Track Derailment Risk Model

Survival analysis is the phrase used to describe the analysis of data regarding the occurrence of a particular event, within a time period after a well-defined time origin[20]. Analyzing survival times is common in many areas, for instance, in biomedical computation, engineering and the social sciences.

In our railway application, each inspection run will "refresh" the track segment since all Class I geo-defects will be repaired. If there is no derailment between two scheduled inspection runs on a track

segment, the track can be considered to have "survived" from one inspection to the next. If any derailment occurs, the track segment is said to have failed in the time period since the last inspection run.

Associated with the survival of a track section at a point in time, we refer to derailment on this track section as a hazard. In survival theory, there are three basic functions: the density function $f(t)$, survival function $S(t)$ and hazard function $\lambda(t)$. For a derailment, density function $f(t)$ expresses the likelihood that the derailment will occur at time *t*. The survival function represents the probability that the track section will survive until time *t*:

$$S(t) = \text{Prob}(T \geq t) = \int_t^{\infty} f(x)dx = 1 - \int_0^t f(x)dx = 1 - F(t) \tag{2a}$$

where *T* denotes the variable of survival time of a track segment after inspection and F(t) the cumulative function of variable *T*. The hazard function is the likelihood that a derailment takes place in time *t* given that it has lasted at least until *t*. By definition, the relationship between these three functions can be written as:

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{d \ln S(t)}{dt}$$

$$S(t) = \exp[-\int_0^t \lambda(x)dx]$$

$$f(t) = \lambda(t)S(t) = \lambda(t)\exp[-\int_0^t \lambda(x)dx]$$

The hazard function represents the instantaneous rate of failure probability at time *t*, given the condition that the event survived to time *t*. Parametric models may be used to specify the density distribution $f(t)$, such as exponential, Weibull, log-logistic, and log-normal distributions, but such pre-defined distributions may not appropriately fit the real world data. Without having to specify any assumptions about the shape of the baseline function, Cox [21] proposed a method for estimating the coefficients of covariates in the model using the method of partial likelihood (PL) rather than maximum likelihood. Hence, the Cox model is sometimes referred to as a semi-parametric model. The Cox model, which assumes that the covariates multiplicatively shift the baseline hazard function, is by far the most popular choice in practice due to its elegance and computational feasibility [22]. It has a considerable advantage compared to parametric approaches in that it does not need an assumption about the baseline hazard function. Furthermore, unlike

non-parametric analysis such as the Kaplan-Meire method and the rank test, it allows both nominal and continuous variables. The hazard function form of Cox model is,

$$\lambda(t;\beta,X)=\lambda_0(t)e^{\beta' X} \tag{2b}$$

where $\lambda_0(t)$ is an unspecified nonnegative function of time called the baseline hazard, and β is a column vector of coefficients to be estimated, and $\beta' X=\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_k x_k$. Because the hazard ratio for two subjects with fixed covariate vectors X_i and X_j,

$$\frac{\lambda_i(t)}{\lambda_j(t)}=\frac{\lambda_0(t)e^{\beta' X_i}}{\lambda_0(t)e^{\beta' X_j}},$$

is constant over time, the model is also known as the proportional hazards (PH) model. In order to estimate β, Cox [21] proposed a conditional (or partial) likelihood function which depends only on the parameter of interest. Originally he speculated that the resulting parameter estimators from the partial likelihood function would have the same distributional properties as full maximum likelihood estimators. Then he provided the mathematical proofs in [23]. The partial likelihood function is described as

$$L_p(\beta)=\prod_{i=1}^{n}\left[\frac{e^{\beta' X_i}}{\sum_{j\in R(t_{(i)})}e^{\beta' X_j}}\right]^{\delta_i}$$

The maximum partial likelihood estimator is found by solving the following equation,

$$\frac{\partial \ln(L_p(\beta))}{\partial\beta}=0$$

Table 3 Estimation results of Cox PH model

covariates	coef	exp(coef) (hazard ratio)	se(coef)	z	Pr(>|z|)
numCII_GAGE_W1	1.01E-01	1.11E+00	1.96E-02	5.165	2.40E-07
amp90_SUPER_X	5.10E-01	1.67E+00	1.68E-01	3.028	0.00246
amp90_GAGE_C	6.33E-01	1.88E+00	2.10E-01	3.013	0.00259
numCII_REV_X	6.66E-01	1.95E+00	3.21E-01	2.077	0.03782
amp90_DIP	9.58E-01	2.61E+00	4.67E-01	2.052	0.04016
amp90_WARP	7.44E-01	2.10E+00	3.69E-01	2.014	0.04404
amp90_HARM_X	5.43E+00	2.29E+02	2.78E+00	1.958	0.05028
numCII_GAGE_W2	2.28E-01	1.26E+00	1.17E-01	1.948	0.05139
amp90_WEAR	1.54E+00	4.65E+00	8.34E-01	1.843	0.06533
amp90_ALIGN	4.17E+00	6.45E+01	2.40E+00	1.736	0.08261

Fitting a Cox model can be handled using existing statistical software. The reader may refer to [24] for details of implementation. As described in Section 2, all the raw geo-defects are spatially aggregated to the section level (2 mile), and temporally into each inspection level. In one aggregated record, the dependent variable is either the time duration between two inspection runs (censored survival time), or time duration between the derailment and the last inspection run before derailment. Selected candidate predictors are listed as follows:

- Monthly traffic in MGT
- Number of Class II geo-defects (starting with "numCII") in each defect category in Table 1
- 90 percentile amplitude (starting with "amp90") of Class II geo-defects in each defect category in Table 1

The final Cox model fit to the censoring derailment data is illustrated in Table 3. An efficient way to evaluate the fitted model is to use Cox-Snell residuals [24]. If the model is calibrated correctly, the Cox-Snell residuals should show a standard exponential distribution with hazard function equal to one, and thus the cumulative hazard of the Cox-Snell residuals should follow a straight 45 degree line. The plot in Figure 4 confirms that most of the step lines are close to the dashed straight line, except for a few tail large ones. As a result, we feel there is no evidence for us to reject the model.

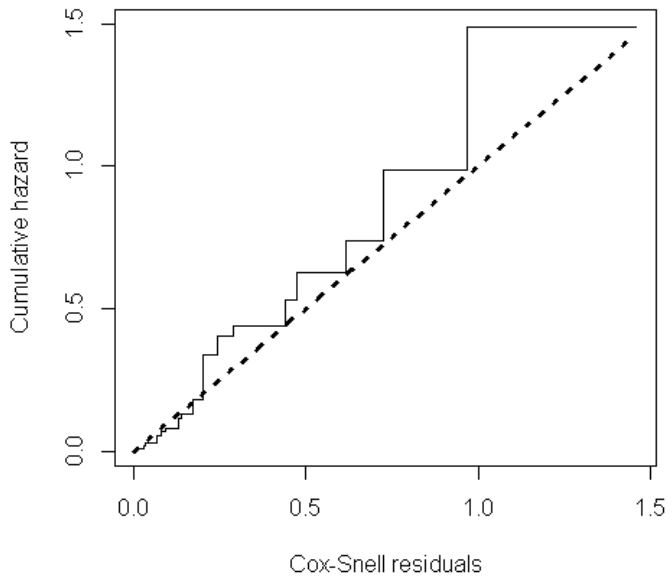


Figure 4 Cumulative hazard of Cox-Snell residuals

The model shown in Table 3 includes ten simple covariates. Each significant covariate represents a particular geo-defect type. To further explain the covariate, a positive coefficient means that the hazard is higher (hazard ratio is greater than 1.0), whereas a negative one indicates a lower hazard (hazard ratio is

less than 1.0). In the fitted model, all the covariates have positive coefficients, indicating that all geo-defect types listed in Table 3 have a strong positive impact on derailment risk. Those 10 significant geo-defect types can be categorized into two groups: the number based and amplitude based groups. On one hand, the number based group consists of GAGE_W1, REV_X and GAGE_W2, in which the number of geo-defect in a track section determines the derailment risk. On the other hand, the amplitude based group includes SUPER_X, GAGE_C, DIP, WARP, HARM_X, WEAR, and ALIGN, and the 90 percentile of geo-defect amplitudes plays an important role to influence track geometry induced derailment.

Therefore, we can derive the derailment probability, $P_i(t)$, on section *i* prior to time *t* given from:

$$P_i(t) = \text{Prob}_i(T \le t) = 1 - S_i(t) \tag{2c}$$

where $S_i(t)$ indicates the survival probability on section *i* prior to time t. Furthermore, the derailment probability after repair alternative *a* is taken, $P_i(t,a)$, can also be calculated in similar fashion,

$$P_i(t,a) = 1 - S_i(t,a) \tag{2d}$$

where $S_i(t,a)$ is the survival probability after repair action *a* is performed on section *i* prior to time *t*.

5. Optimal Track Repair

In the preceding sections, we presented two models: one to predict deterioration of Class II defects into Class I defects, and another to predict track-based derailments of trains. In this section, we use the results of the previous models along with information regarding costs as inputs for an optimization model under uncertainty.

According to FRA regulations, all Class I defects have to be repaired as soon as possible - but which Class II defects should the railways company repair? There are costs associated with repairing Class II defects, but doing so may decrease the probability of derailment, which in turn would decrease expected derailment costs. It may be particularly prudential to repair Class II geo-defects that are likely to soon become Class I defects, since they will have to be repaired in the future anyway. The decision maker in such situations is usually the local track master, who may be in charge of several sections along a line segment. For instance, the track master may be responsible for track repair decisions for 50 miles of track, i.e. 25 sections of 2 miles each.

We now formulate the track repair optimization using a simple single-stage model and describe the parameters and relevant assumptions:

Decision variables

Due to the economies of scale involved, we assume that the decision maker either repairs none of the defects or all the defects of a particular category in that section. Suppose that a section is observed to contain Class II defects of 3 defect categories. In this case, there are $2^3 = 8$ alternatives available to the decision maker, because s/he can choose to repair none or all of the defects of each type in any possible combination. We denote decision variables using indicator variables x^i_a. If alternative $a \in A$ is chosen for section $i \in I$, then $x^i_a = 1$, otherwise $x^i_a = 0$. The decision maker can choose only one of the alternatives for a particular section, and this results in a feasibility constraint: $\sum_{a \in A} x^i_a = 1 \; \forall i \in I$.

Cost and probability parameters

After an inspection run by the geometry car, the decision maker must fix all observed Class I defects. We denote the cost of repairing such defects in section i as C^i_1. If the decision maker repairs all the Class II defects of a certain category in a section, then there are costs associated with repairing these defects:

Cost of Class II defect repair = $\sum_{i \in I, a \in A} x^i_a C^i_{2,a}$, where $C^i_{2,a}$ is the cost of repairing all Class II defects for section i if alternative a is chosen.

If Class II defects are not repaired, then there will be costs associated with repairing them at the next inspection run if they turn into Class I defects. Hence similarly:

Expected cost of Class I defect repair = $\sum_{i \in I, a \in A} x^i_a C^i_{1,a}$. $C^i_{1,a}$ is the expected cost of Class I defect repair for section i if alternative a is chosen, and it includes the probability of deterioration of all Class II defects into Class I defects. Specifically, $C^i_{1,a} = \sum_{j \in J} \sum_{k \in K} p^k_a C^j_{1,a}$, where the summation is over all defect categories j in section i, and for all defects k in category j. p^k_a is the probability that this defect will progress to a Class I defect if alternative a is chosen ($p^k_a = 0$ when defect k is repaired, otherwise $p^k_a = p_k(t)$), and $C^j_{1,a}$ is the cost of fixing a Class I defect of category j when alternative a is chosen. There are also costs associated with derailment:

Expected cost of derailment = $C_D \sum_{i \in I, a \in A} x^i_a P^i_a$, where P^i_a is the probability of derailment in section i if alternative a is chosen, and C_D is the expected derailment cost. A derailment is defined as the

interruption of normal wheel-to-rail interaction, and the consequences of derailments can vary significantly, from slight equipment damage to passenger injury or even death. Figure 5 presents the derailment cost data and a fitted probability density function. The derailment cost distribution follows a heavy-tailed distribution, ranging from hundreds of US dollars to millions of US dollars. The expected derailment cost based on on our data set is around $510K.

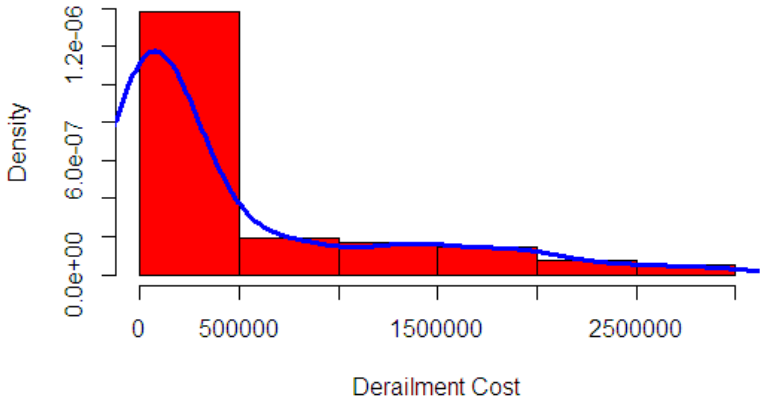


Figure 5 Probability density of derailment cost

All model parameters are summarized in Table 4.

Table 4 Model parameters for the optimization formulation

Type	Notation	Description
Indices	$i \in I$	Index for section in a line segment
	$j \in J$	Index for defect category in the set of categories observed in section i
	$k \in K$	Index for individual Class II defect of type j in section i
	$a \in A$	Index for chosen alternative out of all possible alternatives available to the decision maker
Decision variables	$x^i_a \ \forall i \in I,\ a \in A$	Indicator which is 1 if alternative a is chosen for section i, otherwise 0
Parameters	B	Budget for repair
	C_D	Average derailment cost, assumed identical for all sections
	C^i_1	Cost of repairing all existing Class I defects in section i
	$C^i_{1,a} \ \forall i \in I,\ a \in A$	Cost of repairing all converted Class I defects from existing Class II defects if alternative a is chosen for section i
	$C^i_{2,a} \ \forall i \in I,\ a \in A$	Cost of repairing all Class II defects if alternative a is chosen for section i
	$P^i_a \ \forall i \in I,\ a \in A$	Probability of a derailment in the time from this inspection run to the next, if alternative a is chosen for section i

	$p^k{}_a \ \forall k \in K,\ a \in A$	Probability that Class II defect k will progress to a Class I defect if alternative a is chosen
	$C^j{}_{1,a} \ \forall j \in J,\ a \in A$	Cost of repairing a Class I defect of category j when alternative a is chosen

The formal optimization model is as follows:

$$\text{Minimize total expected cost} = \sum_{i \in I} \sum_{a \in A} x^i{}_a (C^i_{2,a} + C^i_{1,a} + C_D P^i_a) + C^i_1 \tag{3a}$$

Subject to

$$x^i{}_a = \{0,1\} \ \forall \mathrm{i} \in \mathrm{I}, \forall a \in A \tag{3b}$$

$$\sum_{a \in A} x^i{}_a = 1 \ \forall \mathrm{i} \in \mathrm{I} \tag{3c}$$

$$\sum_{i \in I, a \in A} x^i{}_a C^i{}_{2,a} \le B \tag{3d}$$

- Objective (3a) aims to minimize the total expected cost, which is the sum of Class I and II defect repair costs in this period, expected Class I defect repair costs in the next period, and expected derailment costs in the intermediate time period.
- Equation (3b) specifies that the decision variables are binary 0-1 variables.
- Equation (3c) is for feasibility, indicating that only 1 alternative can be chosen for any particular section.
- Inequality (3d) is a capacity constraint where the total repair cost of Class II defects cannot exceed the available budget. If the budget includes both Class I and II defects for the current inspection run, then this constraint can be modified to $\left(\sum_{i \in I, a \in A} x^i{}_a C^i{}_{2,a} \right) + C^i_1 \le B$.

The optimization model is a binary integer programming (BIP) problem which is linear in the objective function as well as constraints, and it can be solved using standard commercial solvers.

Note that when the cost of derailment is a random variable, due to the nature of the optimization formulation, we can simply replace derailment cost C_D with the expected derailment cost $\overline{C_D}$ in the optimization model 3(a)-(d). This is because the total expected cost is given as:

$$\begin{aligned} E\left[\sum_{i \in I} \sum_{a \in A} x^i{}_a (C^i_{2,a} + C^i_{1,a} + C_D P^i_a) + C^i_1 \right] &= \sum_{i \in I} \sum_{a \in A} E\left[x^i{}_a \left(C^i_{2,a} + C^i_{1,a} + C_D P^i_a \right) + C^i_1 \right] \\ = \sum_{i \in I} \sum_{a \in A} x^i{}_a \left(C^i_{2,a} + C^i_{1,a} + E\left[C_D \right] P^i_a \right) + C^i_1 &= \sum_{i \in I} \sum_{a \in A} x^i{}_a \left(C^i_{2,a} + C^i_{1,a} + \overline{C_D} P^i_a \right) + C^i_1 \end{aligned}, \tag{4}$$

since the expectation of a sum is the sum of expectations. According to expression (4), the stochastic objective will eventually be derived the same as deterministic one (3a). Therefore, the proposed BIP can also handle uncertainty derailment costs.

Comparison with heuristic strategies

We compare the solution of our optimization formulation with two "baseline" heuristic strategies that are intuitive and commonly used in practice to make track repair decisions.

Heuristic I: Distance-based strategy (HEUI)

It is intuitive and cost-efficient to repair a Class II defect that is near a Class I defect of the same category. In the distance-based strategy, Class II defects close to Class I defects in the same category are likely to be repaired. We define

$$D_j^i = \min(d_j^{ik}) \quad \forall i, j$$

as the distance index of section *i* and defect category *j*, and d_j^{ik} represents the distance to the nearest Class I defect for Class II defect *k* in section *i* and category *j*. Given the budget, section and defect category, the section and category pairs (*i,j*) are ranked and selected for repair in ascending order of the distance index, until the budget is consumed.

Heuristic II: Severity-based strategy (HEUII)

In the severity-based strategy, an empirical aggregate level measure of defect amplitudes is computed for every section and category. We define the severity index as

$$S_j^i = \exp\left(\frac{\max(\omega_j^{ik})}{\omega_j^{\max}}\right)\ln(n_j^i) \quad \forall i, j$$

where ω_j^{ik} is the amplitude of Class II defect *k* in section *i* and category *j*, $\omega_j^{\max}$ denotes the maximum amplitude of geo-defect category *j*, and n_j^i represents the number of Class II defects in section *i* and category *j*. These severity measures are ranked and each section and category pair (*i,j*) is repaired in descending order of severity, until the budget is consumed.

6 Numerical Examples

Four track segments denoted A, B, C and D, are selected for track repair optimization for geo-defects located in the last inspection run in December 2011. A summary of track information and geo-defects is

shown in Table 5. Segments are divided into no more than 2 mile sections according to traffic data. It is very hard to estimate exact costs of fixing individual geo-defects, due to variations of equipment, fleet and personnel at each local maintenance center. For simplicity, we assume all Class II and Class I defects cost $50 and $100 respectively, according to empirical estimation. Since Class I defects need to be repaired by law, it will be more convenient to fix Class II defects whereever there is a Class I defect in the same category nearby. Therefore, the repair costs are assumed to be reduced by 50% for those Class II defects within 5 miles of a Class I defect in the same category.

Table 5 Summary of test track segments

Name	Length (mile)	Num. of sections	Num. of Class II defects	Num. of Class I defects	Budget (USD)	Geo-defect types observed
A	50	13	60	4	1200	GAGE_C, GAGE_W1, XLEVEL
B	100	21	77	2	2500	GAGE_C, GAGE_W1, HARM_X, OVERELEV, TWIST, XLEVEL
C	300	49	122	37	4500	GAGE_C, GAGE_W1, GAGE_W2, HARM_X, OVERELEV, REV_X, TWIST, XLEVEL
D	500	94	343	38	6500	ALIGN, GAGE_C, GAGE_W1, GAGE_W2, OVERELEV, REV_X, TWIST, XLEVEL

For the sake of comparison, 100 scenarios of derailment costs are randomly sampled from the density function fitted in Figure 5. The proposed optimization model (BIP) can be solved within 1 second for all four instances, by CPLEX 12.3 on a personal computer with quad core 2.4GHz CPU and 3G RAM. Our framework is therefore suitable for online track repair activities.

Figure 6 plots total expected costs for each scenario for our optimal solutions (BIP) and two heuristic strategies (HEUI & HEUII). As expected, note that almost all the points are above the 45 degree straight line, since the optimal solution guarantees better solutions than the heuristic strategies by definition. the difference between the costs is particularly evident for long tracks such as Track D. In Figure 6(d), total expected costs of the optimal track repair solution vary from $200K to $400K, while HEUI and HEUII range from $200K to $900K.

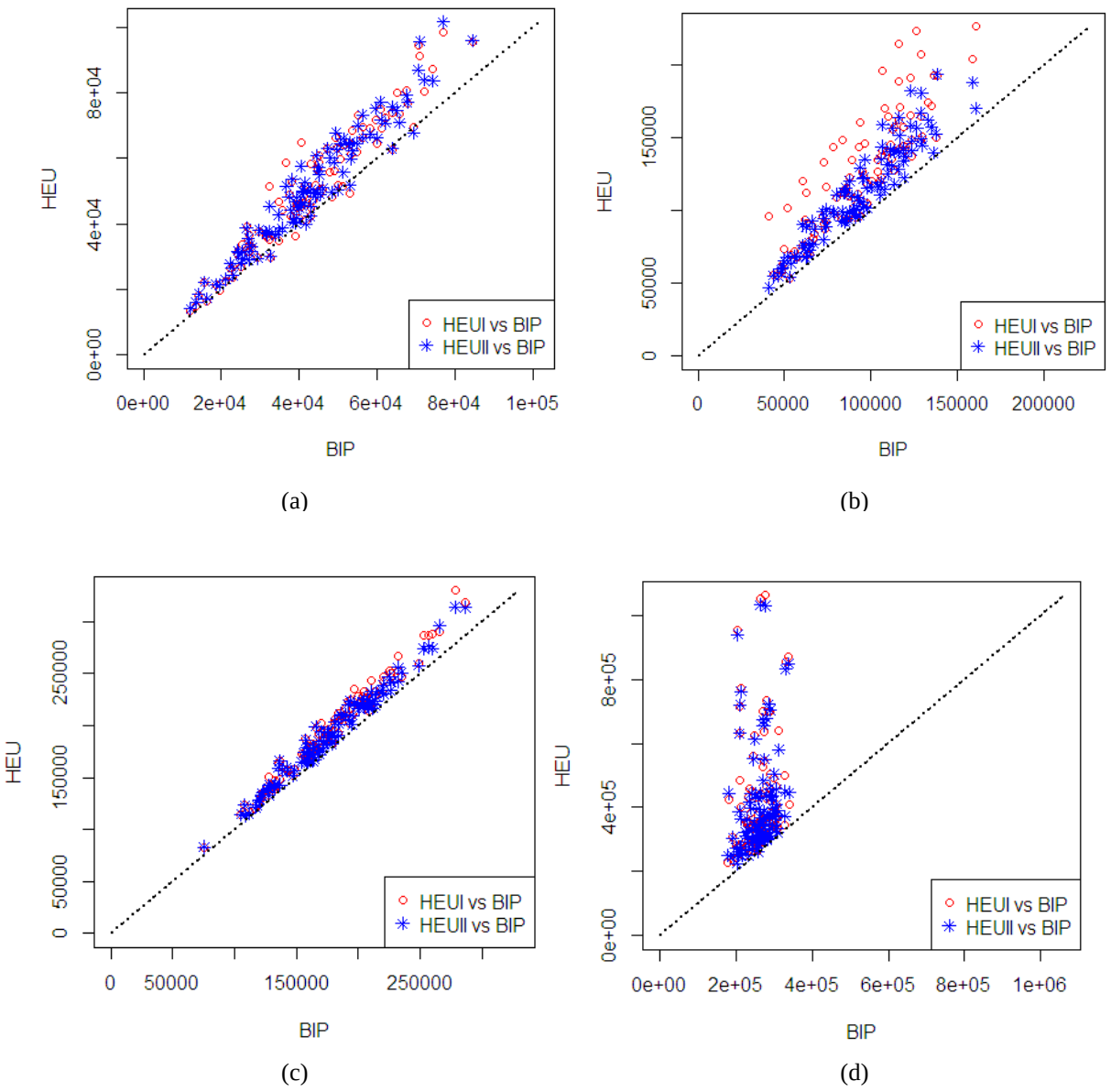


Figure 6 Comparison of total expected costs from Optimal, HEUI and HEUII under 100 different derailment cost scenarios (a) Track A; (b) Track B; (c) Track C; (d) Track D

Figure 7 compares average total costs across different strategies with a bar chart. HEUII produces lower total cost than HEUI does for all four track segments. It means that fixing geo-defects with higher severity can reduce more on derailment costs, which complies with the conclusion from derailment risk model. Among the three strategies, BIP always outperform the baselines of course. The percentage changes of total costs from HEUI & HEUII to optima; are labeled on the top of bars, and they range from -7% to -38%. When the repair problem gets larger in scale, from track A to D, BIP shows a linear increasing curve,

while the baseline costs appear to increase exponentially. On average, BIP decreases around 20% of total costs, which is equivalent to about $50K.

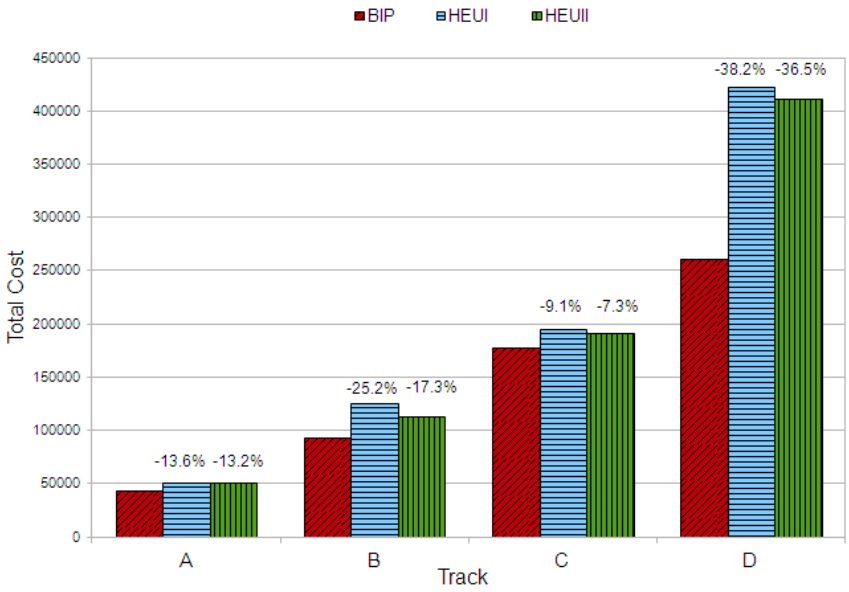


Figure 7 Comparison of average total costs from BIP, HEUI and HEUII

7 Conclusions

This paper presents an analytical framework to address the track geo-defect repair problem using three models. First, a statistical deterioration model is specially designed for track Class II geo-defects, and it aims to model deterioration rates for each geo-defect category. Second, a track derailment risk model is proposed by applying survival analysis. This model incorporates combinations of effects from different types of geo-defects, and represents how derailment risk changes over time. Finally, an optimization model is formulated for making geo-defect repair decisions. A real-world case study has shown that the proposed methodology yields more reliable and less costly solutions than typical heuristic strategies used according to current industry practices. Our proposed track repair model can reduce total costs by 20% or more. Our experiments demonstrate that particularly large savings of around 38% can be incurred for long track sections.

References

[1] Association of American Railroads, “Class I Railroad Statistics.” 2011.

[2] D. Tolliver and D. Benson, “Freight Railway Track Maintenance Cost Model,” North Dakota State University, MPC-357, 2010.

[3] M. Andersson, “Strategic Planning of Track Maintenance - State of the Art,” Department of Infrastructure, Royal Institute of Technology, TRITA-INFRA 02-035, 2002.

[4] F. Beichelt and K. Fischer, “General Failure Model Applied to Preventive Maintenance Policies,” *IEEE Transactions on Reliability*, vol. R-29, no. 1, pp. 39–41, 1980.
[5] G. Budai, D. Huisman, and R. Dekker, “Scheduling Preventive Railway Maintenance Activities,” *Journal of Operational Research Society*, vol. 57, no. 9, pp. 1035–1044, 2006.
[6] B. S. N. Cheung, K. . Chow, L. C. K. Hui, and A. M. . Yong, “Railway Track Possession Assignment Using Constraint Satisfaction,” *Engineering Applications of Artificial Intelligence*, vol. 12, no. 5, pp. 599–611, 1999.
[7] A. Higgins, “Scheduling of Railway Track Maintenance Activities and Crews,” *Journal of Operational Research Society*, vol. 49, no. 10, pp. 1026–1033, 1998.
[8] A. Higgins, L. Ferreira, and M. Lake, “Scheduling Rail Track Maintenance to Minimise Overall Delays,” in *The 14th International Symposium on Transportation and Traffic Theroy*, Jerusalem, Israel, 1999.
[9] F. Peng, S. Kang, X. Li, Y. Ouyang, K. Somani, and D. Acharya, “A Heuristic Approach to the Railroad Track Maintenance Scheduling Problem,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 26, no. 2, pp. 129–145, 2011.
[10] S. Simson, L. Ferreira, and M. Murray, “Rail Track Maintenance Planning: An Assessment Model,” *Transportation Research Record: Journal of the Transportation Research Board*, no. 1713, pp. 29–35, 2000.
[11] J. I. van Zante–de Fokkert, D. den Hertog, F. J. van den Berg, and J. H. M. Verhoeven, “The Netherlands Schedules Track Maintenance to Improve Track Workers’ Safety,” *Interfaces*, vol. 37, no. 2, pp. 133–142, 2007.
[12] J. Zhao, A. H. C. Chan, and M. P. N. Burrow, “Reliability Analysis and Maintenance Decision for Railway Sleepers Using Track Condition Information,” *The Journal of the Operational Research Society*, vol. 58, no. 8, pp. 1047–1055, 2007.
[13] T. Oyama and M. Miwa, “Mathematical Modeling Analyses for Obtaining an Optimal Railway Track Maintenance Schedule,” *Japan Journal of Industrial and Applied Mathematics*, vol. 23, no. 2, pp. 207–224, 2006.
[14] F. Peng, “Scheduling of Track Inspection and Maintenance Activities in Railroad Networks,” Ph.D. Dissertation, University of Illinois at Urbana-Champaign, 2011.
[15] J. Sadeghi and H. Askarinejad, “Development of Improved Railway Track Degradation Models,” *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance*, vol. 6, no. 6, pp. 675–688, 2010.
[16] A. R. B. Berawi, R. Delgado, R. Calçada, and C. Vale, “Evaluating Track Geometrical Quality through Different Methodologies,” *International Journal of Technology*, no. 1, pp. 38–47, 2010.
[17] Federal Railroad Administration, “Development of Objective Track Quality Indices,” Federal Railroad Administration, Washington, DC, RR05-01, 2005.
[18] M. El-Sibaie and Y.-J. Zhang, “Objective Track Quality Indices,” *Transportation Research Record: Journal of the Transportation Research Board*, no. 1863, pp. 81–87, 2004.
[19] D. Collett, *Modelling Survival Data in Medical Research*. Chapman & Hall/CRC, 2003.
[20] D. R. Cox, “Regression Models and Life-Tables,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, no. 2, pp. 187–220, 1972.
[21] M. A. Cleves, W. Gould, and R. Gutierrez, *An Introduction to Survival Analysis using Stata*. College Station, TX: Stata Press.
[22] D. R. Cox, “Partial Likelihood,” *Biometrika*, vol. 65, no. 2, pp. 2269–276, 1975.
[23] T. . Therneau and P. M. Grambsch, *Modeling Survival Data: Extending the Cox Model*. Springer, 2000.
[24] D. R. Cox and E. J. Snell, “A General Definition of Residuals,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 30, no. 2, pp. 248–275, 1968.
[25] A. Shapiro, “Monte Carlo Sampling Methods,” in *Handbooks in Operations Research and Management Science*, vol. 10, 2003, pp. 353–426.