Budgeting the Adoption of Sensors on Connected Trains

Raj Bridgelall, Ph.D., Corresponding Author

Assistant Professor, Department of Transportation, Logistics & Finance

College of Business, North Dakota State University

Fargo, ND 58108; Email: raj@bridgelall.com, ORCID: 0000-0003-3743-6652

Denver D. Tolliver, Ph.D.

Director, Upper Great Plains Transportation Institute, North Dakota State University

Fargo, ND 58108; Email: denver.tolliver@ndsu.edu, ORCID: 0000-0002-8522-9394

Declarations of Interest: None

Word Count: 6109

Budgeting the Adoption of Sensors on Connected Trains

Abstract

Railroads can save millions of dollars by deploying multi-sensor track scanners on connected trains to detect track and roadbed problems that could cause accidents. However, uncertainties about performance and return-on-investment impeded the development and deployment of such sensor systems. This research develops a budget model that both manufacturers and railroads can use to decide on a suitable tradeoff between price affordability and achievable performance. A case study of five Class 1 railroads demonstrates that a payback within two years is achievable at $4,000 per device and an annual maintenance cost of one-quarter the system deployment cost.

Keywords: Benefit-cost analysis; non-destructive evaluation; payback period; positive train control; railroad safety; return-on-investment.

1. Introduction

Railroads use less energy to move more bulk freight across longer distances than any other mode of transportation (BTS 2020). Hence, their safe, efficient, and reliable operation is critical to the economic health of a nation. Nevertheless, U.S. railroads had an annual average of 2,570 accidents over the past decade (Figure 1), based on the authors' analysis of the Federal Railroad Administration (FRA) railroad accident database (FRA 2020). Those accidents resulted in an industry average financial loss of $376 million each year (Figure 1). Human errors caused more than 35% of those accidents. Consequently, railroads recently deployed a government-mandated system called positive train control (PTC) to prevent accidents due to human errors (Bridgelall and Tolliver 2020).

The next dominant accident cause after human errors was track and roadbed (T&R) problems, which consistently accounted for more than 23% of the accidents each year. Railroads currently use multisensory track scanning (MTS) systems on dedicated railroad inspection vehicles to scan for T&R problems (Chia, et al. 2018). However, this method has several disadvantages. First, railroads lose revenue service capacity when dedicated condition monitoring vehicles and track inspectors occupy the maintenance-of-way. Second, the finite allocation of track inspection resources limits the space-time coverage of the network (Peng and Ouyang 2012). Third, infrequent scanning can result in failure to detect problems that develop between scanning cycles (Rahimikelarijani, Mohassel and Hamidi 2020). Furthermore, the prediction of track geometry degradation is challenging (Cárdenas-Gallo, et al. 2017) (Karimpour, et al. 2018). Those issues led to a growing body of research to overcome the challenges of installing MTS systems on revenue service trains (Weston, et al. 2015).

An MTS system consists of a suite of sensors, wireless communications, power supply, and software that can intelligently detect many T&R problems. Sensor types include optical, acoustic, radar, LiDAR, inertial, and electromagnetic (Chia, et al. 2018). There is no one type of sensor that can detect all types of T&R problems. Therefore, MTS systems must combine multiple types of sensors and the appropriate signal processing methods (Li, et al. 2017). Earlier efforts focused on wireless sensor networks (WSN) to combine the data from multiple sensors situated throughout the train (Hodge, et al. 2014). More recently, the Internet-of-Things (IOT) movement generalized the concept of connected railroads (Fraga-Lamas, Fernández-Caramés and Castedo 2017). PTC is an example of connected railroads that can communicate train operating parameters to a cloud-based system for decision making about operating and maintenance strategies (Brezulianu, et al. 2020).

Steady reductions in price, power consumption, and size continue to improve the feasibility of adding sensors to connected trains (Bernal, Spiryagin and Cole 2018). However, railroads have not yet deployed MTS systems on service trains because of uncertainties about the upfront capital needed, performance, and the timing of a return-on-investment (ROI).

To further the development of PTC-compatible MTS systems, manufacturers need to determine a price target that railroads can evaluate for affordability, and conduct engineering to determine product feasibility. The **objective** of this research is to develop a closed-form price budget model that both railroads and engineering can use to quickly explore scenarios that could be mutually beneficial. Section 2 explains how this work relates to similar research involving the benefit-cost analysis (BCA) of railroad technology deployments. The **contributions** of this research are:

- The derivation of a closed-form mathematical price budget model that researchers can customize for their own applications (Section 3)

- Descriptive analysis of trends in Class 1 railroad financial loss from accidents that MTS systems could help to prevent (Section 4.1)

- A case study of railroad-specific device price budgets (DPB) to guide product development and deployment decisions (Sections 4.2 to 4.5).

- A sensitivity analysis for each Class 1 railroad to explore the effects of trading off system performance and annual maintenance cost (Section 4.6)

The discussion (Section 5) provides some additional insights about model usage and some limitations of the data used. Section 6 concludes with a discussion of the generalization of the model for other application areas and briefly describes future work already in progress.

2. Related Works

Early attempts to investigate the use of onboard devices to automatically monitor for T&R condition focused on feasibility assessment and performance evaluation. Lee et al. (2012) mounted accelerometers to the axle box and the bogie of a high-speed train to estimate track geometry measurements from the lateral and vertical accelerations (Lee, et al. 2012). Around the same time, Mori et al. (2013) reported that a compact size, battery-powered device could effectively estimate track irregularities from car body dynamics (Mori, et al. 2013).

The continuous cost reduction of smartphone technology and advancements in their sensing capabilities led to more recent studies about their potential use on trains. Paixão et al. (2019) found that acceleration measurements using smartphones can identify critical situations that increase derailment risks (Paixão, Fortunato and Calçada 2019). Bridgelall & Tolliver (2020) demonstrated that localization errors from GPS reception issues can be improved by combining

the signals from multiple train traversals (Bridgelall and Tolliver 2020). Balouchi et al. (2020) developed a cab-based track monitoring system and found good agreement with ground truth measurements (Balouchi, Bevan and Formston 2020).

The continuous advancements in data mining, artificial intelligence, and machine learning has motivated their application to the analysis of data from vibration, sound, and image sensors aboard passing trains. Tsunashima (2019) reported that machine learning algorithms could classify detected track faults from their vibration signatures (Tsunashima 2019). Farlik and Tabaszewski (2020) trained three independent neural networks with vibration signals to detect track issues and found that the solution is sensitive to train speed (Firlik and Tabaszewski 2020). Sun et al. (2020) applied machine learning algorithms to sound signals and found that support vector machines provided the best performance in predicting track switch condition (Sun, et al. 2020). On the other hand, Bukhsh et al. (2019) found that using tree-based classification techniques to predict the maintenance needs for railway switches can provide greater interpretability than other types of machine learning models (Bukhsh, et al. 2019). Sysyn et al. (2019) found that applying machine learning to high-resolution images of rail crossings can detect surface cracks that are evidence of rail contact fatigue (Sysyn, et al. 2019). Lasisi and Attoh-Okine (2019) combined bagging and boosting ensemble classifiers to enhance the prediction accuracy of annual fatigue track defects (Lasisi and Attoh-Okine 2019).

A recent survey found that even though freight railroads believe that the adoption of automation technologies will bring future benefits, they have significant concerns around training, deskilling, and technology performance (Brooks, et al. 2017). There has been little analysis at the intersection of engineering and railroad decision making about technology adoption. There were no efforts reported in the literature to develop a model that could help

engineering and railroads evaluate the tradeoffs among price, device performance, and ROI. This work will help to reduce uncertainties about the affordability and performance of onboard MTS systems that are suitable for deployment on connected service trains.

3. Methods

The next subsections develop the DPB as a function of the desired payback period, discount rate, annual maintenance cost, and system performance.

3.1 Return on Investment

The theory of BCA determines an ROI by accumulating the annual net benefits from an initial investment or cost to acquire a system. The definition of cumulative discounted net benefit is

$$\text{ROI} = \sum_{i=1}^{Y} \frac{B_i - C_i}{(1 + r)^i} \tag{1}$$

where B_i and C_i are the discounted annual benefits and costs of maintaining the system, respectively. The variable r is the discount rate, i is the future year, and Y is the total number of years for payback. The convention is to equate ROI with the discounted net benefits accumulated after Y years only. Therefore, any additional net benefits realized after the Y years of payback is bonus savings beyond the ROI.

In this application of BCA, the annual benefits are the *average* amount that the MTS system will save a railroad in accident prevention. The annual costs are the *average* annual amount that the railroad will spend to operate and maintain the installed system, including software licensing fees. The payback period is the value of Y that solves

$$\sum_{i=1}^{Y} \frac{B_i - C_i}{(1 + r)^i} = C_\text{T} \tag{2}$$

where C_T is the capital cost to deploy the system. The modeling treats C_T as the system price

budget (SPB) that railroads will evaluate to consider affordability. Analyst use an average value

for benefit and cost when the annual fluctuations are unknown. For example, it is not possible to

predict the exact financial loss from accidents each year. Therefore, using the historic average

annual financial loss serves as the best estimate of future annual financial losses from accidents,

if everything else remains unchanged. Similarly, it is not possible to know the future cost of

maintenance each year. However, an average value based on empirical knowledge of railroad

operations and maintenance would be the best estimate. Furthermore, using an average value for

each year helps to simplify the model and expose smoothed trends.

The average value is unchanged each year, therefore, the model can treat it as a constant.

This also enables *normalization* with other variables. Selecting the estimate of the average

annual benefit as the constant produces the following normalization: C_T can be a proportion η of

B_i and C_i can be a proportion λ of B_i. This normalization transforms Equation (2) to the form

$$\sum_{i=1}^{Y} \frac{B_i - \lambda B_i}{(1+r)^i} - \eta B_i = 0 \tag{3}$$

Given that B_i is a constant each year, it can move outside of the summation. Then, dividing both

sides by B_i reduces the expression to

$$(1 - \lambda) \sum_{i=1}^{Y} \frac{1}{(1+r)^i} - \eta = 0 \tag{4}$$

which is independent of the estimated average annual benefit. The normalization with respect to

the average value for annual benefits allows an analyst to *parameterize* costs in *direct proportion*

to estimates of the average annual benefits. Such parameterization enables future sensitivity

analysis to gain insights.

The variable normalization refines the model with parameterization, but it is still not in closed form. That is, a solution for η as a function of Y is necessary. Such a solution cannot be determined algebraically. Therefore, the strategy was to develop a numerical solution by formulating an optimization problem for a set of Y values as

$$\text{Minimize:} \quad (1-\lambda)\sum_{i=1}^{Y}\frac{1}{(1+r)^i} - \eta$$

$$\text{Subject to:} \tag{5}$$

$$\eta > 0$$

$$Y \in \{1, 2, \dots, 20\}$$

Subsequently, fitting a logarithmic function to the solution set $\{Y, \eta\}$ yielded a closed-form solution as

$$\eta = \alpha \ln Y + \beta \tag{6}$$

where the parameters α and β were determined by minimizing the sum-of-squared error between the solution set and the estimated function.

Accounting for the normalization, the SPB, C_T, is

$$C_T = \eta B_A = B_A(\alpha \ln Y + \beta) \tag{7}$$

where B_A is the estimate average annual benefits realized as expenses prevented from T&R accidents each year. Depending on the average performance of the system, the value of B_A will be in direct proportion, P_n, to the average annual financial loss, A_T, due to accidents that the system can help to prevent. Hence,

$$B_A = A_T P_n \tag{8}$$

and after combining equations, the SPB becomes

$$C_T = A_T P_n(\alpha \ln Y + \beta) \tag{9}$$

where P_n depends on the system performance developed in the next subsection.

3.2   System Performance

The proportion P_n of annual accidents that the system can detect is the probability of detecting an existing problem after some number of attempts. The detection probability is a function of the first-scan probability, P_1, which characterizes the system performance in an environment. That is, P_1 is the probability that the system can detect a problem after the first scan. Hence, the probability of not detecting the problem after the first scan is $(1 - P_1)$. Furthermore, the probability of not detecting the problem after n consecutive and independent scans is $(1 - P_1)^n$. Therefore, the probability of detecting the problem after n consecutive scans is

$$P_n = 1 - (1 - P_1)^n. \tag{10}$$

The theory of signal detection involves a fundamental tradeoff between the false negative rate (missed detections) and the false positive rate (noise detected as signal) of a system. Figure 2 illustrates the relationships between the first-scan probability, the signal detection threshold, and the statistical separation between signal and noise. The example shows a hypothetical distribution of noise amplitude relative to a distribution of signal amplitude for a given data collection environment. The relative distributions can differ under conditions of lighting, electromagnetic interference, and other environmental factors that could affect the system performance, regardless of the detection threshold setting. The value of P_1 is the area under the normalized signal distribution curve and above the signal detection threshold. This example shows that for a normally distributed signal, setting a detection threshold at the mean a_1 will yield an area under the upper half of the curve of $P_1 = 0.50$. Alternatively, setting a detection threshold at the first upper standard deviation a_2 will result in $P_1 = 0.364$. At the a_2 threshold setting ($P_1 = 0.364$) the false positive rate will be zero (or negligibly low) because the noise amplitude never exceeds that threshold. However, decreasing the threshold below a_2 will further

increase P_1 because the area under the signal distribution curve will increase. However, the area under the noise distribution curve, above the lower threshold, will also increase, which will increase the false positive rate.

False positives are costly to railroads because they must send human inspectors to the location to verify a problem detection. Decreasing P_1 by increasing the signal detection threshold will reduce the false positive rate. However, doing so will also increase the number of scans required to detect a problem as governed by Equation (10). The network traffic constrains the average number of scans per day achieved across any segment of track.

3.3 Data Mining

The average annual financial loss from accidents due to T&R problems was determined by mining the FRA railroad accident data (FRA 2020). The database labels the cause of an accident into one of five broad categories: human error, T&R problem, mechanical failure, equipment failure, and signaling error. The analyzed scenarios focused on Class 1 railroads (Friebel, McCullough and Angulo 2014) because they will benefit most from deploying MTS systems. The data mining revealed that only five of the Class 1 railroads consistently reported accidents for the decade 2009 to 2019, so the case study focused on those railroads.

Locomotives (power-units) must host MTS systems because they are typically the only source of power and communications on a train. Consequently, the size of the initial investment for maximum network coverage would be directly proportional to the locomotive fleet size. The number of in-service locomotives L_M for a railroad was determined from their annual report to the U.S. Surface Transportation Board (STB). To minimize the upfront capital required, only one of the power-units of a multi-locomotive train need to host the MTS system. Therefore, the number of MTS systems to install, N_S, is

$$N_S = \frac{L_M}{L_\sigma} \qquad (11)$$

where L_σ is the average number of locomotives per train and L_M is the number of locomotives in service. The value of L_σ was determined as a ratio of the annual locomotive miles to the annual train miles listed in the STB report.

The value of n needed to determine the probabilistic performance of the system is a function of the annual train traffic of a railroad. That is, the average number of scans per day per track segment is the average network traffic measured in trains per day, T_d. Hence, $n = T_d$, which the STB report lists. The average annual train traffic across the network is the ratio of the number of train miles to the number of track miles used to move those trains. The STB report has both values for each operating year.

3.4 System Price Budget

The SPB is the maximum purchase price that would allow a railroad to achieve an ROI or payback in Y years. The SPB is a function of the installation cost, C_I, and the DPB such that

$$C_T = C_I \left\lceil \frac{N_S}{K_S} \right\rceil + N_S C_D \qquad (12)$$

where N_S is the number of devices to install and C_D is the DPB. The model accommodates installation cost by batches of K_S devices. The ceiling function $\lceil \cdot \rceil$ assures that the number of batches is an integer. This model generalizes for a fixed price to retrofit a group of locomotives in a yard. To the extreme, $K_S = 1$ when there is no group pricing for installations. Substituting Equation (9) into Equation (12) yields

$$C_I \left\lceil \frac{N_S}{K_S} \right\rceil + N_S C_D = A_T P_n (\alpha \ln Y + \beta) \qquad (13)$$

3.5 Device Price Budget

Substituting Equation (10) into Equation (13) and solving for the DPB, C_D, yields

$$C_D = \frac{A_T(1-(1-P_1)^n)(\alpha \ln Y + \beta) - C_I\left[\frac{N_S}{K_S}\right]}{N_S} \tag{14}$$

Substituting n for T_d as derived from the STB report and Equation (11) into Equation (14) yields the full closed-form cost model as

$$C_D = \frac{L_\sigma}{L_M}\left[A_T(1-(1-P_1)^{T_D})(\alpha \ln Y + \beta) - C_I\left[\frac{L_M}{L_\sigma K_S}\right]\right]. \tag{15}$$

4. Results

Each subsection of this section parallels those of the methods section to provide the results of the BCA modeling, data mining of the FRA accident database, and the data extraction from the STB reports.

4.1 Return on Investment

Figure 3 shows the numerical solution of the BCA for three scenarios of annual maintenance cost at 5%, 15%, and 25% of the SPB. The overlapping dotted lines are the functions fitted to the series of SPB results for each payback year. Each scenario used the standard FRA discount rate of 7% for BCA (FRA 2016). The scenario for annual maintenance cost at 15% of the SPB uses the recommended FRA discount rate of 3% to assess the sensitivity. The analysis indicates that lowering the discount rate from 7% to 3% produces a relatively small increase in the SPB within the first five years and increases by a factor of 13% two decades later. The insets show the best fit functions to the numerical solutions for each scenario of the BCA. The goodness-of-fit

measure is based on the coefficient of determination, R^2. Table 1 summarizes the parameter estimates in Equation (6) and the R^2 values for each scenario.

4.2 System Performance

As discussed previously, the first-scan probability depends on the signal detection threshold and the amount of separation between signal and noise during data collection. Figure 4 shows the probability of detection, P_n, as a function of the first-scan probability and three values for the number of scans, n, per track segment each day. As observed, the sensitivity of P_n to P_1 increases rapidly as the number of scans exceed 10.

4.3 Data Mining

Three Class 1 railroads did not consistently report data for accidents each year of the past decade, so they did not qualify for the case study. Figure 5 illustrates the cumulative trend in financial loss, A_T, from accidents due to T&R problems. The ranking of financial loss from most to least were from BNSF Railway Company (BNSF), Union Pacific Corporation (UP), CSX Transportation (CSX), Norfolk Southern (NS), and Kansas City Southern (KCS). The table also summarizes the main statistics needed from the 2019 STB reports, and the calculated metrics as previously described. The scenarios analyzed focused on revenue freight trains where locomotives moved unit, way, and through trains across the networks. Table 2 summarizes the financial loss for each railroad, averaged over the past decade.

4.4 System Price Budget

The SPB involved a trade-off between the installation cost and the DPB. That is, given a fixed SPB, an increase in the batch installation cost requires a proportional decrease in DPB. The sensitivity is determined by taking the derivative of DPB with respect to the installation cost where

$$\frac{dC_{\mathrm{D}}}{dC_{\mathrm{I}}} = -\frac{1}{K_{\mathrm{S}}} \tag{16}$$

This result indicates that for $K_{\mathrm{S}} = 1$, each unit of price increase for installation requires an equal

unit of DPB decrease. However, any amount of DPB decrease will increase the difficulty of

developing a high-performance device within a smaller cost budget. Increasing the batch size

will reduce the sensitivity but doing so will also increase the labor requirements for each

installation. This result points to a better strategy, which is to eliminate the installation cost and

rely on a higher annual maintenance fee to recover related expenses over time. Doing so will

maximize the DPB that engineering can work with to develop a viable product.

4.5 Device Price Budget

Table 3 summarizes results of the SPB, DPB, and annual maintenance costs calculated for

nominal parameter values. Those nominal values are a discount rate of 7%, zero installation

charge, an annual maintenance proportion of 25% of the total system cost, a desired payback

period of 2 years, and a first-scan probability $P_1 = 0.50$.

4.6 Sensitivity Analysis

Figure 6 displays the DPB sensitivities to P_1 and λ for each of the railroad analyzed.

The horizontal axis is the desired payback year after deployment. The vertical axis is the DPB.

A regression analysis between the DPB and the accident cost per train revealed a strong

correlation with a coefficient of determination, R^2, ranging from 0.94 to 1.0. Figure 7 shows the

regression for three scenarios of P_1. The regression fit is a linear relationship for a given payback

period and annual maintenance cost proportion of the SPB. For example, if the annual

maintenance cost is 25% of the SPB and the desired payback period is 2 years, the minimum

DPB ranges from \$44,000 to \$6,000 when first-scan probabilities range from 15% to 50%,

respectively.

5. Discussion

Data mining the FRA accident database reveals that since 2013, the analyzed Class 1 railroads

experienced a declining trend in financial loss from accidents due to T&R problems. The decline

was less than \$30 million over 7 years. The trend shows that the financial losses for the five

railroads have plateaued near \$50 million in the recent 3 years. Increasing pressure to move more

freight over the same tracks is likely to cause more wear and increase the burden of frequent

T&R inspections. These factors suggest that increasing the frequency of track inspections with

onboard MTS devices could help to break the plateau in the accident trend without reducing

revenue service capacity.

The numerical solution to the normalized BCA shows that the SPB as a factor of the average

annual benefits from T&R accident prevention best fits a natural log function of payback period

in years. The R^2 values summarized in Table 1 for all scenarios are greater than 98%, thus

showing that the natural log function is a good fit. In general, the SPB increases with the desired

payback period, and the rate of increase slows with an increase in the annual maintenance cost

proportion of the SPB. The increase for all scenarios is nonlinear, with a more pronounced

increase within the first five years.

The optimization separately derived each function for different values of discount rate, r,

and maintenance cost as a proportion of the SPB, η. The FRA recommended a discount rate of

$r = 7\%$, so it is a fixed parameter that reduces the complexity of estimating α and β as a function

of λ. Practically, discrete models for scenarios of $\lambda = \{0.05, 0.15, 0.25\}$ will supply enough

information for decision making because other intermediate values can be interpolated visually

from the trends. The BCA solution shows that for each scenario, a higher value of λ decreases the sensitivity of the SPB to the desired payback period after the first five years. By extension, selecting a payback period within the first five years would produce the largest impact on DPB.

The average train traffic is a gross approximation by the ratio of annual train miles to track miles that the railroad operated. This estimate reflects average daily train traffic across the entire network. However, the distribution of traffic varies across segments of the network. The STB report lists train and locomotive miles separately for track categories A, B, C, D, and E. Railroads define the track categories according to thresholds of freight density that they supported. Therefore, the sensitivity of problem detection probability, P_n, to first-scan probability, P_1, will vary according to Figure 4, for different track segments. That is, an increase in P_n with respect to P_1 will be more pronounced for those track segments with higher traffic and vice versa.

To encourage adoption, manufacturers can reduce some of the upfront capital required for system deployment by instead recovering installation costs through higher annual maintenance fees. Taking that concept to the extreme, manufacturers may elect to adopt a subscription model that moves the initial system cost into an annual maintenance fee. In that case, the model must change to reflect that the railroad will accumulate a net benefit G as a function of deployment years Y such that

$$G(Y) = \sum_{i=1}^{Y} \frac{B_i - \lambda B_i}{(1+r)^i} \tag{17}$$

and there will be no initial capital. Instead, the manufacture will seek an ROI R as a function of the deployment years Y such that

$$R(Y) = \sum_{i=1}^{Y} \frac{\lambda B_i}{(1 + r)^i} = C_\mathrm{T} = \eta B_i \tag{18}$$

and a numerical solution can follow as before to complete the DPB model.

The sensitivity analysis for all railroads analyzed shows that a lower value of P_1 decreases the sensitivity of the DPB to the desired payback period after the first five years. Similarly, a higher maintenance cost proportion, λ, of the SPB decreases the sensitivity of the DPB to the desired payback period after the first five years. Hence, the device manufacturer can balance the maintenance cost proportion with an achievable DPB based on a desired payback period and the anticipated P_1 performance achievable.

The regression analysis shows that the DPB is predictable as a linear function of the accident cost per equipped train, given a desired payback period and maintenance cost proportion of SPB. If a larger DPB is necessary to engineer a feasible product, then the trade-off will be one or more combinations of increasing the desired payback period, decreasing the annual maintenance cost proportion, and increasing the first-scan probability. The DPB will be lower for railroads that experienced few accidents per train. Therefore, initially targeting railroads with a high number of accidents per train will seed the marketplace and allow for subsequent economies of scale to influence further cost reduction and greater market penetration.

One limitation of using the FRA accident data is that the benefits realized per railroad does not include externalities such as improved inspection efficiency and reduced track closures. Another limitation is that the FRA dataset may not list accidents with damages below $10,500 because railroads need not report those. The reported costs include the loss and/or repair of cars and locomotives, the repair of signal systems and other structures, and the repair of roadbed and track. The reported financial losses do not include those associated with clean up, lost freight,

societal damages, fatalities, injuries, and line closures. Therefore, the DPB for the railroads analyzed may be a conservative estimate.

6. Conclusions

Connected trains are part of the Smart Cities and associated Internet-of-Things (IoT) movements that seek benefits in operational safety and efficiency through intelligent sensing and data-driven decision making. The recent deployment of positive train control (PTC) enables trains to communicate their operational situation in real-time and prevent accidents due to human errors. However, the system cannot help to prevent accidents from other causes such as track and roadbed (T&R) problems.

During the past decade, researchers have focused on the viability of placing equipment onboard revenue service trains to detect T&R problems. However, railroads have not adopted the approach because of uncertainties in the achievable performance and the return-on-investment. This research developed a price budget model to guide the development of viable and affordable multi-sensor track scanning (MTS) systems. Manufacturers can use the model to determine the engineering feasibility of MTS systems with adequate performance, within a specific price budget. Railroads can use the model to determine affordability and a payback period based on the anticipated level of performance. The model is based on deriving a closed-form model from a set of point-based benefit-cost analyses with a standard discount rate and several levels of annual maintenance cost. A case study for five Class 1 railroads found that the minimum device price budget for early adopters would be $4,000. The scenario for this price budget is a 2-year payback period and an annual maintenance cost of one-quarter the initial system deployment cost.

Readers can use the same approach to create closed-form models that can help determine device price budgets for many other applications where users will realize financial benefits from

deploying some quantity of devices. The model helps both the device manufacturer and the first adopters to evaluate the tradeoff price for payback period as a function of performance expectations. Future work will develop additional closed-form models to guide technology deployments in other forms of intelligent transportation systems.

REFERENCES

[1]   Balouchi, Farouk, Adam Bevan, and Roy Formston. 2020. "Development of railway track condition monitoring from multi-train in-service vehicles." *Vehicle System Dynamics* 1-21. doi:10.1080/00423114.2020.1755045.

[2]   Bernal, Esteban, Maksym Spiryagin, and Colin Cole. 2018. "Onboard Condition Monitoring Sensors, Systems and Techniques for Freight Railway Vehicles: A Review." *IEEE Sensors Journal* 19 (1): 4-24. doi:10.1109/JSEN.2018.2875160.

[3]   BNSF. 2019. *Class 1 Railroad Annual Report to the Surface Transportation Board for the Year Ending December 31, 2019.* Forth Worth, Texas: BNSF Railway Company, 119.

[4]   Brezulianu, Adrian, Cristian Aghion, Marius Hagan, Oana Geman, Iuliana Chiuchisan, Alexandra-Ligia Balan, Doru-Gabriel Balan, and Valentina Emilia Balas. 2020. "Active Control Parameters Monitoring for Freight Trains, Using Wireless Sensor Network Platform and and Internet of Things." *Processes* 8 (6): 639. doi:10.3390/pr8060639.

[5]   Bridgelall, Raj, and Denver D. Tolliver. 2020. "Closed form models to assess railroad technology investments." *Transportation Planning and Technology* 43 (7): 639-650. doi:10.1080/03081060.2020.1805541.

[6]   Bridgelall, Raj, and Denver Tolliver. 2020. "Accuracy Enhancement of Anomaly Localization with Participatory Sensing Vehicles." *Sensors* 20 (2): 409. doi:10.3390/s20020409.

[7]   Brooks, James D., Hannah Groshong, Andrew Liu, Paul Houpt, and Chuck Oman. 2017. "Survey of Future Railroad Operations and the Role of Automation." *Transportation Research Record* 2608 (2608): 10-18. doi:10.3141/2608-02.

[8]   BTS. 2020. *Freight Facts and Figures.* October 13. Accessed October 13, 2020. https://www.bts.gov/product/freight-facts-and-figures.

[9]   Bukhsh, Zahaarah Allah, Aaqib Saeed, Irina Stipanovic, and Andre G. Doree. 2019. "Predictive maintenance using tree-based classification techniques: A case of railway switches." *Transportation Research Part C-emerging Technologies* 101: 35-54. doi:10.1016/J.TRC.2019.02.001.

[10]   Cárdenas-Gallo, Iván, Carlos A. Sarmiento, Gilberto A. Morales, Manuel A. Bolivar, and Raha Akhavan-Tabatabaei. 2017. "An Ensemble Classifier to Predict Track Geometry Degradation." *Reliability Engineering & System Safety* 161: 53-60. doi:10.1016/j.ress.2016.12.012.

[11]   Chia, Leonard, Bhavana Bhardwaj, Pan Lu, and Raj Bridgelall. 2018. "Railroad Track Condition Monitoring Using Inertial Sensors and Digital Signal Processing: A Review." *IEEE Sensors Journal* 19 (1): 25-33. doi:10.1109/JSEN.2018.2875600.

[12]   CSX. 2019. *Class 1 Railroad Annual Report to the Surface Transportation Board for the Year Ending December 31, 2019.* Jacksonville, Florida: CSX Transportation, 123.

[13]   Firlik, Bartosz, and Maciej Tabaszewski. 2020. "Monitoring of the technical condition of tracks based on machine learning." *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 234 (7): 702-708. doi:10.1177/0954409719866368.

[14]   FRA. 2020. *Accident Data as reported by Railroads.* Federal Railroad Administration (FRA) Office of Safety Analysis. July 31. Accessed October 11, 2020. https://safetydata.fra.dot.gov/officeofsafety/publicsite/on_the_fly_download.aspx.

[15]   FRA. 2016. *Benefit-Cost Analysis Guidance for Rail Projects.* Washington, D.C.: Federal Railroad Administration, U.S. Department of Transportation.

[16]   Fraga-Lamas, Paula, Tiago M. Fernández-Caramés, and Luis Castedo. 2017. "Towards the Internet of Smart Trains: A Review on Industrial IOT-Connected Railways." *Sensors* 17 (6): 1457. doi:10.3390/s17061457.

[17]   Friebel, Guido, Gerard McCullough, and Laura Padilla Angulo. 2014. "Patterns of Restructuring The US Class 1 Railroads from 1984 to 2004." *Journal of Transport Economics and Policy* 48 (1): 115-135.

[18]   Hodge, Victoria J., Simon O'Keefe, Michael Weeks, and Anthony Moulds. 2014. "Wireless Sensor Networks for Condition Monitoring in the Railway Industry: A Survey." *IEEE Transactions on Intelligent Transportation Systems* 16 (3): 1088-1106. doi:10.1109/TITS.2014.2366512.

[19]   Karimpour, Mostafa, Lalith Hitihamillage, Najwa Elkhoury, Sara Moridpour, and Reyhaneh Hesami. 2018. "Fuzzy Approach in Rail Track Degradation Prediction." *Journal of Advanced Transportation* 2018. doi:10.1155/2018/3096190.

[20]   KCS. 2019. *Annual Report of The Kansas City Southern Railway Company to the Surface Transportation Board for the Year Ended December 31, 2019.* Kansas City, Missouri: The Kansas City Southern Railway Company, 111.

[21]   Lasisi, Ahmed, and Nii Attoh-Okine. 2019. "Machine Learning Ensembles and Rail Defects Prediction: Multilayer Stacking Methodology." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 5 (4): 4019016. doi:10.1061/AJRUA6.0001024.

[22]   Lee, Jun Seok, Sunghoon Choi, Sang-Soo Kim, Choonsoo Park, and Young Guk Kim. 2012. "A Mixed Filtering Approach for Track Condition Monitoring Using Accelerometers on the Axle Box and Bogie." *IEEE Transactions on Instrumentation and Measurement* 61 (3): 749-758. doi:10.1109/TIM.2011.2170377.

[23]   Li, Chunsheng, Shihui Luo, Colin Cole, and Maksym Spiryagin. 2017. "An Overview: Modern Techniques for Railway Vehicle On-Board Health Monitoring Systems." *Vehicle System Dynamics* 55 (7): 1045-1070. doi:10.1080/00423114.2017.1296963.

[24]   Mori, Hirotaka, Yasuhiro Sato, Hiroyuki Ohno, Hitoshi Tsunashima, and Yoshihisa Saito. 2013. "Development of Compact Size Onboard Device for Condition Monitoring of Railway Tracks." *Journal of Mechanical Systems for Transportation and Logistics* 6 (2): 142-149. doi:10.1299/jmtl.6.142.

[25]   NS. 2019. *Annual Report of Norfolk Southern Combined Railroad Subsidiaries ("NS Rail") to the Surface Transportation Board for the Year Ended December 31, 2019.* Norfolk, Virginia: Norfolk Southern Combined Railroad Subsidiaries, 123.

[26]   Paixão, André, Eduardo Fortunato, and Rui Calçada. 2019. "Smartphone's Sensing Capabilities for On-Board Railway Track Monitoring: Structural Performance and Geometrical Degradation Assessment." *Advances in Civil Engineering* 2019: 1-13. doi:10.1155/2019/1729153.

[27]   Peng, Fan, and Yanfeng Ouyang. 2012. "Track Maintenance Production Team Scheduling in Railroad Networks." *Transportation Research Part B: Methodological* 46 (10): 1474-1488. doi:10.1016/j.trb.2012.07.004.

[28]   Rahimikelarijani, Behnam, Ahmad Mohassel, and Maryam Hamidi. 2020. "Railroad Track Geometric Degradation Analysis: A BNSF Case Study." *Journal of Transportation Engineering, Part A: Systems* 146 (2): 4019068. doi:10.1061/JTEPBS.0000303.

[29]   Sun, Yongkui, Yuan Cao, Guo Xie, and Tao Wen. 2020. "Condition Monitoring for Railway Point Machines Based on Sound Analysis and Support Vector Machine." *Chinese Journal of Electronics* 29 (4): 786-792. doi:10.1016/J.TRC.2019.02.001.

[30]   Sysyn, Mykola P., Ulf Gerber, Olga S. Nabochenko, Dmitri Gruen, and Franziska Kluge. 2019. "Prediction of Rail Contact Fatigue on Crossings Using Image Processing and Machine Learning Methods." *Urban Rail Transit* 5 (2): 123-132. doi:10.1007/S40864-019-0105-0.

[31]   Tsunashima, Hitoshi. 2019. "Condition Monitoring of Railway Tracks from Car-Body Vibration Using a Machine Learning Technique." *Applied Sciences* 9 (13): 2734. doi:10.3390/app9132734.

[32]   UP. 2019. *Class 1 Railroad Annual Report R-1 to the Surface Transportation Board for the Year Ending December 31, 2019.* Omaha, Nebraska: Union Pacific Railroad, 119.

[33]   Weston, Paul, Clive Roberts, Graeme Yeo, and Edward Stewart. 2015. "Perspectives on Railway Track Geometry Condition Monitoring from In-Service Railway Vehicles." *Vehicle System Dynamics* 53 (7). doi:1063-1091. 10.1080/00423114.2015.1034730.
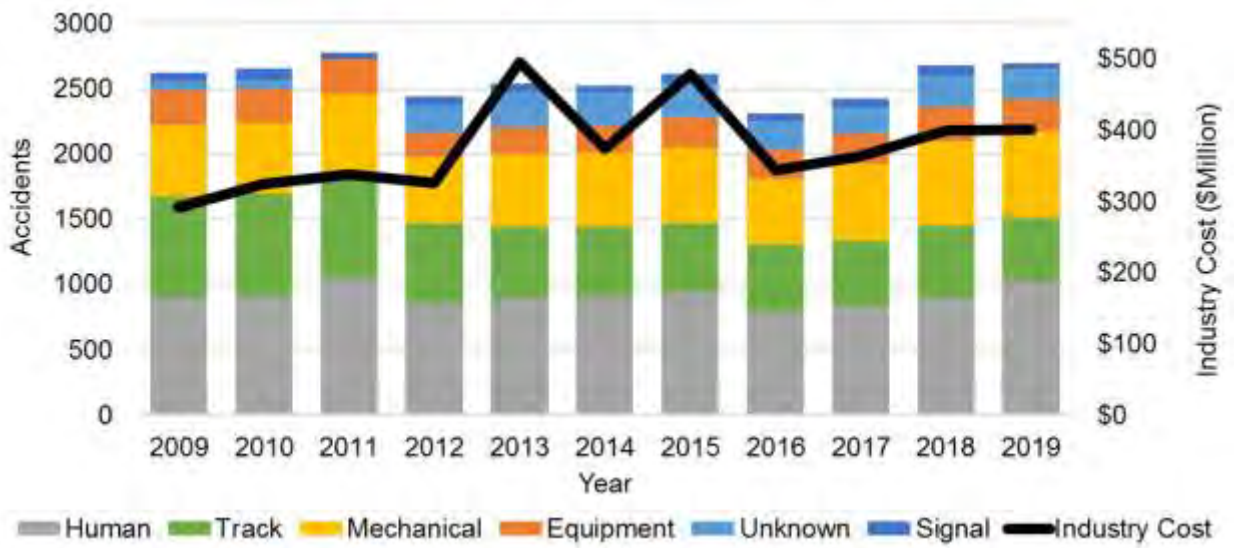
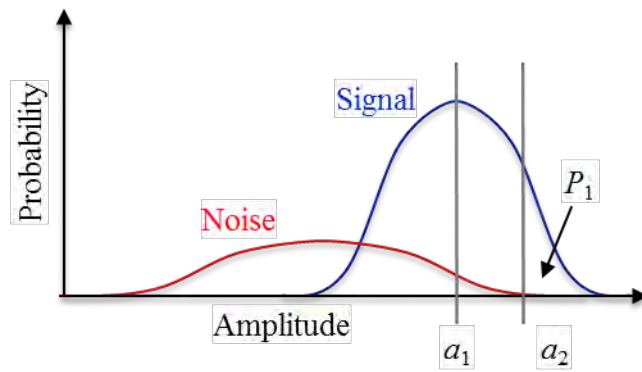Figure 1. Railroad accident statistics.

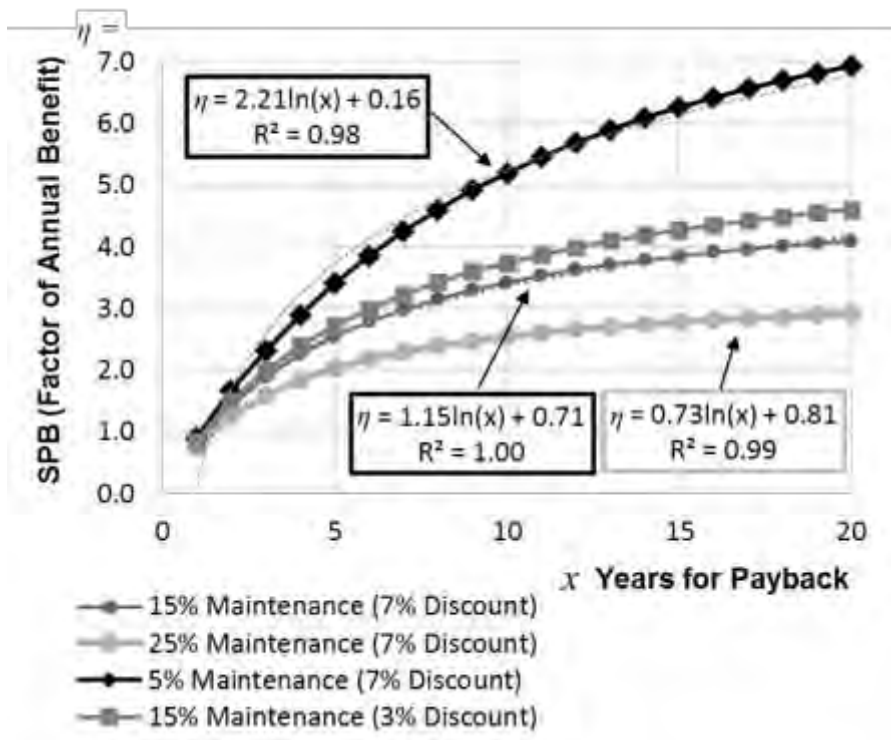Figure 2. Tradeoff between false positive and false negative rates.

Figure 3. SPB factor of annual savings as a function of payback period.
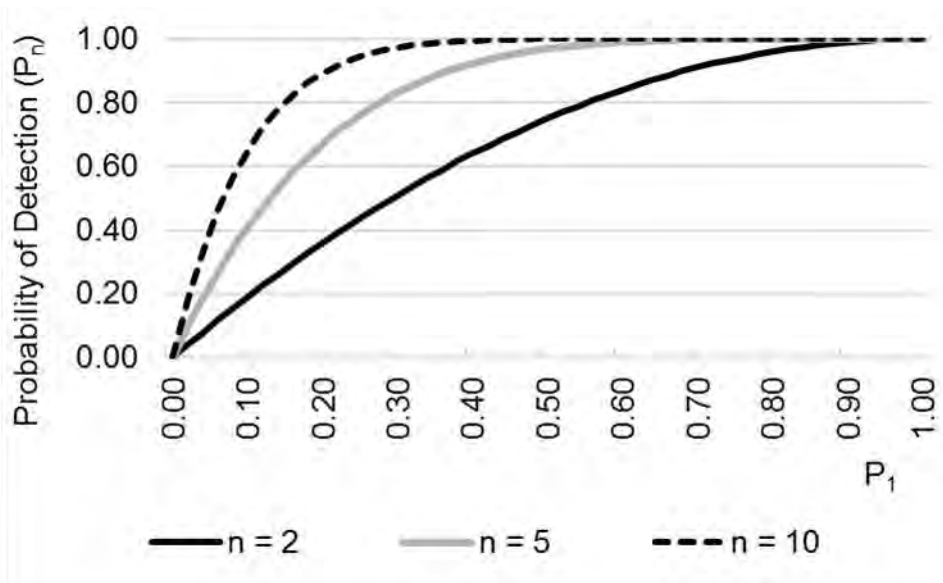
Figure 4. Probability of detection as a function of first-scan probability and number of scans.
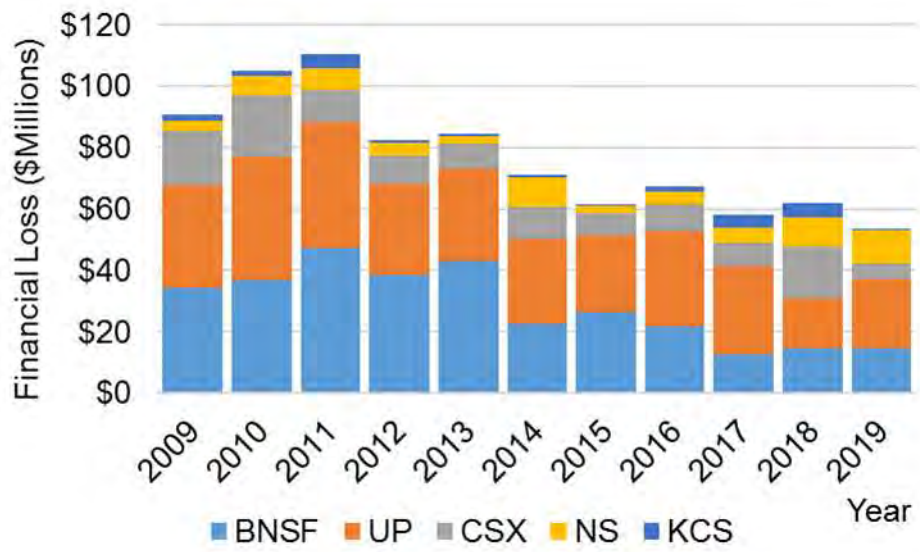
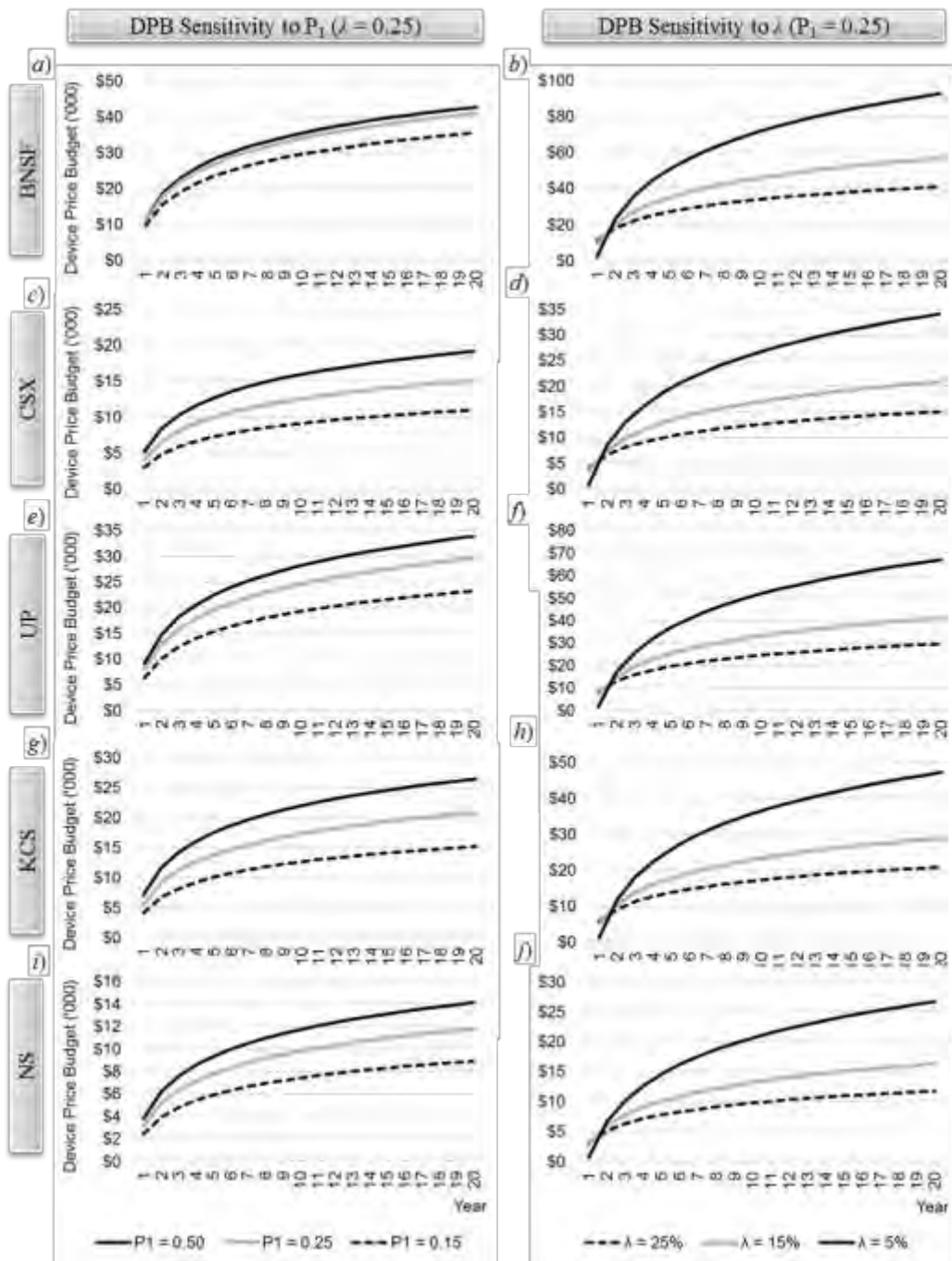Figure 5. Financial loss from accidents due to T&R problems, by railroad.

Figure 6. Railroad specific sensitivity analysis to first-scan probability and maintenance cost proportion.
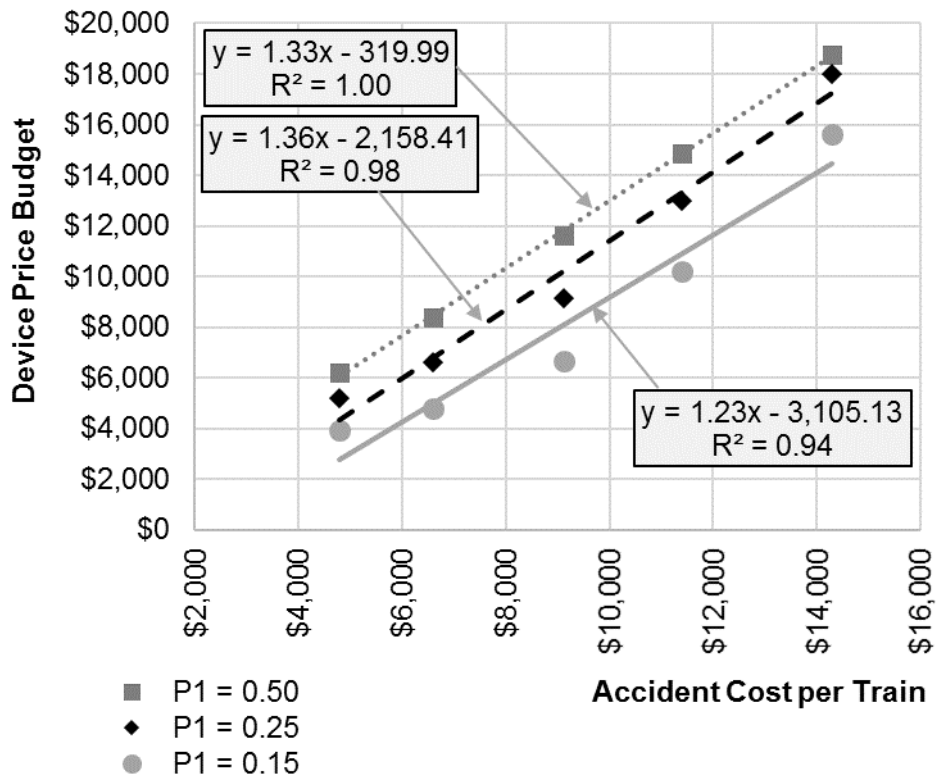
Figure 7. Prediction of DPB by accident cost per train with first-scan probability scenarios ($\lambda = 0.25$).

Table 1. BCA Parameter Estimates by Scenario

Scenario	r	λ	α	β	R^2
1	0.07	0.25	0.73	0.81	0.99
2	0.07	0.15	1.15	0.71	1.00
3	0.07	0.05	2.21	0.16	0.98

Table 2. Data Mined for Select Class 1 Railroads

Parameter	BNSF	CSX	UP	KCS	NS
Data Source	(BNSF 2019)	(CSX 2019)	(UP 2019)	(KCS 2019)	(NS 2019)
T&R Accident (mean)	$28,471,572	$10,962,435	$29,665,807	$1,906,782	$6,003,679
Locomotives Operated	6,578	3,440	7,969	535	2,812
Track Miles Operated	39,858	28,949	43,625	4,084	28,422
Train Miles	161,392,821	61,671,368	115,212,273	8,654,805	70,588,362
Locomotive Miles	532,407,992	127,435,787	352,728,786	22,083,995	158,376,020
Locomotives/Train	3.3	2.1	3.1	2.6	2.2
Train Trips	4,049.2	2,130.3	2,641.0	2,119.2	2,483.6
Trains/Day	11.1	5.8	7.2	5.8	6.8
Accident Cost/Train	$14,279	$6,588	$11,401	$9,123	$4,791

Table 3. SPB, DPB, and Annual Maintenance Cost for Individual Class 1 Railroads

Parameter	BNSF	CSX	UP	KCS	NS
Sensors Installed, N_S	1,984	1,664	2,602	209	1,253
Detect Proportion, P_n	1.000	0.969	0.992	0.969	0.984
DPB, C_D	$18,781	$8,399	$14,887	$11,631	$6,207
SPB, C_T	$37,450,294	$13,975,734	$38,735,200	$2,430,909	$7,777,392
Annual Maintenance	$5,617,544	$2,096,360	$5,810,280	$364,636	$1,166,609