University of Illinois at Urbana-Champaign (UIUC)

Rail Transportation and Engineering Center (RailTEC)

RailTEC Faculty and Student Papers and Presentations IHHA 2017 Cape Town 2-6 September 2017

RAILTEC
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

University of Illinois at Urbana-Champaign
Rail Transportation and Engineering Center (RailTEC)

Papers Presented at the 11th International Heavy Haul Association (IHHA) Conference
Cape Town, South Africa
2 – 6 September 2017

MONDAY - 4 September

TUESDAY - 5 September

WEDNESDAY - 6 September

Lateral load performance of concrete sleeper fastening systems under non-ideal conditions

Brevel G. J. Holder, Yu Qian, Marcus S. Dersch & J. Riley Edwards
University of Illinois Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT: The fastening system is an essential component of the track superstructure that facilitates load transfer from the rail to the sleeper while holding the rail in place. Previous research has focused on investigating the performance of different fastening systems under laboratory and field loading environments when the fastening systems are properly installed. However, with the increased traffic and challenging service environments often experienced in North America, it is likely not all the fastening systems can remain intact throughout their service life of the track, thus missing fastening system components can occur. To date, the performance of different fastening systems under non-ideal conditions, such as track with missing fastening system components, has not been thoroughly investigated. In order to better understand the behaviour of concrete sleeper fastening systems under different non-ideal loading conditions similar to what is seen in the field, an on-going research project is currently being conducted at the University of Illinois at Urbana-Champaign. This paper presents the preliminary laboratory results of the lateral load performance of the Skl-style fastening system on track with missing fastening system components at one or more sleepers. Lateral load redistribution was quantified for different test scenarios. The results from this study will improve the understanding of lateral load distribution under non-ideal conditions and can be used in future fastening system design and field maintenance practices.

1 INTRODUCTION

The fastening system is an essential component of track superstructure and facilitates load transfer from the rail to sleeper while holding the rail in place. Previous research has studied the lateral load distribution when the fastening systems are properly installed (Holder et al. 2017). However, missing fastening system components on one or more adjacent sleepers can occur in the field, and the performance of track under these non-ideal conditions has not been investigated. To better understand the performance of the fastening system under non-ideal conditions, researchers at the University of Illinois at Urbana-Champaign (UIUC) are conducting an experimental study to investigate the magnitude and distribution of the lateral load through the track superstructure when a portion of the fastening system components are missing. The performance of the fastening system after one or more reinstallations (i.e. clamps and angled guide plates that have been removed and reapplied for multiple times) is also being studied to better understand the loss of clamping force over the service life of the track and how it affects the lateral load distribution and rail rotation. It is the expectation of the authors that the information in this paper will assist the rail industry in improving fastening system design, performance, and maintenance for heavy-haul freight railroad applications through the use of quantitative loading data as inputs for future practice. The primary objectives of this project are to quantify the clamping forces of the Skl-style fastening system involved in the removal and reinstallation investigation, as well as to gain a better understanding of the lateral loads distribution under non-ideal conditions when fastening system components are removed.

2 EXPERIMENTATION PLAN

2.1 *Fastening System and Concrete Sleepers*

Experiments were performed using concrete sleepers equipped with Skl-style fastening systems. The concrete sleepers have dimensions of 2,590 mm long, 279.4 mm wide, and 222.3 mm high. The Skl-style fastening system is comprised of five major components that ensure the longevity of their performance. These components are the tension clamps, angled guide plates, lag screw/dowel, abrasion plate and rail pad (see Figure 1). As described by Van Dyk et al.

(2015), the tension clamps are designed to have high fatigue strength, which allows them to maintain their clamping ability over extended periods of time. The field and gauge angled guide plate transfer the lateral force experienced by the rail to the concrete sleeper. The lag screw and dowel combination hold the tension clamp to the sleeper and help decrease the transverse stress on the concrete sleeper

The rail pad is designed to provide appropriate resilience while the abrasion plate is a protection layer between the rail pad and the rail seat. Additionally, the rail pad is designed to withstand the high pressures that are associated with heavy haul trains, and the abrasion pad is a critical part in mitigating Rail Seat Deterioration (RSD), a major problem of the concrete sleeper in North America (Greve et al. 2015).

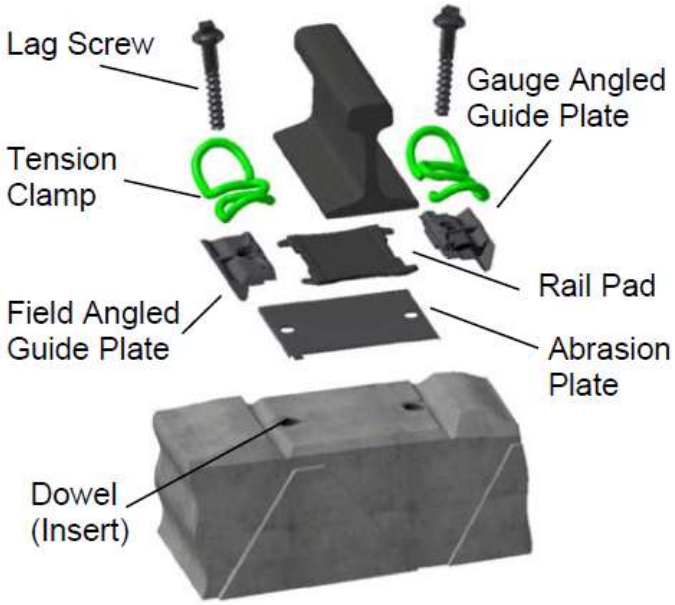


Figure 1. Vossloh Fastening Systems, Inc. W 40 (Van Dyk et al. 2015)

2.2 *Lateral Load Path*

To better understand the magnitude and distribution of lateral loads through the track superstructure, understanding the load path through the Skl-style fastening system is important. The load path has not been well defined as of yet, however, based on the results from the past experiments conducted at the University of Illinois, a hypothetical lateral load path is presented in Figure 2 (Williams et al. 2014). When a lateral load is applied to the head of the rail, it is hypothesized that the load is transferred to the base of the rail and is primarily resisted by the field side angled guide plate of the fastening system. For this reason, field side guide plates are often designed to be larger than the gauge side guide plates decrease the compressive stress on the concrete sleeper. UIUC researchers also believe that a small portion of the lateral load applied to the rail could possibly be transferred through the tension clamp into the sleeper as well as through frictional forces between the rail pad – rail seat interface.

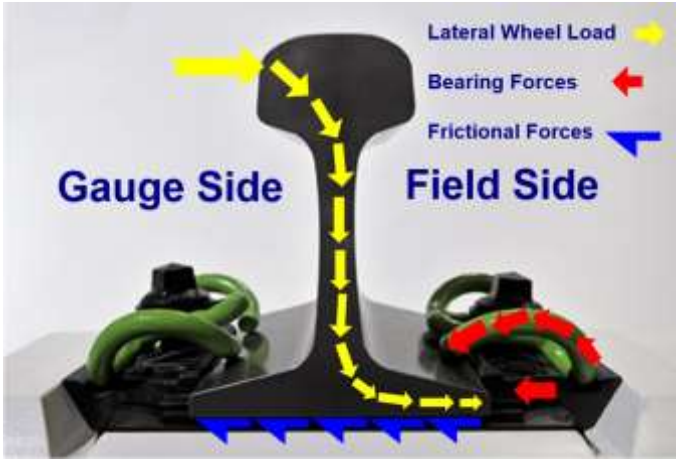


Figure 2. Hyperthetic lateral load path

2.3 *Measuring Lateral Forces*

In order to quantify the lateral loads that are applied to the field side angled guide plate, a device called the Lateral Load Evaluation Device (LLED) has recently been developed at the University of Illinois. Strain gauge bridges (applied to the top and bottom sections of the LLED, Figure 3) are used to measure the bending strain of the instrument, which in turn is used to calculate the lateral force experienced by LLEDs with pre-developed calibration curves. Two LLEDs are installed on each of the modified field side angled guide plates to capture any possible uneven loading conditions. The modified field side angled guide plates with LLEDs are also designed to have a similar stiffness of the original guide plate.

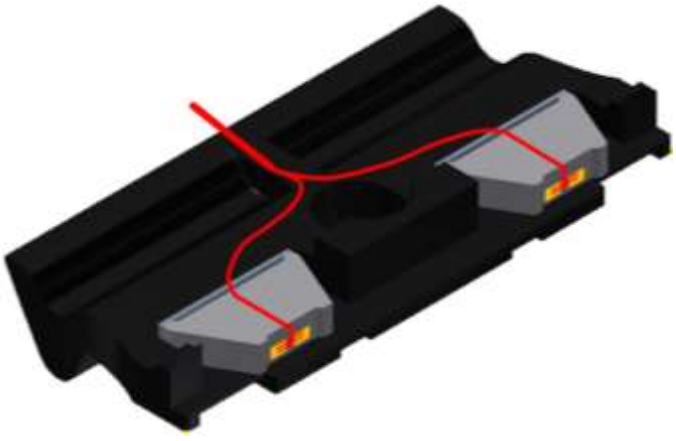


Figure 3. LLED placed in a angled guide plate

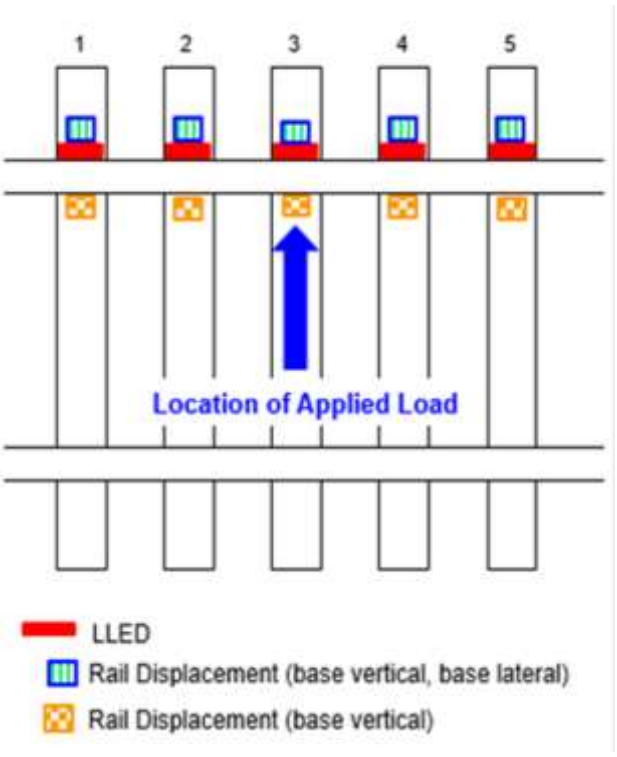
2.4 *Laboratory Setup*

Previous studies have shown that the lateral load primarily is distributed into three sleepers with elastic fastening systems, thus a total of five sleepers were used in the laboratory test setup to be conservative. Figure 4 (a) shows a picture of the test setup and Figure 4 (b) gives the schematic drawing. The

five sleepers were secured to the floor with space of 609.6 mm (24 in.) between each other. A section of 136 RE rail with the length of 2590 mm (102 in.) was installed on one side of the sleepers with the Skl-style fastening systems properly tightened according to the supplier's guidance. Other than the aforementioned two LLEDs installed in the field side guide plate, customized aluminium brackets were installed in each sleeper to measure rail movement. The gauge and field brackets were equipped with vertical potentiometers to measure rail rotation, while two other potentiometers were installed laterally on the field side brackets to measure rail base displacement. A hydraulic jack aligned with the center of the middle sleeper was used to apply lateral load from rail head. All loads were applied at this location and were controlled manually using a hand pump. Due to the limitation on length, this paper will present the results LLEDs only. Rail deflection and rotation will be discussed in future publications.

the typical case for a new installation. Figure 5 shows the percentage of lateral wheel load resisted by the field angled guide plate on the y-axis and its corresponding sleeper on the x-axis. The percentage of lateral wheel load resisted by the field angled guide plate is calculated by dividing the summation of the lateral load measured by the two LLEDs installed in each field angled guide plate by the lateral load applied from the hydraulic jack. The summation of the percentages may not be 100% due to the fact that not all the applied lateral force was transferred into the field angled guide plate. A portion of the applied lateral load is assumed to be transferred into the sleeper through the tension clamps, and the friction between the rail and fastening system components. Each of the five trials have been plotted in Figure 5 to show the repeatability of this tests. Mean value and the standard deviation are also provided in Table 1. It is redundant to plot all the repeated test results like Figure 5 for all the types of tests. Instead, the standard deviation between all five trials for each

(a) Laboratory test setup



(b) Schematic drawing of the test setup

Figure 4. Laboratory test setup

3 PRILIMINARY RESULTS

3.1 *Repeatability of test*

In order to ensure the accuracy and reliability of the data, all experiments were repeated five times. For the five repeated trials, the test setup remains intact without disassembling any components. Figure 5 gives an example of the results from the five repeated tests under the fully fastened condition. The fully fastened test represents how the fastening system performs when there is no missing component, as is

sleeper under different test scenarios are provided in Table 1. It should be noted that the largest standard deviation manifested between five trails is only 3.07; therefore, suggesting that the tests performed are both replicable and accurate. All the graphs presented in this paper later are the averaged results from the five trials for each type of tests.

Table 1. Mean value and standard deviation of lateral load percenage for all test

Sleeper	1		2		3		4		5	
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
Fully Fastened	1%	0.22	8%	0.29	39%	0.13	27%	0.40	3%	0.35
Missing Center Clamps	2%	0.23	16%	0.57	33%	0.13	26%	0.28	7%	1.71
Missing Center Plates	0%	0.07	22%	0.13	0%	0.02	39%	0.79	49%	0.73
Missing Adjacent Clamps	7%	0.14	23%	0.66	45%	0.46	34%	3.07	0%	0.00
Missing Adjacent Plates	13%	0.51	0%	0.02	62%	0.03	0%	1.59	0%	0.00
5 Replacements	0%	0.00	20%	0.14	34%	0.07	26%	0.54	0%	0.00
10 Replacements	0%	0.00	20%	0.14	34%	0.07	26%	0.54	0%	0.00
15 Replacements	0%	0.00	22%	0.19	34%	0.04	28%	0.23	0%	0.00

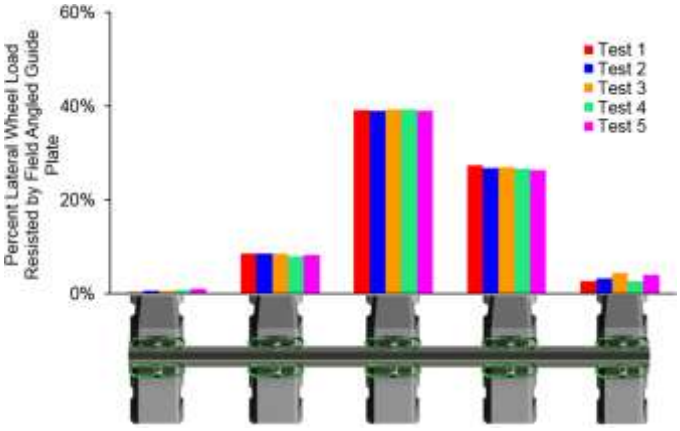


Figure 5. Fully fastened condition

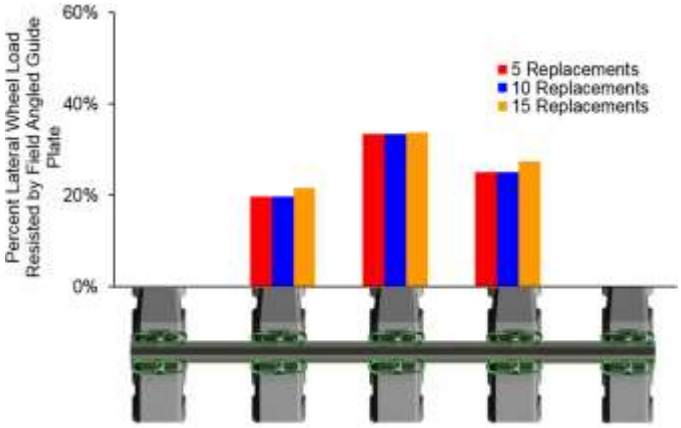


Figure 6. Loss of clamping force

3.2 *Reusability of component*

To investigate the role that lateral forces have on the Skl-style tension clamps under different test scenarios, components such as tension clamp and angled guide plate need to be removed and reinstalled several times. However, the reusability of those components has not been thoroughly investigated in the literature. In this study, the reusability of the components for the skl-style fastening system was investigated. Experiments were performed by completely removing each fastening systems and reinstalling them. After the first five removals and installations were completed, lateral force was applied to the rail and the lateral load distribution was recorded. Load was applied again after an additional 5 removals and reinstallations.

Figure 6 presents all three-replacement tests results. Mean values and standard deviations are also provided in Table 1. It is clear that the differences between each of the three tests are minor, especially when the five and ten replacement tests are compared. The test results shown in Figure 6 and Table 1 indicate the components for the Skl-style fastening system can be reused for at least 15 times under the loading magnitude in this study without significant loss of clamping forces.

3.3 *Missing Center Clamps*

As briefly discussed previously, properly installed fastening systems with no missing components can represent new installation scenario. However, with the accumulation of tonnage during service life, fastening systems with no missing components are not the only scenario found in the field. It is possible for some components to experience failure caused by fatigue, fracture, or crushing which can lead to missing fastening system components on the track superstructure. Thus the performance of the fastening system under non-ideal conditions is worth being investigated. Various possible non-ideal conditions were simulated and tested as listed in Table 1.

The first test performed is the “missing center clamps” test. For this test, the tension clamps on both sides of the center sleeper were removed and the angled guide plates were left with screws tightened on the concrete sleeper. Figure 7 shows the percent lateral load for each sleepers under the “missing center clamps” scenario. On that same graph, the averaged fully fastened test results are also provided for comparison. Figure 7 shows that once the clamps are removed, the percent lateral load resisted by the center sleeper was reduced from 39% to 33%, the percent lateral load resisted by one of adjacent sleeper (sleeper 4) also reduced by a

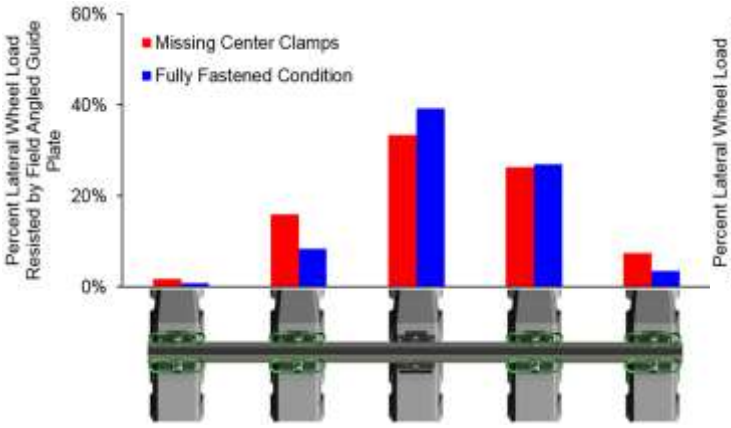


Figure 7. Missing center clamps

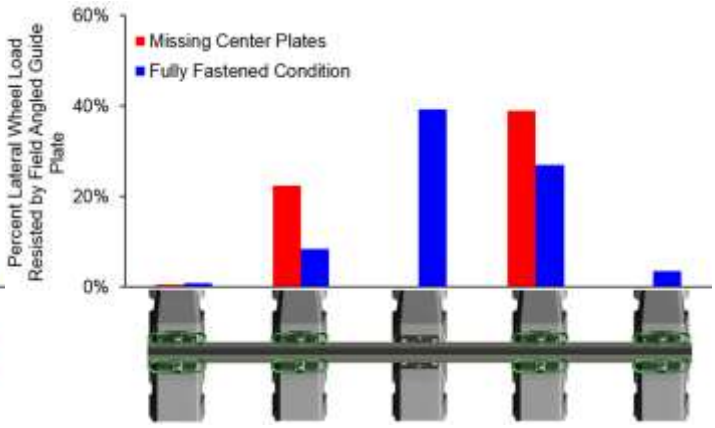


Figure 8. Missing center plates

small portion, from 27% to 26%. However, the percent lateral load resisted by sleeper 1, 2, and 5 all increased, from 1% to 2%, 8% to 16%, and 3% to 7%, respectively, which means the lateral forces become more evenly distributed for the missing center clamps scenario than the fully fastened scenario. The change of lateral force distribution may be caused by the change of lateral stiffness at the center when the clamps were removed. However, the majority of the lateral load was still resisted by the middle three sleepers. It was observed that there was more rail rotation at the center sleeper, which will be discussed in future publications.

3.4 Missing Center Clamps and Plates

The second test performed was for missing center sleeper tension clamps and angled guide plates. For this test the tension clamps and angled guide plates were completely removed for the center sleeper. Since there are no angled guide plates in either the field or gauge side of the sleeper, there will be no lateral load measurement at this location, but lateral resistance from the friction is still possible. Figure 8 shows the percent lateral load for each sleepers under the missing center clamps and plates scenario. After removing the center clamps and plates, the percent lateral load clearly transferred into the adjacent sleepers. The percent lateral load for sleepers 2 and 4 increased from 8% and 27% to 22% and 39%, respectively, while the values for sleeper 1 remained similar. It is interesting to see that there is no value of percent lateral load for sleeper 5 after removing the center clamps and plates in Figure 8. This may be due the significant reduction in lateral restraint at the center sleeper which allowed the rail more freedom to twist or move, causing the field side angled plate of sleeper 5 to lose contact with the rail as load was being applied. This scenario may also happen in the field considering the longitudinal stiffness of rail is relatively low.

3.5 *Missing Adjacent Clamps*

In order to help this study to cover a wide range of possibilities, it was important not only to include the non-ideal conditions on the center sleeper but also considering the two adjacent sleepers. Figure 10 illustrates the distribution of lateral loads when the two adjacent sleepers have had their tension clamps removed, representing a possible failure mode for both adjacent sleepers. When two out of five sleepers had missing clamps, the total percent of lateral load resist by angled guide plate increased from 79% to nearly 100%. The reason that the portion of lateral load resisted by the angled guide plate increased was possibly due to the loss of clamping force when the two adjacent clamps were missing. Similar to missing the clamps for the center sleeper only in which case the percent lateral load on the center sleeper reduced (from 39% to 33%, see Figure 7), the percent lateral load for sleeper 3 decreased from 39% to 38%. Sleepers 1, 2 and 4 increased, 1% to 6% for sleeper 1, 8% to 20% for sleeper 2 and 27% to 29% for sleeper 4 (Figure 9). One interesting observation is the percent lateral load of sleeper 5 reduced to 0 after clamps were removed from sleeper 2 and 4, while a considerable percent lateral load increased from 1% to 6% for sleeper 1.

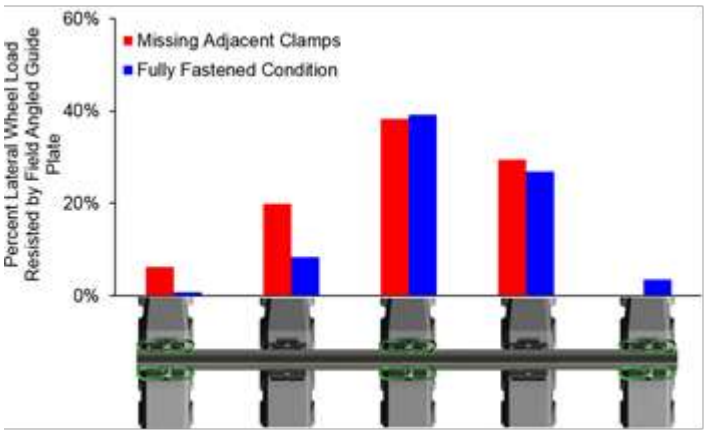


Figure 9. Missing adjacent clamps

One possible reason for this behavior could be that the rail moved during the application of load and one side (sleeper 1 and 2) of the test setup became more engaged. This can also be confirmed by comparing the total percent lateral load changed for sleeper 1 and 2 was 17%, while the values for sleeper 4 and 5 was 1%.

3.6 *Missing Adjacent Clamps and Plates*

Figure 10 illustrates the distribution of lateral loads when the two adjacent sleepers have had both the tension clamps and angled guide plates removed. This test is the most "severe" scenario in this study. With two sleepers lose ability to restraint lateral movement dramatically, the total percent lateral load measured by LLEDs changed from 79% to 89%. The percent lateral load restrained by the angled guide plate of the center sleeper increased from 39% at the fully fastened scenario to 62%, while the value at the missing adjacent clamps was 38%. A similar increase at sleeper 1 was from 1% at the fully fastened scenario to 13%, while the value at the missing adjacent clamps was 6%. This is because when sleeper 2 and 4 significantly lose their ability to sustain lateral load, fastening system at sleeper 1 and 3 became more engaged. Similar as the missing adjacent clamps test, no noticeable lateral load was measured from sleeper 5.

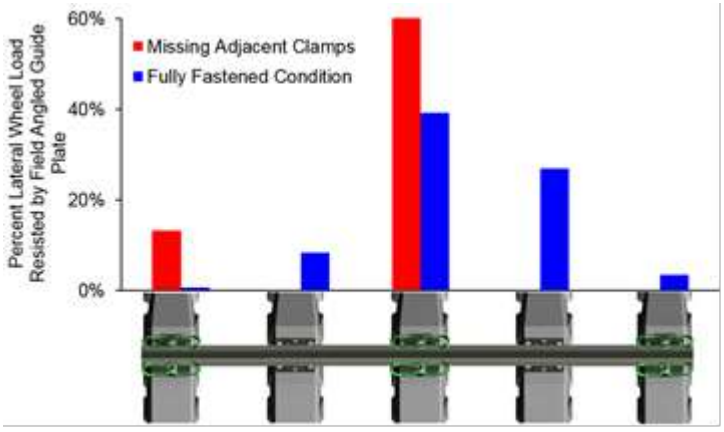


Figure 10. Missing adjacent clamps

4 CONCLUSIONS AND FUTURE WORK

This paper presents an experimental study focused on characterizing lateral load distribution with the Skl-style fastening system under non-ideal conditions. These non-ideal conditions include missing one or more components at one or more sleepers. Lateral load was measured and analysed for each sleeper under different test scenarios. The following conclusions can be drawn based on the preliminary results from this study:

The performance of the Skl-style fastening system was consistent under the specific test conditions in this study. Loss of clamping force during reinstallation for multiple times was not observed. Skl-style fastening system relies more on the angled guide plate to resist lateral load than frictional forces. The majority of the lateral load is distributed into three sleepers under single point lateral load. Missing components will redistribute the lateral load considerably, especially when the angled guide plate is missing. When some component or components are missing, lateral load will mainly be transferred into two adjacent sleepers. However, depending on the initial position of the rail, the redistribution of the lateral load could influence up to five sleepers.

5 ACKNOWLEDGMENTS

This project was partially sponsored by funding from the Federal Transit Administration (FTA), part of the United States Department of Transportation (US DOT). The published material in this report represents the position of the authors and not necessarily that of the DOT. The sleeper and fastening system were provided by Rocla Concrete Tie Inc. and Vossloh North America, respectively. The authors would like to acknowledge Donovan Holder, Brendan Schmit, Alamo DiTarso, Michael Parisotto, Daniel Savio, Jacob Allen, Tim Prunkard and the UIUC machine shop for their help in experimentation setup.

6 REFERENCES

Holder, D.E., Williams, B.A., Dersch, M.S., and Edwards, J.R. 2015. Quantification of lateral Forces in Concrete Sleeper Fastening Systems under Heavy Haul Freight Loads. In: *Proceedings: The 11th International Heavy Haul Association Conference,* Perth, Australia.

Greve, M.J., Dersch M.S., Edwards J.R., Barkan C.P.L., Thompson, H., Sussmann, T., and McHenry M. 2015 Examination of the Effect of Concrete Crosstie Rail Seat Deterioration on Rail Seat Load Distribution. In*: Transportation Research Record – Journal of the Transposrtation*

Van Dyk,B., Bosterling,W., Harrass, M.,Wiethoff, N., and Wroblewski A. 2015. Rail Fastening System Design Considerations to Reduce The Stress Within Haul Track Supersstructure. In: *Proceedings: International Heavy Haul Conference Fastening System Session*, Perth, Australia, Jun 2015, pp. 1308-1315

Williams, B.A., Edwards, J.R., Kernes, R.G., and Barkan, C.P.L. 2014. Analysis of the Lateral Load Path in Concrete Crosstie Fastening Systems. In: *Proceedings: The 2014 Joint Rail Conference*, Colorado Springs, CO, USA,

Influence of the tie-ballast interface on transition zone performance

S.T. Wilk & T.D. Stark
University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

J.G. Rose
University of Kentucky, Lexington, Kentucky, USA

T.R. Sussmann, Jr.
Volpe Center, Washington D.C., USA

H.B. Thompson II
Federal Railroad Administration, Washington D.C., USA

ABSTRACT: Many locations of railroad track consist of abrupt changes in track stiffness and underlying substructure called transition zones. These regions historically experience greater track settlement than the surrounding track requiring more frequent track maintenance from railroad companies. This paper emphasizes how discontinuities at the tie-ballast interface can be indicative of future track geometry problems and further deteriorate transition zone locations. This is supported by non-invasive field instrumentation using video cameras to measure tie displacement, accelerometers to measure tie acceleration, and three-dimensional dynamic numerical modeling. Transition zones with reoccurring track geometry defects often display unsupported ties and field instrumentation and numerical modeling suggest this condition can further increase tie loading and ballast deterioration. Potential solutions to avoid discontinuities at the tie-ballast interface are balancing transient and permanent displacements between the bridge and transition zone, providing a cushioning layer at the tie-ballast interface, and ensuring good substructure support and ballast condition underneath the transition zone ties. Examples of these transition zones are included in the paper.

1 INTRODUCTION

Railroad track transitions are track locations that experience a rapid change in track structure. This often refers to bridge transition zones but can also include asphalt crossings, culverts, and transitions from ballasted to unballasted track. Track transitions are a common topic of study because they often experience accelerated track geometry deterioration and represent an expensive maintenance location for railroads (Li & Davis, 2005; Mishra et al., 2012; Stark & Wilk, 2016).

Multiple studies have investigated root causes of accelerated settlement in the transition approach, i.e. about 3 to 6-m from bridge abutment (Kerr & Bathurst, 2001; Li & Davis, 2005; Plotkin & Davis, 2008; Coelho et al., 2011). Results suggest that while the specific causes of deterioration is site dependent , the three general root causes are: (1) lack of track settlement on the bridge, (2) increased dynamic loads in the approach, and (3) reduced-performance substructure conditions in the approach. Reduced-performance conditions are defined as ballast or subgrade conditions that result in a reduced stiffness or increased settlement rate than what is anticipated from a compacted substructure in the open track, i.e. track location with no adjacent track structure. Examples of reduced-performance ballast conditions are ballast degradation, fouling, and increased moisture. Examples of reduced-performance subgrade conditions are inadequate compaction due to the abutment or increased moisture. Typically, a combination of all three mechanical root causes and construction issues play a role in the accelerated settlement but the interaction and magnitude of each factor can vary between sites.

A common chain of events that lead to transition zone track geometry deterioration is described below. After resurfacing or being put into service, the ballast and subgrade in the approach will compact and densify from repeated train loading, resulting in some magnitude of differential settlement between the bridge and approach. This magnitude depends on the initial ballast and subgrade density and characteristics (Indraratna et al., 2012). The differential top-of-rail (TOR) elevation at the approach-bridge interface produces rail-tie or tie-ballast gaps as the first few ties “hang” from the stiff rail that is supported by the higher elevated bridge. These gaps can redistribute the wheel loads throughout the track system, produce impacts as the tie establishes contact with the ballast, and promote ballast deterioration from tie-ballast abrasion due to unrestricted tie movement. The increased loads and ballast breakdown in the approach can produce a negative feedback loop that requires frequent resurfacing to maintain track geometry.

This paper presents a general overview of a recently completed study investigating the root causes of the differential movement at transition zones and the benefits of various mitigation techniques (Wilk, 2017). This paper briefly covers data collected from long-term instrumentation (Section 2), short-term instrumentation (Section 3), and numerical modelling (Section 4) along with some recommendations of design, remedial, and resurfacing techniques (Section 5). Both high-speed passenger and heavy haul freight are covered but causes and remedial techniques are similar for both cases.

2 LONG-TERM INSTRUMENTATION

This section presents the results of a long-term instrumentation project to monitor a bridge approach transition zone experiencing historical track geometry problems (Mishra et al., 2012; Stark and Wilk, 2016; Wilk, 2017). The site is located at the Upland Street bridge approach on Amtrak's high-speed passenger line (177 km/hr / 110 mph) with instrumentation located 4.5 m (15 ft) and 18.3 m (60 ft) away from the bridge abutment, representing bridge approach and open track locations, respectively. To keep notation with previous publications (Mishra et al., 2012; Stark and Wilk, 2016), the terms Upland (15 ft.) and Upland (60 ft.) will be used herein. The instrumentation consists of strain gauges attached to the rail to measure wheel loads and strings of LVDTs installed at depth to measure the permanent and transient displacement of individual substructure layers, e.g. ballast, subballast, and three subgrade layers. On two trips, an accelerometer was attached to the Upland (15 ft.) tie to measure tie acceleration.

The project objective was to identify the depth of movement, determine root causes of differential movement, and propose and test solutions to prevent and mitigate differential movement. This section presents a summary of the results from the instrumentation.

2.1 *Permanent Ballast Displacements*

The permanent displacements after 446 days of monitoring of the uppermost LVDT (LVDT #1) are displayed in Figure 1. LVDT #1 measures the displacement from the top of the concrete tie to the bottom of the ballast layer and is considered a reasonable estimate of ballast settlement. The majority of overall track settlement occurred in LVDT #1 and not the subballast (LVDT #2) or subgrade (LVDT #3 through #5) so the ballast layer is considered the region of interest for this particular site (Stark and Wilk, 2016). The instrumented site has been in service for almost a century so it is expected that the subgrade has fully compacted, explaining the minimal subgrade movement. This behaviour may not be true if the transition zone is on a recently constructed

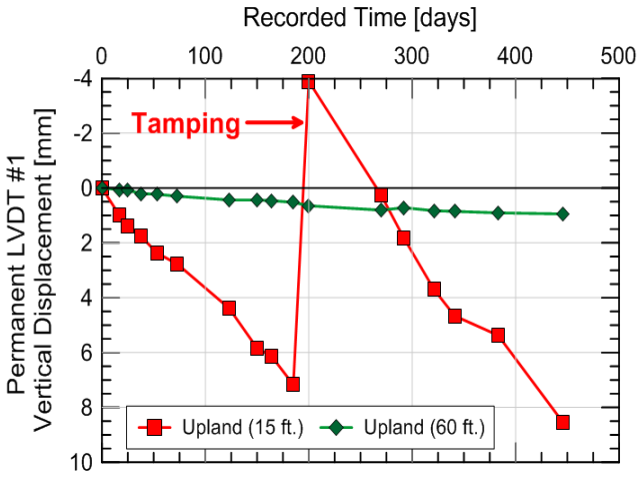


Figure 1. Permanent LVDT #1 vertical displacements of Upland (15 ft.) and Upland (60 ft.)

fill or subgrade that has not previously experienced train loading (Moale et al., 2016).

Comparisons between the approach and open track locations show greater ballast settlement at Upland (15 ft.) with 14 mm/yr than Upland (60 ft.) with 1 mm/yr. The trends show a consistent Upland (15 ft.) settlement rate while the Upland (60 ft.) appears to show decreased settlement with time. This suggests that Upland (60 ft.) has reached a near "equilibrium state" while Upland (15 ft.) has not.

2.2 *Transient Ballast Displacements*

Transient displacements, i.e. displacements from a passing train, were recorded five times during the period of data collection. The results from LVDT #1, measuring the top of concrete tie to bottom of ballast layer, showed non-linear behaviour as the tie must close any tie-ballast gap prior to transferring the load to the ballast. Plotting the peak wheel loads and corresponding LVDT #1 displacements for a passing train allows for an estimation of the tie-ballast gap height ($\delta_{P=0}$) to be calculated. The load-displacement curve of Upland (15 ft.) is displayed in Figure 2.

Once the two LVDT #1 components of tie-ballast gap height ($\delta_{P=0}$) and ballast displacement components (δ_{mob}), i.e. displacement of frictionally mobilized ballast, were separated, the tie-ballast gap (LVDT #1), ballast displacement (LVDT #1), and subgrade displacement (LVDTs #2 through #5) magnitudes could be compared with time and location. Figure 3 compares the transient displacement components at Upland (15 ft.) and Upland (60 ft.). The results show the majority of variation occurs within the tie-ballast gap component and the gap height at Upland (15 ft.) appears to gradually increase with time. This suggests an influence from the tie-ballast gap as there was little to no evidence of a relation between permanent displacements and ballast or subgrade stiffness (Wilk, 2017).

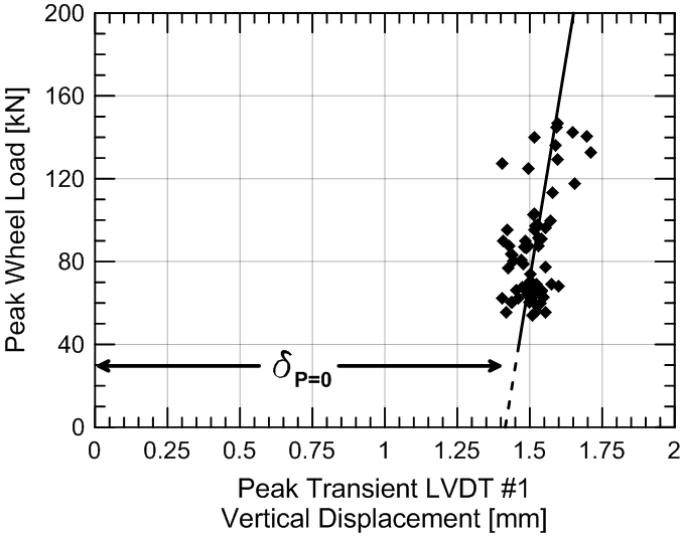


Figure 2. Typical load-displacement curve of Upland (15 ft.) displaying tie-ballast gap

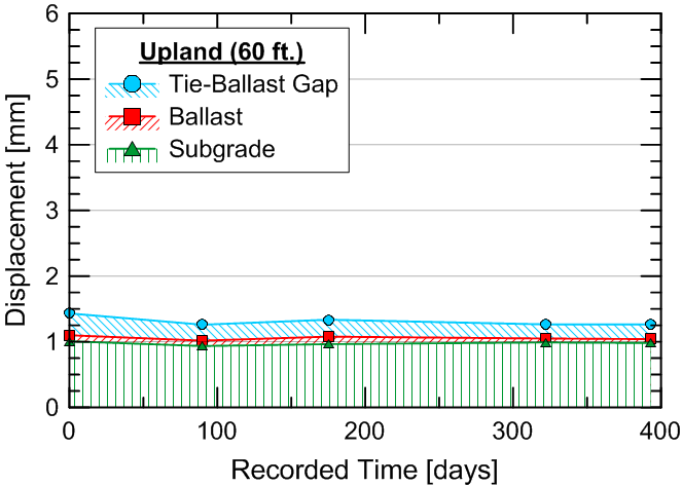
The development of tie-ballast gaps at transition zone locations is anticipated as the rail and connected ties will hang from the track that is supported by the rigid open-deck bridge and open track. However, the influence of tie-ballast gaps on accelerated ballast settlement is still relatively unclear. For example, laboratory testing by Selig & Waters (1994) showed ballast settlement rates can increase up to five times with the inclusion of a 1 to 4 mm gap as opposed to a tie that continually in contact with the ballast. The additional movement from the poor contact could increase ballast abrasion and breakage.

Impacts from the tie-ballast contact and uneven distribution of loads may also increase and concentrated load on particular ties. These two mechanisms are explored in the subsequent sections. An additional factor producing accelerated settlement at Upland (15 ft.) is degraded and moist ballast from ballast breakdown, fine infiltration, and blocked drainage. All factors likely played an interconnected role but deeper look into each is required for better standing of the transition zone system deterioration.

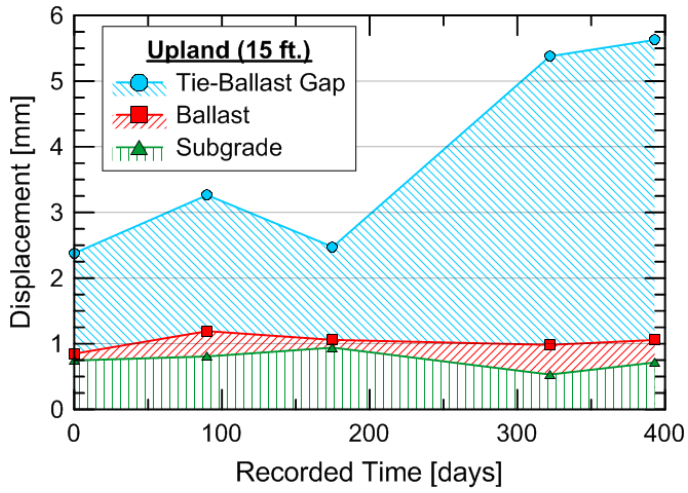
2.3 *Tie Accelerations*

Accelerometers were attached to the Upland (15 ft.) tie on the last two days of transient data collection and Upland (60 ft.) on the last day of transient data collection to measure tie accelerations. The benefits of using accelerometers is they can measure impacts and vibrations occurring within the track, giving additional insight into track behaviour and loading. At the Upland (15 ft.) sites, the accelerometer measured an acceleration spike at the moment the tie contacted the ballast from every wheel pass (Wilk et al., 2016).This shows that some impact load can be present during tie-ballast contact, which may increase the load being distributed to the ballast or damage, degrade, and abrade the ballast.

Figure 4 plots the average peak tie acceleration of each passing wheel from impact/loading measured at



(a)



(b)

Figure 3. Transient displacement components for (a) Upland (60 ft.) and (b) Upland (15 ft.)

Upland (15 ft.) and Upland (60 ft.) with various tie-ballast gap heights. The results show an increase in tie acceleration with increasing tie-ballast gap height. It must be emphasized that no trend is proposed because this relation will be site specific and dependent on train velocity and support conditions of surrounding ties. However, this does suggest tie-ballast closure can play a detrimental role in transition zone performance.

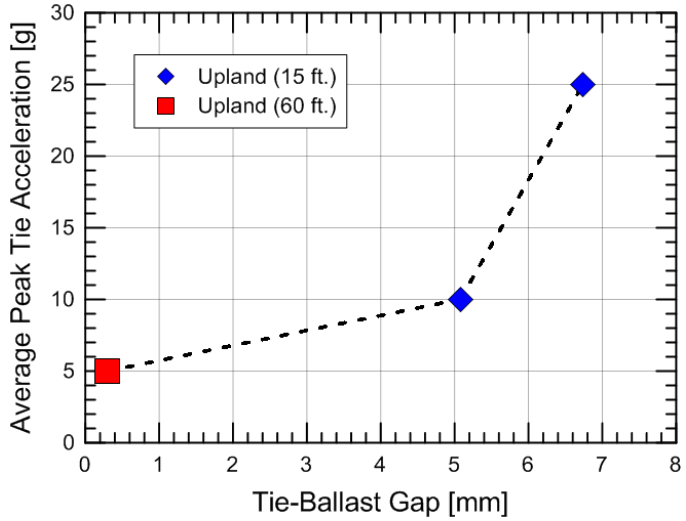


Figure 4. Relation between tie-ballast gap and peak tie acceleration at Upland Street Bridge.

3 SHORT-TERM INSTRUMENTATION

A second system of measurement used for the study is short-term instrumentation. Opposed to the long-term instrumentation introduced in the previous section, the objective of the short-term instrumentation is to evaluate track performance while eliminating invasive measurement techniques and minimizing setup time and track fouling.

The short-term instrumentation setup consists of two high-speed video cameras to measure rail and tie displacements and eight accelerometers to measure track vibrations and impacts along with estimating tie displacements along the track. This setup is used at eight different transition zone locations in the United States, with varying performance. The first three setups involved well-performing transition zones with no known track geometry maintenance since installation. The next three setups involved transition zones requiring reoccurring track geometry maintenance. The last two sites involving recently renovated transition zones with track geometries that have not yet required maintenance, but more time is required before any conclusion can be made.

Table 1 displays the eight sites along with various site attributes. For example, the well-performing track locations, to date, all had ballasted-deck bridges, confining wing walls, and either HMA or geoweb underlayment. The one exception is Site #7 which solely uses under-tie pads (UTPs) in the approach. The three sites requiring maintenance did not have any significant transition designs. Additionally, Site #8 includes four different approaches at a single bridge.

The goal of the short-term instrumentation is to monitor the amount of movements and vibrations in the track, the variation along the track, and identify regions of particular interest. This may include the impact or free-body vibration of an unsupported tie or the impact from a rail joint. Accelerometers are typically evenly spaced along the bridge, approach, and open track while high-speed video cameras usually consists of an approach and open track location.

General results from the short-term instrumentation are as follows:

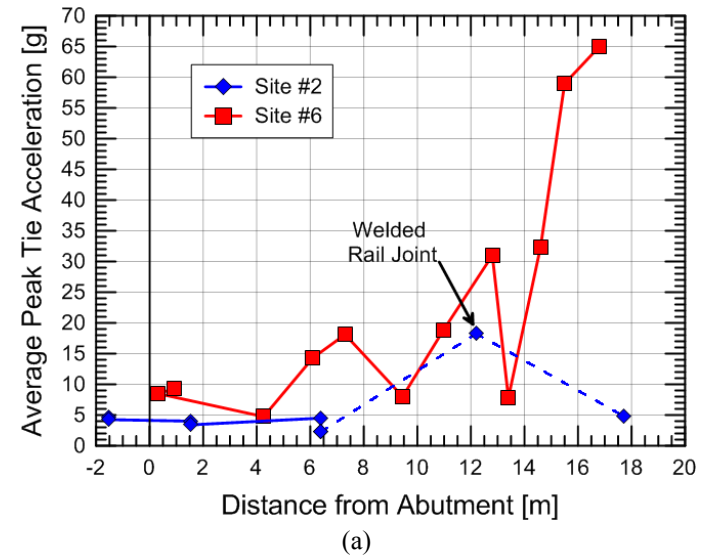
- Well-performing sites had minimal variation in track response along the track and tie accelerations were typically below 5g. This is attributed to the even load distribution and lack of relative movement between track components, i.e. good tie support.
- Sites with historical track geometry problems had variation in response with distance and time along with higher accelerations from impacts and loading vibrations. This variation in response is believed to be caused from uneven load distribution and impacts due to relative movement between track components. These higher and uneven tie acceleration magnitudes can indicate the potential of increased loads in the track.
- Besides track support, train velocity appears to be the most influential factor with tie acceleration magnitudes. For Site 5, if impacts are excluded, i.e. train loading vibrations only, the trend appears to be about 0.06 to 0.09g/km/hr while about 0.25 g/km/hr if impacts are included. These values will likely vary from site-to-site and more information is required before general conclusions can be made.

To show the first two bullets, the average tie acceleration response between Site #2 and Site #6 is compared in Figure 5a. While train velocity and other factors prevent a true comparison, the difference in behaviour is apparent. The well-performing Site #2 shows consistent tie accelerations below 5g except for the locations of a welded rail joint, which

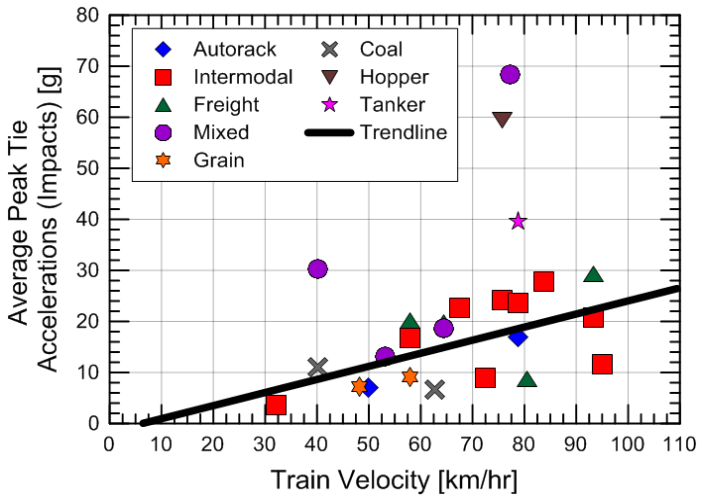
Table 1. List of measured sites along with basic attributes and transition designs

Site	Geometry Problems	Train Velocity [km/hr]	MGT	Tie Type	Bridge Deck*	Wing Wall [m]	Approach Modification*	Ballast Condition
1	No	40	7	Timber	Ballasted	8.2	HMA	Clean
2	No	40	70	Concrete	Ballasted	7.3	HMA	Clean
3	No	40	70	Concrete	Ballasted	5.2 (1 side)	HMA	Clean
4	Yes	16	?	Timber	Open	1.8	N/A	Unknown
5	Yes	177	?	Concrete	Open	N/A	N/A	Fouled
6	Yes	100	60	Timber	Open	N/A	N/A	Fouled
7	N/A	177	?	Concrete	Open	N/A	UTP	Clean
8a	N/A	40	15	Timber	Ballasted	2.7	Geoweb	Clean
8b	N/A	40	15	Timber	Ballasted	2.1	HMA	Clean
8c	N/A	40	15	Timber	Ballasted	N/A	Geoweb	Clean
8d	N/A	40	15	Timber	Ballasted	2.1	Soil Grout	Clean

* HMA represents hot-mixed asphalt and UTP represents under-tie pad



(a)



(b)

Figure 5. (a) Comparison of average response at Site #2 and Site #6 and (b) relation between average peak tie acceleration and train velocity at Site #6.

displays values of about 18g. Site #6, which displays historical track geometry issues, has varying tie acceleration magnitudes with distance. The low tie acceleration values near the approach are due to load distribution from rail-tie gaps and then impacts and uneven loads are expected further from the abutment.

Figure 5b shows the change in average tie acceleration (average for all eight accelerometer locations) with increasing train velocity for various types of trains. This trend is expected to be different for each site, especially at locations with impacts.

4 3D DYNAMIC NUMERICAL MODELING

A progressive settlement analysis is implemented to simulate the change in loading environment during the settlement of a transition zone. The analysis uses an iterative procedure that predicts loading, displacement, and settlement at 0.4 MGT increments. The settlement model used is based from Sato (1997) and modified by Dahlberg (2001) and outputs ballast settlement as only a function of tie load. The equation is shown in Equation 1:

$$y = 5.87E^{-9} * (\mathrm{P} - 25)^4 \quad (1)$$

where y is tie displacement in mm and P is tie load in kN. The load at each tie from the front and back wheels of the passing train truck are used to calculate the settlement under each tie and the geometry is updated. This procedure is repeated until 28 MGT is reached.

The numerical model incorporates the entire track system including the secondary suspension system of a train truck, the rail, concrete ties, and the substructure. The train truck passes from the open track onto a timber tie open-deck bridge. The tie loads of the first ten ties from the abutment are measured allowed the ballast settlement to be calculated in each iteration.. The tie and ballast are modeled as separate entities and discontinuities between the two surfaces are allowed. All elements are modeled as homogenous, isotropic, linear elastic materials. In open track, the rail and substructure stiffness along with tie spacing causes each tie to receive about 40% of the peak wheel load. The ratio of maximum tie load / static wheel load is defined as Maximum Normalized Tie Load and is the assumed load distribution for when all ties are in intimate contact. Increases and decreases of tie load are normalized by this 40% value (see Fig. 6).

The progressive settlement analysis is simulated at 0.4 MGT increments (20,000 wheel passes) for a total of 28 MGT (1.4 million wheel passes). The results show the ballast settlement in the bridge approach settles in manner that evenly distributes the wheel load amongst all the underlying ties, therefore minimizing the tie load. This occurs because if a particular tie experiences greater load than the surrounding ties, it will also experience greater ballast settlement. Then, the ties with greater settlement will receive less load in the next iteration as wheel load gets redistributed from ties with less support, i.e. greater settlement, to ties with better support, i.e. less settlement. This process produces a ballast settlement profile that minimizes tie loads and keeps the approach in a state of equilibrium.

However, this analysis assumes the ballast is perfectly homogenous and does not change properties over space and time. Physically, this is rarely true as ballast properties vary both spatially and temporally. To introduce the effect of heterogeneity, the ballast surface calculated at 28 MGT is randomly varied by ±0.5 mm under each tie and five sensitivity analyses (SA) were conducted. The results in Figure 6 show small deviations from the original ballast surface profile could increase tie loads up to 80%. This suggests that heterogeneities in the ballast or subgrade

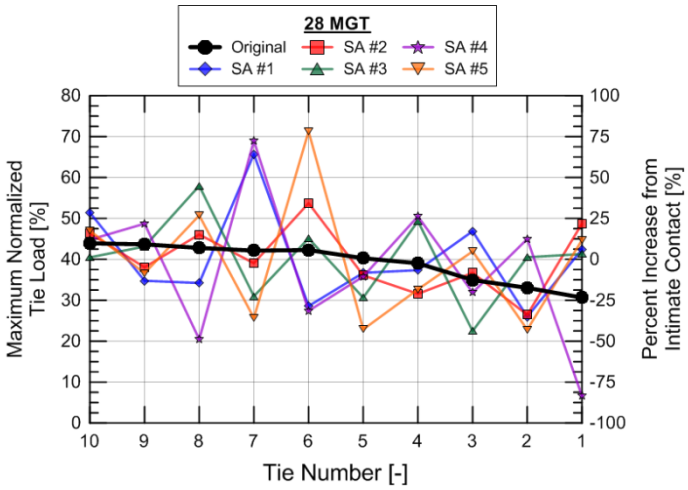


Figure 6. Normalized Tie Load Sensitivity Analysis.

can concentrate and increase tie loads in the approach.

5 REMEDIAL SOLUTIONS

The previous two sections presented field and numerical data that indicates poor tie support, i.e. hanging ties, or local differential settlement can play a detrimental role in transition zone performance. Once a transition zone is put into service and experiences loading, the approach will settle from ballast densification and possible subgrade compaction as opposed to the minimal settlement of the track on a fixed structure such as an open-deck bridge. The differential rail elevation will produce tie-ballast or rail-tie gaps because the stiff rail causes the tie to hang over locations of greater ballast settlement. Then, increased relative movement between track components can increase tie loads and ballast damage. The degraded ballast is then expected to settle at higher rates than its non-degraded
counterparts, especially if wet due to blocked drainage from the abutment This can produce a negative feedback loop in which remediation is difficult without replacement of the entire ballast structure.

Other factors that can expedite this process is rapid settlement immediately after tamping as the ballast densifies, gradual subgrade settlement as subgrade compacts or hydrocompresses, degraded and fouled ballast, and blocked drainage.

5.1 *Design and Remedial Techniques*

Design or remedial techniques to prevent the differential settlement from occurring typically requires balancing the entire transition zone system. This means that the stiffness and settlement between the bridge and approach should ideally be equal and differential settlements within the approach should be minimal. Additionally, the transition zone must be recognized as a system in which all components play an important role and disregarding a single component, e.g. ballast or subgrade layer, can initiate the negative feedback loop. This helps prevent increased dynamic loads and prevents the formation of tie-ballast gaps.

The two typical methods of achieving balance are to soften the track on the bridge and better support the approach. Softening the track on the bridge can include the conversion of open-deck bridges to ballasted-deck bridges or using under-tie pads and/or ballast mats. Better supporting the approach can involve using concrete wing walls to better confine the approach, hot-mixed asphalt/geoweb/geogrid underlayment, and subgrade stiffening solutions. Typically, combinations involving three or more solutions have consistently produced transition zones with minimal required maintenance (Stark et al., 2016). The ballasted-bridge deck, confining wing walls, and HMA combination from Site #2 in Table 1 is shown in Figure 7a.

A second potential method is to directly mitigate the negative effects of the tie-ballast interface by installing under-tie pads (UTPs). UTPs are essentially a thin rubber pad that is connected to the bottom of the tie that serves as a cushioning layer. Anticipated benefits of UTP include: (1) a reduction in approach ballast and tie degradation by better distributing the load to the ballast, (2) increased vibration damping, and (3) reduced contact stress between the tie and individual ballast particles. Therefore, UTPs should reduce the ballast pressure and abrasion resulting in a reduction of ballast settlement. Photographs of UTPs installed on concrete ties are displayed in Figure 7b.

(a)

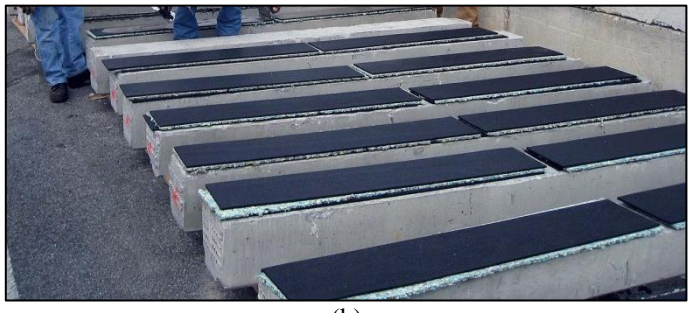

(b)

Figure 7. Photographs of (a) bridge approach with ballasted-deck bridge, confining wing walls, and HMA underlayment and (b) under-tie pads.

5.2 *Resurfacing Techniques*

Another potential underappreciated technique for better maintaining track geometry at transitions is improving resurfacing techniques. Current methods of resurfacing typically involve either automated or pneumatic tamping. Automated tamping raises the rail elevation essentially by loosening the ballast underneath the tie. Pneumatic tamping raises the rail elevation by pneumatically pushing new ballast underneath the rail seat. Both methods tend to disturb and loosen the ballast and result in rapid track settlement that reverts back to its original elevation and resurfacing is required within a year (Stark et al., 2015).

Automated tamping is widely used because of its speed and mechanization and is desirable when resurfacing long stretches of track in short amounts of time with minimal labor. However, a major drawback in the automated tamping procedure is that the ballast is disturbed from its post-compaction equilibrium state and loosened, forcing the ballast to repeat the compaction/post-compaction cycle after each resurfacing event. In addition the current tamping technique degrades and breaks down the ballast every resurfacing event, producing degraded ballast that will settle at a quicker rate (Selig & Waters, 1994).

While replacement of tamping on a wide-scale it is not anticipated because of its speed and cost-effectiveness, recommendations on how to improve resurfacing techniques at specialized location such as transition zones that typically experience track geometry problems are introduced below. Some of these solutions are new and not tested but are potentially viable because they address some of the key factors causing ballast settlement immediately after resurfacing.

Spot tamping bridge approaches two weeks after resurfacing events could potentially extend the service life of the track between maintenance cycles. The purpose of spot tamping is to identify local regions within the approach that has experienced settlement since resurfacing and to re-tamp those specific areas. This can address the weak spots of the newly tamped approach, such as the bridge-approach interface. New pneumatically tamping devices that better distributes the ballast underneath the tie without breaking the ballast could also be developed. This could be accomplished by using different tamping heads and vibration frequencies.

Stoneblowing is an alternative resurfacing method that has been developed and implemented in the United Kingdom and parts of Europe (McMichael, 1991). One of the main benefits of stoneblowing is that it leaves the ballast in its post-compaction state and adds additional stone material to fill the tie-ballast gap. The goal of this procedure is to reduce the ballast compaction stage after resurfacing. An additional benefit of stoneblowing is the reduction in ballast degradation with each resurfacing cycle. Recently, innovative ideas of combining the benefits of stoneblowing and UTPs by blowing stone mixed with rubber pellets have been tested in the laboratory and has shown to further reduce the breakdown and settlement of the ballast (Sol-Sanchez et al., 2016).

One potential drawback from stoneblowing is the stones can still degrade, breakdown, and even fall within the gaps of the underlying ballast. An alternative idea of stoneblowing is installing rubber or plastic shims, defined as hanging tie shims (HTS), underneath the tie during resurfacing. The primary obstacle is getting the shim fully underneath the tie as the underlying angular ballast particles will catch the shim as it slides underneath the tie. Therefore, this technology is considered a work-in-progress.

6 SUMMARY

This paper presents a general overview of a recently completed study investigating the root causes of the differential movement at transition zones and the benefits of various mitigation techniques. Long-term monitoring, short-term monitoring, and numerical modelling were used to investigate these problems. A summary of findings are below:

- The three general root causes are: (1) lack of track settlement on the bridge, (2) increased dynamic loads in the approach, and (3) reduced-performance substructure conditions in the approach.
- The development of tie-ballast gaps in the approach can redistribute the loading throughout the track system, produce impacts as the tie establishes contact with the ballast during loading, and promote ballast deterioration from tie-ballast abrasion due to unrestricted tie movement.
- Design and remedial techniques should attempt to balance the approach and bridge by using a combination of bridge softening and approach supporting techniques.
- Under-tie pads (UTPs) can directly mitigate against the negative effects of the tie-ballast interface. Anticipated benefits are (1) a reduction in approach ballast and tie degradation by better distributing the load to the ballast, (2) increased vibration damping, and (3) reduced contact stress between the tie and individual ballast particles.
- Improvements in resurfacing techniques could extend the geometry life between resurfacing events. This can include new methods of pneumatic tamping, stoneblowing, or the insertion of shims underneath the tie.

7 ACKNOWLEDGEMENTS

The authors would like to acknowledge the Federal Railroad Administration (FRA) BAA funding for the "Differential Movement at Railway Transitions" research project (DTFR53-11-C-0028) and the project supervision provided by Cameron Stuart. The research team also gratefully acknowledges the assistance of Deb Mishra, Erol Tutumluer, Mike Tomas, Marty Perkins, Carl Walker, and Steve Chrismer of Amtrak for their assistance with installation of the field railroad track instrumentation, monitoring the instrumentation, and interpretation of the results.

8 REFEENCES

Coelho B., Hölscher P, Priest J., Powriw W., & Barends F. 2011. An Assessment of Transition Zone Performance. *Proc IMechE Part F. J Rail Rapid Transp.* Vol 225. Pp. 129-139.

Dahlberg, T. 2001. Some railway settlement models – a critical review. *Proc IMechE Part F. J Rail Rapid Transp.* Vol 215. Pp. 289-300.

Indraratna, B., Salim W. & Rujikiatkamjorn C. 2012. *Advanced Rail Geotechnology – Ballasted Track*. CRCPress/Balkerma, Leiden. The Netherlands.

Kerr A.D & Bathurst L.A.A. 2001. Method for Upgrading the Performance at Track Transitions for High-Speed Service. *DOT/FRA/ORD-02-05*, September 2001.

Li, D. & Davis D. 2005. Transition of Railway Bridge Approachs. *Journal of Geotechnical and Geoevniornmental Engineering,* ASCE, November 2005. 131(11): pp. 1392-1398.

McMichael P.L. 1991. The economics of stoneblowing for the maintenance of way. *Proceedings of the 1991 International Heavy Haul Railway Conference*, June 1991, Vancouver, B.C., Canada.

Moale C., Smith D., Stark T.D., Wilk S.T., & Rose J.G. 2016. Design and Performance of Three Remediated Bridge Approaches. *Proc: 2016 American Railway Engineering and Maintenance-of-Way Association Conference*, Orlando, FL.

Mishra D., Tutumluer E, Stark T.D., Hyslip J.P., Chrismer S.M., & Tomas M. (2012). Investigation of differential movement at railroad bridge approaches through geotechnical instrumentation. *J Zhejiang Univ-SciA (Appl Phys & Eng)*, 12(11), pp. 814-824.

Plotkin D. & Davis D. 2008. Bridge Approach and Track Stiffness. DOT/FRA/ORD-08-01, February 2008.

Sato, Y. 1997. Optimatization of track maintenance work on ballast track. *In Proceedings of the world Congress on Railway Research (WCRR '97)*. Florance, Italy, 16-19 November 1997, Vol B., pp. 405-411.

Selig E.T. & Waters J.M. *1994. Track Geotechnology and Substructure Management*. Long: Thomas Telford.

Sol-Sánchez M, Moreno-Navarro F, and Rubio-Gámez MC. 2016. Analysis of ballast tamping and stone-blowing process on railway track behavior: the influence of using USPs. *Geotechnique*, Vol 66, No 6, pp 481-489.

Stark T.D., Wilk S.T., Rose J.G., & Moorhead W. 2015. Effect of Hand Tamping on Transition Zone Behavior. *Proc: 2015 American Railway Engineering and Maintenance-of-Way Association Conference*, Minneapolis, MN.

Stark T.D. & Wilk S.T. 2016. Root cause of differential movement at bridge transition zones. *Proc IMechE Part F. J Rail Rapid Transp.* Vol 230(4). Pp. 1257-1269.

Stark T.D., Wilk S.T., & Rose J.G. 2016. Design of Well-Performing Railway Transitions. *Transportation Research Record: Journal of the Transportation Research Board*. Transportation Research Board of the National Academia. Washington DC.

Wilk S.T., Stark T.D., & Rose J.G. 2016. Evaluating tie support at railway bridge transition zones. *Proc IMechE Part F. J Rail Rapid Transp.* Vol 230(4). Pp. 1336-1350.

Wilk S.T. 2017. *Mitigation of Differential Movement at Railroad Bridge Transition Zones*. PhD Thesis. University of Illinois at Urbana-Champaign, Urbana, IL.

Laboratory mechanical fatigue performance of under-ballast mats subjected to North American loading conditions

A. de O. Lima, M. S. Dersch, Y. Qian, E. Tutumluer & J.R. Edwards
University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT: Under-ballast mat applications have seen growth in the North American freight market, primarily being employed in ballasted concrete bridge decks and tunnels as a solution to lower the track stiffness while reducing the stress state of the ballast and reducing ground-borne vibrations. However, current standard procedures quantifying the under-ballast mat mechanical fatigue performance are provided solely by the German DIN 45673 standard which is tailored to European Mainline freight and passenger service. This in-turn provides challenges in implementing such test procedures to test materials intended for North American heavy haul freight lines, where the under-ballast mats will be exposed to higher axle loads. As part of this research, laboratory mechanical fatigue experiments were conducted on under-ballast mat samples based on recommended procedures from the DIN standard. Two load magnitudes were applied to the under-ballast mats in this study: loads representing the European mainline loading environment and loads representing the North American heavy haul freight loading environment. The fatigue performance was assessed using three criteria: a qualitative visual assessment of the sample's physical damage, a comparison of bedding modulus values measured prior-to and after repeated load cycles, and lastly, the impacts to the life cycle of the ballast. Samples tested showed no significant physical damage after testing. Further, although there was a significant change in the bedding modulus of the North American loaded sample immediately after the completion of the repeated loading cycles, the bedding modulus change in both the European and North American samples was practically the same when tested one week after the completion of the tests. Further, a gradation analysis revealed little impact to the ballast material other than particle surface wear. Therefore, the results from this work should provide information for future North American recommended testing of ballast mats subjected to heavy haul loads.

1 INTRODUCTION

Railroads continually look for ways to extend the life of their track infrastructure given poor track performance can lead to reduced transportation efficiencies which are vital to the success of rail transport (Sawadisavi 2010). North American (N.A.) heavy haul freight corridors are subject to both an increase in axle loads and higher expectations for reliability between service failures. To address these challenges and further increase the service-life of track components, it is important to reduce the stress state of the entire track structure, including the ballast (Indraratna et al. 2014). Maintenance and renewal expenses related to track ballast add up to 2% of the total annual spending across N.A. Class I railroads (Association of American Railroads 2016). Excessive degradation of the ballast can contribute to fouling and settlement, which consequently may increase impact loading due to the uneven track surface (Giannakos 2010, Le Pen & Powrie 2011). Hence, increasing the life of the ballast is of great interest.

An extension in ballast life may be accomplished through a variety of methods, and one emerging solution is the use of energy absorbing resilient materials in the track structure, primarily under-sleeper pads (USPs) and under-ballast mats (UBMs) (Esveld 2001). The former is an elastic pad bonded to the bottom surface of the sleeper while the latter is the focus of this study and is an elastic mat inserted below the ballast layer or concrete slab. Various researchers have already reported the benefits of introducing resilient pads in the track structure, including both UBMs and USPs (Sasaoka & Davis 2005, Auersch 2006, Dahlberg 2010, Marschnig & Veit 2011, Nimbalkar et al. 2012, Schilder 2013, Indraratna et al. 2014, Li & Maal 2015).

Marschnig & Veit (2011) reported in an assessment conducted for the Austrian Federal Railways that the implementation of USPs increased the time between tamping cycles by at least 100%. Further, Nimbalkar et al. (2012) concluded that the benefits of introducing resilient pads to the track structure were twofold: (i) attenuation of the impact forces and (ii)

reduced magnitude and duration for the impact force. Additionally, Nimbalkar et al. (2012) demonstrated a higher efficiency of UBMs in reducing impact magnitudes and ballast damage when installed over stiff supports (e.g. stiff subgrade or structure). Similarly, Indraratna et al. (2014) quantified the impacts of the component on the ballast material degradation under drop-hammer impact loads, reporting reduction values between 46.5% and 65.0% for hard and weak support conditions respectively. Indraratna et al. (2014) also concluded that the use of resilient pads provided more benefits in hard support conditions as the hard support promotes higher particle breakage and the weak support acts as an additional energy absorption medium.

Nevertheless, studies concerning the evaluation of the mechanical fatigue performance of UBMs are limited. The few published reports available are based on measurements obtained from samples from a single supplier recovered from field installations after many years of service (Wettschureck et al. 2002, Dold & Potocan 2013). Furthermore, through conversations with many in the industry, the majority of laboratory studies conducted have solely been performed for product development purposes, and have not been widely made available to the industry. Finally, the limited literature on this topic is constrained to European applications and testing procedures, with no reports providing insight into the component's performance under higher loading conditions (i.e. N.A. freight heavy axle loads or HAL).

Application cases of UBMs in N.A. freight lines have grown over the last two decades. The UBM growth has primarily been driven through the installation on ballasted bridge decks and tunnels, though there have been cases where the UBM was chosen due to its capabilities for mitigating ground borne noise and vibration. In fact, multiple Class I railroads have employed the component for new ballast deck bridge and/or tunnel construction or retrofit (Nunez 2014, Hanson et al. 2006).

Given this increase in installation frequency and lack of N.A. performance evaluation, this paper presents results from laboratory mechanical fatigue tests conducted to compare component performances under European mainline axle loads and N.A. HALs.

2 OBJECTIVE AND SCOPE

The primary objective of this study is to quantify the effects of increased load (i.e. N.A. HAL) on the mechanical fatigue performance of UBMs relative to European testing specifications. The results presented are only a portion of a larger research effort at the University of Illinois at Urbana-Champaign (UIUC) aimed at evaluating and quantifying the overall performance of UBMs and their benefits to the track structure, while exploring testing procedures for the N.A. environment.

During this study, laboratory mechanical fatigue tests were performed on three UBM samples that originated from the same lot. Each sample was subjected to a different load range representing European and N.A. loads, respectively. A visual assessment of the sample was performed to assess the physical damage incurred as a result of the repeated load cycles. Although potentially not as critical in reducing the ballast stress state in the heavy haul environment, the changes in the UBM bedding modulus were quantified to assess the UBMs ability to mitigate noise and vibration. Values were obtained directly prior-to, within 12-hours after, and 7 days after the repeated loading under a 30-cm (12-inch) ballast layer. Bedding modulus results were used to determine the relative performance of the component. Finally, ballast material characteristics were measured before and after the tests to quantify material degradation incurred due to the increased loading.

3 MATERIALS

3.1 *Under-ballast mat*

UBM samples intended for freight traffic loading conditions, "Type A", were used in this study (Figure 1). The samples comprised of a profiled mat bonded to a flat protective rubber layer with a synthetic fibre grid between. Table 1 provides details of the sample geometry, including its dimensions and thickness.

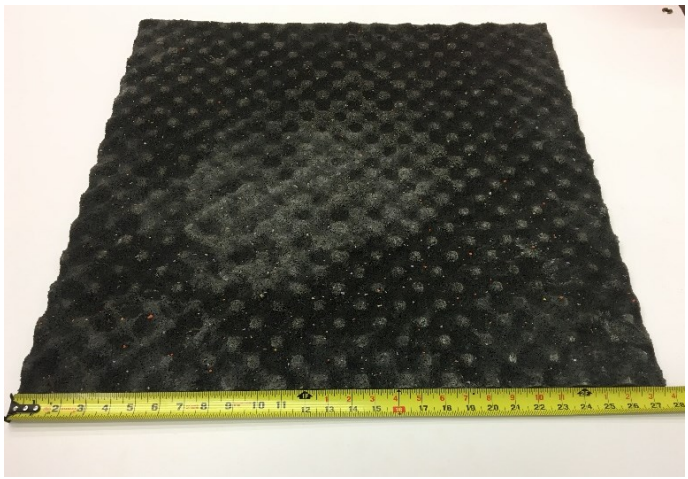


Figure 1. Type A under-ballast mat designed for freight traffic loading

Table 1. Under-ballast mat sample properties.

Label	Mat Thickness		Sample Size	Construction
	Minimum	Maximum		
	mm (in.)		mm (in.)	
Type A	5 (0.197)	10 (0.394)	699x699 (27.5x27.5)	Profiled mat bonded to flat protective layer

3.2 *Ballast*

Ballast material used for this investigation originated from a quarry commonly used by a N.A. Class I railroad and was stored in a stockpile at the laboratory facility. The coarse aggregate material consisted of crushed granite with uniformly graded particle size distribution compliant with the American Railway Engineering and Maintenance-of-way Association (AREMA) No. 4A gradation recommendations (AREMA 2016). Figure 2 shows the original gradation for the ballast material employed along with the AREMA specified gradation limits for No. 4A ballast. To ensure the quality and uniformity of the ballast used for each test, all ballast was washed, oven dried, and sieved to remove all fines from its initial state. For the purpose of this research study, fines were considered as all particles smaller than 9.5 mm or passing the ⅜-in. sieve (Qian et al. 2014). Ballast material separated during the sieving process was recombined and mixed using the recommended practices from AASHTO T 248, mixing and quartering procedures from Method B were employed due to the large size of the sample (AASHTO 2011).

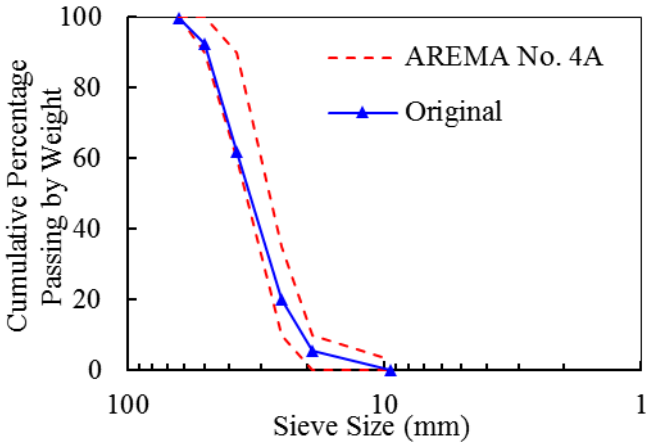


Figure 2. Ballast gradation curve

4 LABORATORY EXPERIMENTATION

Laboratory tests performed as part of this study followed modified recommendations from the German Deutsches Institut für Normung (DIN) 45673-5 standard (hereinafter referred to as DIN) for the determination of the mechanical fatigue resistance of under-ballast mat samples (DIN 2010).

4.1 *Test setup*

Due to space constraints of the test frame available for testing, the ballast box and loading plate had to be scaled down. Hence, a new design was conceived with the intent to maintain most of the considerations of the original design, notably ballast depth, and pressures at the tie/ballast and ballast/UBM interfaces. The newly designed apparatus consisted of a 30.5-cm (12-in.) diameter loading plate and a ballast box of 71 cm (28 in.) sides and 35.6 cm (14 in.) depth supporting a full 30.5-cm (12-in.) thick ballast layer section and capable of accommodating the thickest UBM sample available to the researchers at this time. Figure 3 shows the newly designed ballast box - named as the UIUC ballast box - and loading plate.

Figure 3. UIUC ballast box design and loading plate

The UBM sample was placed on the bottom of the box over the flat steel bottom. Neoprene sheets, 6.35 mm (¼ in.) thick, were placed on the sidewalls, as specified by the DIN 45673-5, to provide elasticity to the ballast layer and better simulate particle confinement experienced in the field. Clean ballast was added and compacted for 90 seconds in three 10.2-cm (4-in.) lifts; an adjustable formwork vibrator attached to a steel plate provided a 4.4-kN (1000-lbf) compaction force at 60 Hz.

4.2 *Test procedures*

Mechanical fatigue testing procedures in the DIN standard comprise of two stages of cyclic loading at incremental load levels and constant frequency in the range of 3 to 5 Hz. The two test stages apply 10,000,000 and 2,500,000 cycles, respectively, on the top-of-ballast in the setup. This leads to continuous testing lasting between 29 and 48 days depending on the loading frequency employed.

Consequently, due to the substantial amount of time required to perform the complete test procedure it has become common practice to restrict testing to the second stage loading (i.e. 2.500,000 cycles), which reduces the testing time to ⅕ of the original. This protocol is still considered to provide an appropriate indication of component performance, especially in cases of relative comparison such as the one presented in this paper; further, a similar number of cycles is used elsewhere in fatigue testing of resilient components (BS EN 16730 2016). Therefore, this work presents results of tests performed using only the second stage of testing recommended in the DIN standard.

Both qualitative and quantitative assessments of the UBM performance were performed during the

tests conducted. Primarily, a qualitative assessment of physical damages incurred to the specimens tested was performed after each of the tests. Additionally, in order to quantify the relative change in the component's vibration mitigation performance, static bedding modulus values for each sample were determined prior-to and subsequent the applied fatigue loading as specified in the DIN standard.

Throughout this research, bedding modulus is determined as the secant modulus of the stress-displacement curves within the specified load ranges for which the component is intended. Table 2 presents the evaluation ranges considered for each of the two scenarios investigated. It is worth noting that even though the evaluation ranges employed are individual to each scenario, both tested samples were loaded to the full load range of the N.A. scenario to maintain consistency of testing and enabling researchers to later evaluate bedding modulus values in additional load scenarios.

Table 2. Bedding modulus evaluation ranges employed

Loading Scenario	Evaluation Range Minimum	Maximum	Loading Rate	No. of Cycles
	kN (kips)	kN (kips)	psi/s (MPa/s)	Applied/ Recorded
European	0.9 (0.2)	12.9 (2.9)	1.45 (0.01)	3/1
N. American*	0.9 (0.2)	16.9 (3.8)	1.45 (0.01)	3/1

*Load range employed for all tests.

Additionally, after the completion of each test (i.e. 2,500,000 cycles), the ballast was collected and the effects of the increased loads to the degradation of the ballast aggregate were quantified by sieve analysis as per ASTM C136. It is believed that this material could contribute further to ballast degradation (Selig et al. 1988, Selig & Waters 1994, Qian et al. 2014).

4.2.1 *Loading conditions*

To provide a means for quantifying the effects of European and N.A. loads on the fatigue performance of the component, both load scenarios were simulated. For the first scenario, the maximum load was of 100kN (22.5 kips) which was obtained from the DIN standard. To maintain the same stress level of the DIN recommendations (i.e., 354 kPa or 51.3 psi) with the reduced-size loading plate of the UIUC ballast box, the DIN recommended load was scaled based on the loading plate areas. The resulting load value to be employed during testing was determined as 25.8 kN (5.8 kips).

Next, the equivalent N.A. scenario load was determined based on the assumption of the 95th percentile nominal N. A. HAL of 356 kN (80 kips) (AREMA 2016) and a back-calculation of the DIN-employed impact factor. The main considerations used by the DIN 45673-5 standard procedure and applied in the impact factor backcalculation are listed below:

- 22.5 tonnes (49.6 kips) European mainline axle load, and
- Loading plate area which corresponds to the support area under one rail seat of the German B70 sleeper.

Given these assumptions, and the assumption that the sleeper directly below the loading axle supports 50% of the axle load, a dynamic impact factor of 1.8 was calculated and used in the determination of the equivalent load of 161 kN (36.3 kips) for the N.A. scenario. Subsequently, the applied test load was scaled based on the same considerations previously described for the European load scenario due to the reduced loading plate size resulting in an applied load of 41.6 kN (9.4 kips). Table 3 presents additional details of both loading scenarios used.

Table 3. Fatigue loading procedures employed.

Loading Scenario	Loading Range Minimum	Maximum	Sinusoidal Frequency	No. of Cycles
	kN (kips)	kN (kips)	Hz	
European	1.8 (0.4)	25.8 (5.8)	5	$2.5x10^6$
N. American	1.8 (0.4)	41.6 (9.4)	5	$2.5x10^6$

5 RESULTS AND DISCUSSION

After the deconstruction of the ballast box setup of each test, ballast materials were collected and the UBM samples were thoroughly evaluated for physical damage. The sample tested to European loads displayed minor surface wear and compression spots immediately after testing. However, all areas initially displaying wear and compression were able to recover after just a few days of rest (i.e. no loading). Likewise, little signs of physical damage could be assessed on the sample tested to N.A. loads. In like manner to the European sample, most compression marks observed in the N.A. sample were able to recover, however, even after a few days of rest, there were still clear ballast particle imprints and minor superficial tears present around some of the existing compression marks (Figure 4). However, even the initial damages were found to be smaller than 12.7 mm

Figure 4. Superficial damage incurred to N.A. sample

(0.5 in.) long, 2.5 mm (0.1 in.) wide and 2 mm (0.08 in.) deep, and so not able to puncture through even the protective layer. Nevertheless, all observed damage incurred to either sample is not seen as degrading to the performance of the component.

As mentioned previously, bedding modulus values were calculated for both evaluation ranges of each mat prior to and after fatigue testing. Furthermore, in order to explore the effects of sample rest period on bedding modulus, two values were obtained for each loading scenario. First, immediately after the completion of the fatigue testing (i.e. test to be completed within 12 hours of the completion of the test) and second after approximately one week of test completion. Due to schedule constraints during the testing, the acquisition of immediate results from the first European sample tested was not possible, hence, a second sample had to be tested separately to provide the immediate results for the European loading scenario. All obtained results are presented in Table 4. It is worth mentioning that even though all samples were evaluated for both ranges, the percent change in bedding modulus is most relevant within the range compatible to the fatigue loading scenario of each particular sample (e.g. European Evaluation Range is most applicable to the European Loading Scenario, etc.).

Table 4. Bedding modulus results

Loading Scenario	Stage	Evaluation Range			
		European		N. American	
		C_{stat} N/mm³ (lbs/in³)	%Δ	C_{stat} N/mm³ (lbs/in³)	%Δ
European*	Initial	0.084 (309)		0.098 (360)	
	After	0.092 (340)	10%	0.107 (395)	10%
European	Initial	0.080 (295)		0.095 (349)	
	1-week	0.086 (315)	7%	0.099 (365)	5%
N. American	Initial	0.086 (316)		0.100 (370)	
	After	0.147 (540)	71%	0.167 (617)	67%
	1-week	0.093 (342)	8%	0.108 (398)	8%

*Additional sample necessary to provide results for the immediately after case of the European loading scenario.

From the presented results, there is a clear difference in the bedding modulus performance metric immediately after the completion of the fatigue loading. This can be observed across the two tests with variation in bedding modulus being higher for the N.A loading scenario than the DIN recommended European loading condition. It is hypothesized that larger amounts of elastic deformation with lower rate of recovery develop due to the higher loads, which in turn temporarily stiffens the component as is attested by an increase in bedding modulus results immediately after the test.

Conversely, results obtained after a one week rest period of the samples depict very similar percent changes of bedding modulus values – 8% and 7% for N.A and European samples respectively – values much smaller than the obtained immediately after the completion of the fatigue loading. Hence, elastic recovery of the deformations occur during the sample's rest period as supported by the bedding modulus results obtained after one week of test completion.

To provide researchers with additional insight into the effects of the higher loads, ballast gradation results were obtained in both tests. These results are presented in Figure 5 indicating no significant damage to the ballast particles had occurred due to the repeated loading. A qualitative visual assessment conducted during the collection of the particles after testing showed no signs of particle breakage. However, the presence of fines within the ballast material was noted after both load levels. This assessment, together with small shifts in the gradation curve, are thought to be related to particle surface wear of the aggregates caused by the relative movement between particles during loading and unloading cycles.

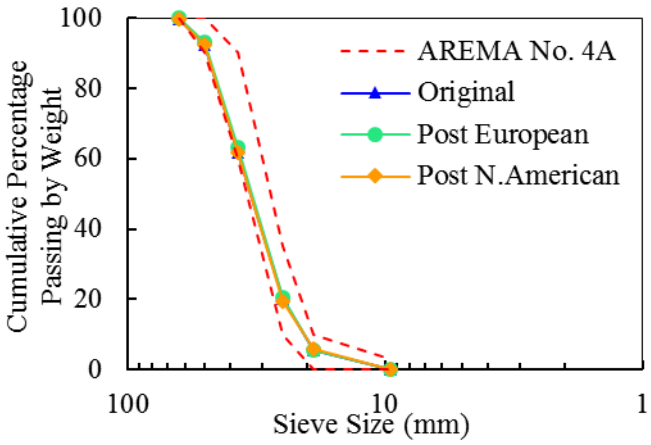


Figure 5. Ballast gradation curves from all tests conducted

As previously mentioned, material passing the 9.5-mm (⅜-in.) sieve was considered to be fines and discarded prior to construction of the box. Accordingly, an estimate measure of fine material produced can be drawn from the difference in weight between the initial and final conditions of the material. Such conclusion can only be drawn based on the assumption that loss of material was due to the generation of particles finer than the employed sieve threshold. Unfortunately, due to issues during the laboratory procedures, an exact loss amount cannot be provided for each individual case. Yet, for both tests the loss in weight of the original material employed was below 1.5%.

6 CONCLUSION

As part of this study, UBM samples were subjected to repeated loading in a ballast box simulating a section of track. Two load scenarios were employed representing North American freight heavy axle loads and the standard procedure (i.e. DIN 45673-5) representing European mainline axle loads. The objective was to quantify the effect of increased loads on the UBM physical health and the change in bedding modulus of the samples. Additionally, degradation trends of the

ballast material employed during testing were also monitored.

Although there were slightly more areas of damage as a result of the North American loading, both samples displayed negligible physical damage as a result of the load through a qualitative visual assessment. Therefore, given the fact that the samples showed no significant damage this particular UBM could withstand both North American and European loading environments.

The UBM subjected to North American loading did display a larger reduction in vibration mitigation performance when quantified immediately after the completion of the fatigue testing when compared to the UBM subjected to European loading (67% change vs 7%, respectively). However, this differences became negligible for the test case after approximately one week. Undoubtedly, these results are important when considering that a rest period exists between revenue service load applications and can allow the recuperation of the component. Further, given vibration attenuation is not typically the primary function of UBMs on heavy haul lines, this UBM should be able to serve the primary purpose of reducing the stress state on ballasted bridge decks or in tunnels. Finally, the gradation analysis results demonstrated that no significant ballast breakage occurred during either test further supporting the effectiveness of the UBM in surviving the loading environments.

This testing has provided researchers and practitioners with information into the importance of case-specific testing procedures for proper assessment of the fatigue performance of UBMs. Additionally, the compelling effects of sample rest period to the determination of changes in the bedding modulus parameter were also demonstrated and should be carefully considered when developing recommended practices.

7 REFERENCES

American Association of State Highway and Transportation Officials (AASHTO) T 248. 2011. Reducing samples of aggregate to testing size. Washington, D.C., USA.

American Railway Engineering and Maintenance-of-Way Association (AREMA). 2016. *Manual for Railway Engineering*. The American Railway Engineering and Maintenance-of-Way Association, Lanham, MD, USA.

American Society for Testing and Materials (ASTM) C136. 2014. Standard Test Method for Sieve Analysis of Fine and Coarse Aggregates. *Annual Book of ASTM Standards* 4 (2). West Conshohocken, PA, USA.

Association of American Railroads (AAR). 2016. *Total Annual Spending - 2015 Data*. Washington DC, USA.

Auersch, L. 2006. Dynamic Axle Loads on Tracks With and Without Ballast Mats: Numerical Results of Three-Dimensional Vehicle-Track-Soil Models. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 220 (2): 169–183.

British Standards Institution. BS EN 16730: Concrete sleepers and bearers with under sleeper pads, 2016.

Dahlberg, T. 2010. Railway Track Stiffness Variations – Consequences and Countermeasures. *International Journal of Civil Engineering* 8 (1): 1–12.

Deutsches Institut für Normung (DIN) 45673-5. 2013. Mechanical vibration - resilient elements used in railway tracks - Part 5: Laboratory test procedures for under-ballast mats. Berlin, Germany.

Dold, M. & Potocan, S. 2013. Long-term Behaviour of Sylomer Ballast Mats. *Rail Technology Review* 53.

Esveld, C. 2001. *Modern Railway Track*, 2. ed. MRT-Productions, Zaltbommel.

Giannakos, K. 2010. Stress on Ballast-Bed and Deterioration of Geometry in a Railway Track. *Journal of Civil Engineering and Architecture* 4 (6): 31.

Hanson, C., Towers, D.A. & Meister, L.D. 2006. *Transit Noise and Vibration Impact Assessment*.

Indraratna, B., Nimbalkar, S., Navaratnarajah, S. K., Rujikiatkamjorn, C. & Neville, T. 2014. Use of Shock Mats for Mitigating Degradation of Railroad Ballast. In: *Sri Lankan Geotechnical Journal - Special Issue on Ground Improvement*, 6 (1): 32–41.

Le Pen, L.M. & Powrie, W. 2011. Contribution of base, crib, and shoulder ballast to the lateral sliding resistance of railway track: a geotechnical perspective. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 225 (2): 113–128.

Li, D. & Maal, L. 2015. Heavy Axle Load Revenue Service Bridge Approach Problems and Remedies. In: *Proceedings of the 2015 Joint Rail Conference*, American Society of Mechanical Engineers.

Marschnig, S. & Veit, P. 2011. Making a case for under-sleeper pads. *International Railway Journal* 51 (1): 27–29.

Nimbalkar, S., Indraratna, B., Dash, S.K. & Christie, D. 2012. Improved Performance of Railway Ballast under Impact Loads Using Shock Mats. *Journal of Geotechnical and Geoenvironmental Engineering* 138 (3): 281–294.

Nunez, J. 2014. Gripping Fastening Systems. *Railway Track and Structures* (July 2014): 14–18.

Qian, Y., Boler, H., Moaveni, M., Tutumluer, E., Hashash, Y. & Ghaboussi, J. 2014. Characterizing Ballast Degradation through Los Angeles Abrasion Test and Image Analysis. *Transportation Research Record: Journal of the Transportation Research Board* 2448: 142–151.

Sasaoka, C.D. & Davis, D. 2005. Implementing Track Transition Solutions for Heavy Axle Load Service. In: *Proceedings, AREMA 2005 Annual Conference,* Chicago, IL, USA.

Sawadisavi, S.V. 2010. *Development of Machine-Vision Technology for Inspection of Railroad Track.* Master's Thesis. University of Illinois at Urbana-Champaign, Urbana, IL, USA.

Schilder, R. 2013. USP (Under Sleeper Pads) - A Contribution To Save Money In Track Maintenance. In: *AusRAIL PLUS 2013 – Driving cost out of rail*, Sydney, Australia.

Selig, E. T. & Waters, J. M. 1994. *Track Geotechnology and Substructure Management*. London: Thomas Telford Publications.

Selig, E. T., Collingwood, B. I. & Field, S. W. 1988. Causes of fouling in track. *AREA Bulletin* 717.

Wettschureck, R. G., Heim, M. & Tecklenburg, M. 2002. Long-term properties of Sylomer® ballast mats installed in the rapid transit railway tunnel near the Philharmonic Hall of Munich, Germany. *Rail engineering international* 31 (4): 6–11.

Predicting the occurrence and cost of temporary speed restrictions on North American freight lines

A.H. Lovett, C.T. Dick, & C.P.L. Barkan
University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT: Temporary speed restrictions, or slow orders, are a major concern of North American heavy-haul freight railroads because they reduce capacity and increase costs. However, current quantitative understanding of predicting the occurrence and cost of slow orders is insufficient for consideration in track maintenance planning. This paper discusses a method for determining expected costs for slow orders related to each of the three major track components (rail, sleepers, and ballast) using probabilistic models, direct and delay costs, and assumed maintenance schedules. Slow order costs vary greatly between the three track components due to both the slow order duration and occurrence rate. The quantified change in slow order costs due to changes in maintenance schedules illustrates how the maintenance planning process can consider these effects.

1 INTRODUCTION

Average train speed as reported by the major North American railroads is a key metric of network fluidity (Association of American Railroads (AAR) 2016). Lower average train speeds increase the number of crews, locomotives, and railcars required to move a given volume of freight during a set period, as well as increasing other associated operating costs (Lovett et al. 2015a). Given the impact of slowing trains, it is not surprising that temporary speed restrictions, or "slow orders," are a strategic concern for North American heavy-haul railroads. However, it is difficult to isolate the costs specific to slow orders. There are few instances in the literature that attempt to quantify the expected impact, or risk, of slow orders or other disruptions. In particular, Lovett et al. (2015b) found that slow orders related to timber crossties (sleepers) do not have sufficient impact on railroad operations to materially influence track maintenance and operating decisions. This lack of quantitative support for industry practice indicated that further research was required to determine how slow orders affect network operations. One way to estimate future impacts is through risk analysis, which considers both the probability, or frequency, and the impact of an event (Ang & Tang 2007). The authors have previously examined the effects of slow orders on rail traffic flow and operating costs (Lovett et al. 2017), so this paper will focus on estimating the rate of slow order occurrence related to rail, crosstie, and ballast defects.

Slow orders are applied to a track segment when it is found to be unsuitable for operation at the posted maximum allowable speed (MAS). These conditions arise after the track structure has been disturbed for maintenance or when track defects are detected. Slow orders caused by track disturbance typically require speeds to be reduced to 10-20 mph (16-32 km/h) for approximately 0.2 million gross tons (MGT) of traffic while the track stabilizes (Selig & Waters 1994). This process is a routine part of maintenance activities such as tamping and crosstie renewal and can be incorporated into the cost of these activities during the maintenance planning process. Therefore, slow orders for track disturbed by routine maintenance activities are not explored in detail in this paper.

Defect-caused slow orders are unexpected events that are difficult to predict and explicitly consider in maintenance planning. Various analytical and probabilistic models can estimate the frequency of track defects that require the railroad to impose a slow order. In this paper, this rate of defects resulting in slow orders is termed the "slow order rate." The estimated average slow order rate on a specific track segment can be used to determine the expected cost of slow orders and unplanned maintenance due to track defects in a given year. Understanding how the slow order rates change over time, and the factors that influence them, will also give insight into how capital maintenance timing affects the total cost of track ownership and operation. This paper will examine how to predict the slow order rate for three major track components: rail, crossties (sleepers), and bal-

last, and apply it to capital track maintenance planning. For this paper, ballast defects include alignment and surface defects, and maintenance activities to repair these defects are classified as ballast maintenance.

Although railroads can have their own maintenance standards that establish criteria for when to impose slow orders, they are also subject to government-defined standards designed to ensure a minimum level of safe train operations. Since the United States Federal Railroad Administration (FRA) Track Safety Standards (TSS) are typically the same as the Canadian regulations and apply to more miles of track, they will be taken as representative of typical North American operations (Transport Canada 2011; Federal Railroad Administration (FRA) 2014). Generally, the track geometry tolerances in the TSS vary according to track classes with each track class having a prescribed MAS. Internal rail defects are the exception because the type and size of the defect, rather than the operating speed, determines the remedial action. As the track class, and associated MAS, increases the allowable tolerances decrease. When the measured in-service track geometry exceeds tolerances, prescribed remedial actions are required on that track segment until maintenance can correct the defect (Federal Railroad Administration (FRA) 2014).

2 SLOW ORDER COSTS

Although this paper will focus on the slow order occurrence rate, it is helpful to understand the costs associated with slow orders since both rate and consequence are required to estimate risk. As with most disruptions to rail traffic, slow orders result in both direct and indirect costs that vary with the nature of the defect as well as maintenance and operational factors.

2.1 *Direct maintenance costs*

Direct costs are those associated with performing localized maintenance to repair the defect and remove the slow order, including labor, materials, and equipment. This localized, or "spot," maintenance is typically not intended to return the track to a perfect state. Spot maintenance is also relatively inefficient due to its small scale, short work windows, and reactive nature (Shimatake 1969; Esveld 2001; Zoeteman 2004; Burns & Franke 2005; Lovett et al. 2015c).

Direct slow order costs follow a traditional risk formulation since the expected costs are the defect rate times the cost per defect. These costs are largely dependent on the track component associated with the defect since different types of remedial action are required for each defective track component. For internal rail defects, a new section of rail, approximately 20 feet (6 m) long, is welded in to replace the section containing the defect (American Railway Engineering and Maintenance-of-Way Association (AREMA) 2012). Ballast-related defects are typically corrected by localized tamping. Other components, such as crossties, require local replacement of a sufficient number of the defective units to meet the required specifications (Riley & Strong 2003; Federal Railroad Administration (FRA) 2014). Railroads usually track the cost of these activities and can apply them in maintenance planning.

2.2 *Indirect costs*

Train delay is the primary indirect cost for slow orders. Lovett et al. (2017) developed a closed-form model for estimating train delay associated with a given number of slow orders and operating conditions. Since this formulation includes the slow order rate, the risk is effectively the output. It also considers the interaction between slow orders, the effects of which will be discussed further in Section 4. After the amount of train delay is computed, it must be multiplied by a train delay cost that considers the operational characteristics of traffic operating on the line (Lovett et al. 2015a).

3 PREDICTION MODELS

To predict the approximate number of slow orders on a track segment in a given year, probabilistic models were used to determine the average annual defect rate per mile. While interactions between track components may increase the local occurrence of defects once one component fails, no models were found that consider these interactions. Therefore, this paper treats each of the major track components independently.

3.1 *Rail slow order prediction*

There are a variety of rail defect types identified by the FRA, each with one or more possible remedial actions based on defect severity (Federal Railroad Administration (FRA) 2014). This analysis will focus on transverse fissures as most rail defects are given this categorization until they are removed from service for further examination (Sperry Rail Service 1999). Orringer (1990) developed a model to calculate the expected number of defects per mile based on the accumulated tonnage on the rail, inspection interval, and historical ratio of service to detected defects, which was modified by Lovett et al. (2017). Detected defects are found through inspection, while service defects are those that result in a broken rail. The Orringer model focuses on detail fractures, a subset of transverse fissures, because they were the type of rail defect causing the most rail breaks when the analysis was performed (Liu et al. 2014), but the concept can be applied to all rail defects. Only detected defects will be addressed here because service defects may

require more extensive remedial actions including stopping service on the line (Federal Railroad Administration (FRA) 2014). Orringer's original formulation was modified to use the cumulative distribution function, rather than a probability density function, which makes the model more accurate and computationally simpler. The detected rail defect slow order rate is calculated by:

$$R_{SO,R}(y_R) = N_{Rail}\frac{e^{-\left(\frac{y_R N_A}{\beta_R}\right)^{\alpha_R}} - e^{-\left(\frac{(y_R+1)N_A}{\beta_R}\right)^{\alpha_R}}}{1+\lambda(\Delta N-\theta)} \quad (1)$$

where $R_{SO,R}$ = annual detected rail defect rate per mile; N_{Rail} = number of rail sections per mile (273 (Orringer 1990)); y_R = years since rail replacement was performed; N_A = annual tonnage (MGT); ΔN = average tonnage between rail inspections (MGT); θ = minimum inspection interval (10 MGT (Orringer 1990)); λ = proportionality factor (0.014 (Orringer 1990)); α_R = Weibull shape factor (3.1 (Davis et al. 1987; Liu et al. 2014)); and β_R = Weibull scale factor (2150 (Davis et al. 1987; Liu et al. 2014)). While the model is dated, it is still used by the FRA to determine rail flaw inspection intervals (Volpe Center 2014), and the parameter values are the most recent that could be found in the literature. New research is ongoing to develop new rail defect prediction models that can be used for this purpose (Davis et al. 2016).

3.2 *Crosstie slow order prediction*

The FRA TSS require a minimum number of crossties in good condition within each 39-foot section of track based on the MAS and track curvature (Federal Railroad Administration (FRA) 2014). The Forest Service Products Curve (FSPC) can be used to determine the failure probability of timber crossties as a function of the ratio of the crosstie age to the average crosstie life (MacLean 1957), but this only gives the probability of failure for crossties of a single age. The nature of crosstie renewals is that only one-quarter to one-third of the crossties are replaced during each cycle, leading to multiple crosstie cohorts of varying ages. Lovett et al. (2015b) developed a process to determine the probability of an FRA TSS defect occurring over a 39-foot section of track given a certain amount of time has elapsed since a crosstie renewal using:

$$R_{SO,T}(y_T) = \left(P_{39}(y_T+1) - P_{39}(y_T)\right) \times \frac{5280}{39} \quad (2)$$

$$P_{39}(y) = 1 - \sum_F \prod_{j=1}^{k} \binom{n_j}{i_j} p_j^{i_j}(y)\left(1-p_j(y)\right)^{n_j-i_j} \quad (3)$$

$$p_j(y) = 1 - \exp\left[-\left(\frac{y+(j-1)c}{\beta_T A}\right)^{\alpha_T}\right] \quad (4)$$

where: $R_{SO,T}$ = annual number of crosstie related slow orders per mile; y_T = number of years since crosstie renewal; F = set of failed crosstie combinations not resulting in an FRA TSS defect in a given 39 foot (12 m) section of track; k = number of crosstie age groups; n_j = number of crossties in age group j; i_j = number of failed crossties in age group j; c = time between capital crosstie replacement; A = average crosstie life; α_T = crosstie Weibull shape factor (4.56); β_T = crosstie Weibull scale factor (1.02); and other variables as previously defined. Equation (4) represents the Weibull distribution approximation of the FSPC used as the occurrence probability for the Binomial distribution in (3).

Since the original FSPC found failure rates based on the age of a crosstie relative to the average crosstie life, the shape and scale factors in (4) do not need to consider the operating conditions directly because they can be factored into the average crosstie life. This model assumes regular crosstie replacement cycles where a set number of crossties are replaced per mile in each crosstie renewal. If the replacement cycle or number of crossties replaced is not constant, (4) will need to be modified to consider the initial age of each crosstie cohort at the beginning of the analysis period.

3.3 *Ballast slow order prediction*

Similar to rail defects, there are a variety of defect types associated with the track geometry surface and alignment. However, all track geometry defects attributable to ballast defects require the same general types of remedial actions and corrective maintenance (Federal Railroad Administration (FRA) 2014). Previous research in this area has focused on the standard deviation of various alignment measurements (Shimatake 1969; Oh et al. 2006; Chang et al. 2010). However, North American track geometry tolerances are based on absolute deviations (Federal Railroad Administration (FRA) 2014), so a new model was developed based on the methodology of Alemazkoor et al. (2015). The data set used was originally released for determining defect progression and did not explicitly include maintenance data (INFORMS Railway Applications Section 2015). To infer the timing of maintenance from the supplied data, if an FRA TSS defect was detected and no defects were detected within 100 feet on either side on a subsequent inspection, it was assumed that capital maintenance was performed. The data were then fit to a Weibull distribution.

Since there is no defined average life of a ballast defect, as is the case in the FSPC, the scale factor will need to vary based on the operating conditions. This can be done by having the scale factor be a function of the specific explanatory variables that are most significant for a particular route or section of track (Mishalani & Madanat 2002; Kleinbaum & Klein 2012; Alemazkoor et al. 2015). For the data used, only considering the time since capital maintenance resulted in the most accurate model, but this will not always be the case, so a more general form is presented here. Unlike rail and crossties, typical ballast maintenance to eliminate track geometry defects does not involve

replacing the ballast section outright with new material. Since the ballast is not truly "new," it is assumed that ballast defects will return each subsequent year that capital maintenance (undercutting) is not performed. This means that all expected ballast defects since capital maintenance was performed need to be considered in a cumulative manner, rather than just those occurring for the first time in a given year as in the rail and crosstie models. The ballast-related slow order rate is calculated by:

$$R_{SO,B}(y_B) = \left(P_{200}(y_B + 1)\right) \times \frac{5280}{200} \quad (5)$$

$$P_{200}(y) = 1 - \exp\left[-\left(\frac{y*365}{\beta_B}\right)^{\alpha_B}\right] \quad (6)$$

$$\beta_B = \exp(\boldsymbol{\Phi X}) \quad (7)$$

where $R_{SO,B}$ = annual number of ballast-related slow orders per mile; $P_{200}(y)$ = probability of a given 200-foot section of track developing one or more surface or alignment related defects at time y; y_B = years since undercutting was performed; α_B = ballast shape factor (1.088); β_B = ballast scale factor (8,862); $\boldsymbol{\Phi}$ = row vector of coefficients; and $\mathbf{X}$ = column vector of explanatory variables.

4 CASE STUDY

The models discussed in Section 3 were applied to a hypothetical 100-mile (160 km) section of 40 mph (64 km/h) track (FRA Class 3) handling 60 MGT annually. Based on industry averages, this tonnage level equates to approximately 24 one-mile long trains per day (Association of American Railroads (AAR) 2015). Rail defect slow orders result in a speed reduction to 30 mph (48km/h), while crosstie and ballast-related slow orders result in 25 mph (40 km/h) maximum speeds (Federal Railroad Administration (FRA) 2014). This case study assumes all trains operate at the MAS but average operating speeds could also be used. Rail defect inspections occur every 20 MGT and rail defects cost $895 to repair (Liu et al. 2014). Crossties have a 20-inch (51 cm) spacing on-center, 30-year average life, and a nine-year renewal cycle (Lovett et al. 2015b). Crosstie defects are corrected by replacing three crossties for a total cost of $285 (Zeta-Tech Associates Inc. 2006). Ballast slow orders cost $1,200 to repair based on an industry source for the cost of spot tamping. All slow orders are applied on the 0.1 mile (0.16 km) section of track surrounding the defect. The duration of rail, crosstie, and ballast slow orders are assumed to be one, ten, and five days, respectively. It is assumed that normal operations use 65% of the line capacity of the route (Cambridge Systematics 2007; Lovett et al. 2017), accelerating and decelerating out of and into slow orders adds an additional 30 minutes to the run time (Lovett et al. 2017), and train delay costs $950 per train-hour (Lovett et al. 2015a).

4.1 *Direct, delay, and total cost comparisons*

The defect rates for each component under the above case study parameters were calculated over a range of conditions expected during the duration of a typical maintenance cycle for that component (Fig. 1). The "defect repair" curves correspond to the equations in Section 3. These curves can be compared to the "no repair" curves that show what the theoretical defect rate would be if the defects were not repaired. The ballast curve is the exception since it is assumed that the defect rate will include both the new defects that develop during the year and all of the previously maintained ballast defects as well reoccur during the year. If the ballast defects were not maintained, the number of defects would increase at approximately the same rate but the severity would increase. Realistically, the components degrade until an acute failure, such as a rail break, occurs so the "no repair" situations will not be examined further.

Comparing the defect repair curves for each component reveals that they each perform differently. Around year 15 of the rail defect repair case, enough of the original rail has been replaced that the change in annual slow order rate begins to decrease. However, other rail related defects, such as broken welds, will likely increase as more replacement rail is welded in. Crossties exhibit a similar effect, except there are almost no defects during the first 12 years after a crosstie renewal. This is because Class 3 track only requires eight crossties in good condition per 39 feet (12 m) to be free of defects (Federal Railroad Administration (FRA) 2014). For a defect to develop, almost all of the crossties installed before the most recent renewal would need to fail. Once the crossties from the two most recent renewals have a larger probability of failure, the compounded failure probability increases dramatically. This also explains why (Lovett et al. 2015b) found that crosstie slow order

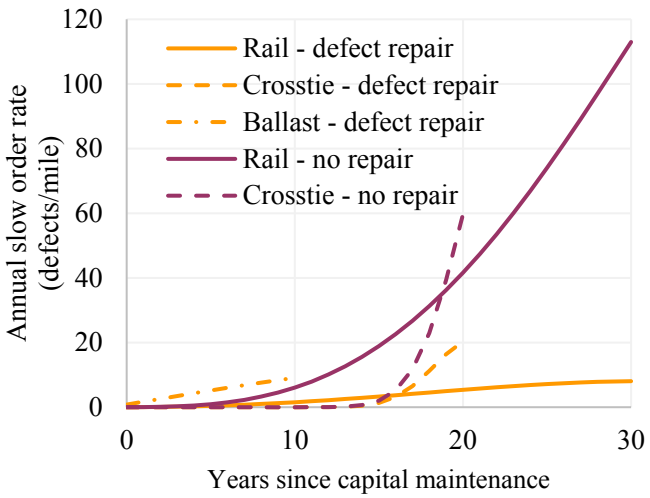


Figure 1: Slow order defect rate for the major track components with and without spot maintenance

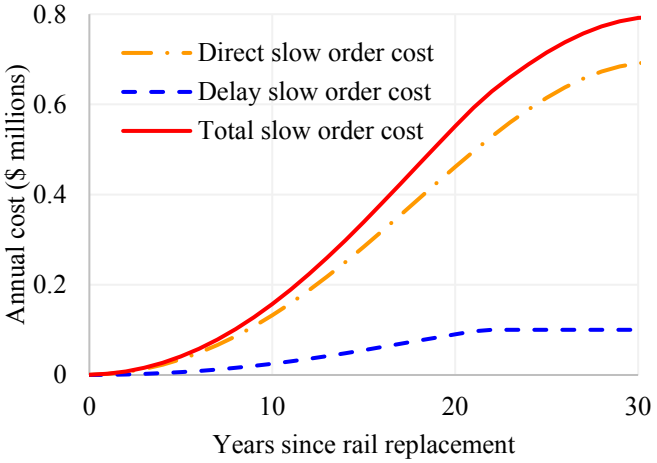


Figure 2. Annual cost of rail-related slow orders vs. years since capital maintenance

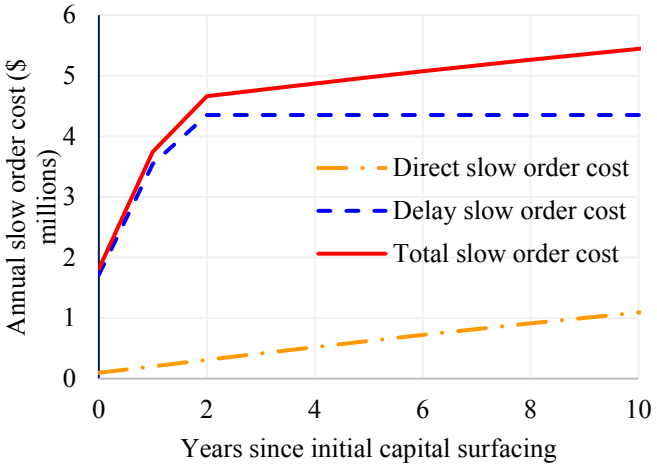


Figure 4 Annual cost of ballast-related slow orders vs. years since capital maintenance

risk would not materially influence maintenance decisions because they only calculated slow order costs until the ninth year after a crosstie renewal.

Further insight is gained by comparing the total, direct, and delay slow order costs for each component (Figs. 2-4). Each plot shows that the defect rate increases until the there are enough defects with overlapping slow orders such that entire route is effectively subject to speed restrictions, as evidenced by the plateau in the delay cost curve. The shape of the delay cost curve, including the plateau location, changes based on the traffic, train performance, and slow order characteristics (Lovett et al. 2017).

For rail (Fig. 2), the train delay costs are relatively low compared to the direct costs due to the short slow order duration (one day). The other extreme is observed for crossties (Fig. 3) where delay accumulated over the ten-day length of each slow order renders the direct costs of repair almost negligible. An increase in delay costs would be expected since the crosstie slow orders are left in place longer, but the increase is vastly disproportionate to the relative increase in the slow order duration. Ballast slow orders (Fig. 4) are consistent with this trend as the ratio between delay and direct costs are much lower than would be expected when comparing just the slow order durations of the ballast and crosstie defects. This disproportionate relationship is consistent with Lovett et al. (2017).

4.2 *Comparison of alternative maintenance timings*

Although it is interesting to look at how the slow order costs change over time, a primary benefit of these curves is to aid in capital maintenance planning. In Figures 2-4, the area under the total cost curve represents the slow order cost for each component in a given planning period. Performing capital maintenance during the planning period will decrease the slow order cost associated with the new component during subsequent years but the savings need to be balanced against the expense of performing the capital maintenance. This can be done by comparing the slow order costs for different capital maintenance timings within the planning period (Figs. 5-7).

For rail (Fig. 5) and crossties (Fig. 6), performing maintenance earlier initially reduces the slow order cost by a noticeable amount. However, comparing the slow order costs for rail in later years shows that the

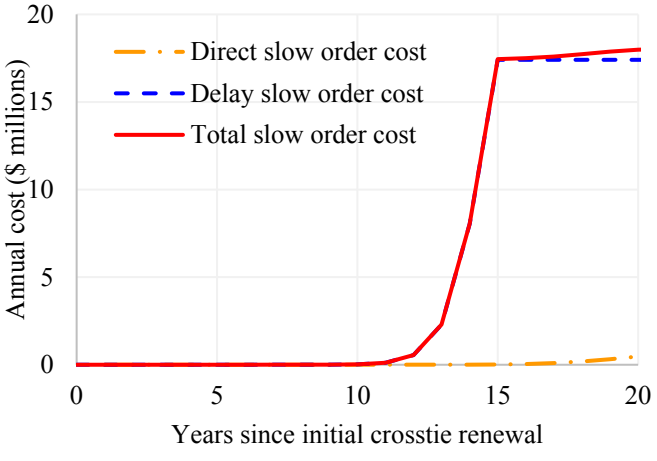


Figure 3. Annual cost of crosstie-related slow orders vs. years since capital maintenance

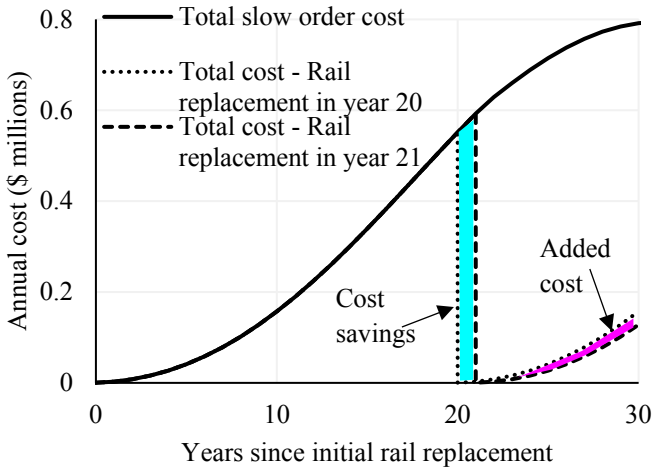


Figure 5: Rail slow order cost under different rail replacement schedules

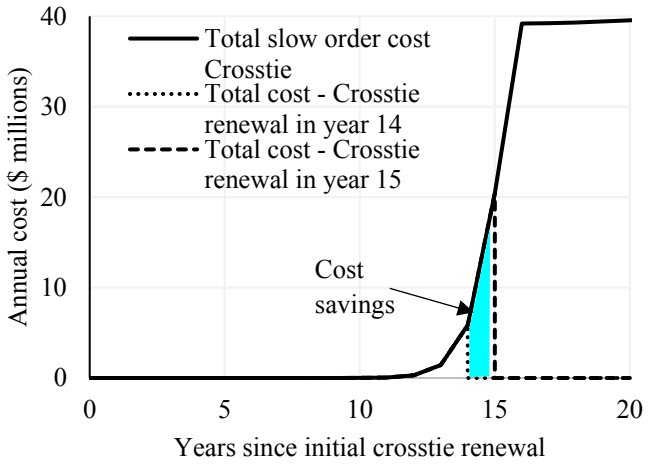


Figure 6: Crosstie slow order cost under different crosstie renewal schedules

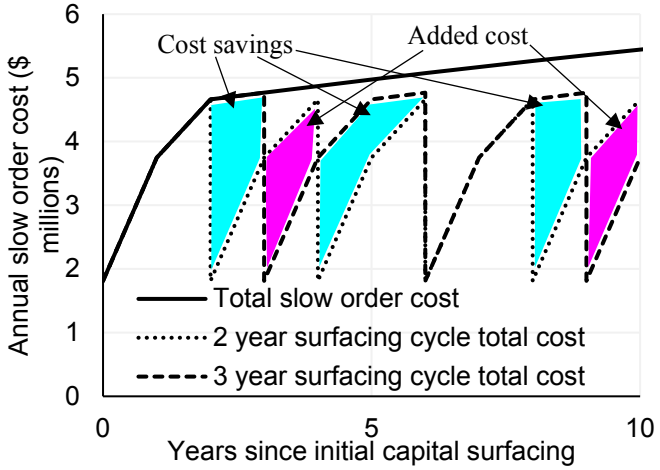


Figure 7: Ballast related slow order cost under different capital surfacing schedules

annual slow order cost is higher for the earlier replacement curve. That is to be expected since it has been longer since capital maintenance was performed. Over time, the higher slow order cost combined with costs to perform capital maintenance earlier may counteract the initial slow order savings, showing that a longer-term perspective is required for maintenance planning.

Comparison of ballast maintenance schedules (Fig. 7) shows a different perspective because capital surfacing is performed multiple times within the illustrated 10-year planning period. Since the 2-year capital surfacing interval would require more maintenance instances than the 3-year interval, the capital costs will be higher, further offsetting the slow order cost reduction. This shows that the selection of the planning window is also an important factor when comparing proposed maintenance schedules.

5 CONCLUSIONS AND FUTURE WORK

One of the key findings of this research is the impact of train delay on the cost of slow orders. For slow orders of short duration, such as those caused by rail defects, delay accounts for a relatively small proportion of the total annual slow order cost. However, as the slow order duration increases, delay cost increases disproportionately until it dominates the total cost, as is the case for slow orders related to crossties. The substantial contribution of train delay to total costs shows how important it is to consider the operational impacts of slow orders and track defects when planning maintenance intervals.

The effects of train delay and the nature of the operational impact of slow orders provide key inputs to a maintenance plan. While performing capital maintenance earlier will decrease the immediate slow order costs, additional costs are incurred in later years after the track components have degraded. Quantification of the slow order impacts allows for the capital maintenance plan to be optimized by balancing the slow order and capital maintenance costs. Additionally, if spot maintenance is made more efficient, effective, and timely it can reduce the overall costs and recurrence of slow orders while increasing the time between capital maintenance activities.

One area where this work can be made more robust is by gathering new data from the railroads and either validating these findings or developing new models that reflect the current quality of materials and maintenance practices. A new analysis could also take advantage of "big data" techniques such as machine learning that were not available for development of the rail and crosstie models referenced by this paper. Analyzing new data would also allow for comprehensive slow order models that consider the condition and maintenance history of the entire track structure rather than a single component. Applying the findings and methodology from this research to new probabilistic models will allow railroads to more effectively optimize their maintenance strategy by using a more holistic planning approach.

6 ACKNOWLEDGMENTS

Research supported by AAR and National University Rail (NURail) Center, a US DOT-OST Tier 1 University Transportation Center. Lead author also supported by the CN Research Fellowship in Railroad Engineering and Dwight David Eisenhower Transportation Fellowship Program.

7 REFERENCES

Alemazkoor, N., Ruppert, C.J., & Meidani, H. 2015. *2015 RAS Problem Solving Competition Report: Track Geometry Analytics* [WWW Document]. URL https://www.informs.org/content/download/310742/2968335/file/Report_IRA_Group.pdf (accessed 2.1.16).

American Railway Engineering and Maintenance-of-Way Association (AREMA). 2012. Recommended Repair of Defective or Broken Rail in CWR. *Manual for Railway Engineering*, Vol. 4, Ch. 4, Sec. 7 Lanham, MD, USA.: American Railway Engineering and Maintenance-of-Way Association.

Ang, A.H.-S. & Tang, W.H. 2007. *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*, 2nd ed. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Association of American Railroads (AAR). 2016. *Railroad Performance Measures* [WWW Document]. URL http://www.railroadpm.org/ (accessed 11.23.16).

Association of American Railroads (AAR). 2015. *AAR Analysis of Class I Railroads, 2014*. Washington, DC, USA: Association of American Railroads.

Burns, D. & Franke, M. 2005. Analyzing the blitz approach to m/w. *Railway Track & Structures* October: 33–38.

Cambridge Systematics, 2007. *National Rail Freight Infrastructure Capacity and Investment Study*. Cambridge, MA, USA: Association of American Railroads.

Chang, H., Liu, R., & Li, Q. 2010. A multi-stage linear prediction model for the irregularity of the longitudinal level over unit railway sections. *Computers in Railways XII*. 114: 641–650.

Davis, D., Banerjee, A., & Liu, X. 2016. *Rail Defect Prediction Model Evaluation, TD-16-043*. Technology Digest. Pueblo, CO, USA: Transportation Technology Center, Inc.

Davis, D.D., Joerms, M.J., Orringer, O., & Steele, R.K. 1987. *The Economic Consequences of Rail Integrity, R-656*. Chicago, IL, USA: Association of American Railroads.

Esveld, C. 2001. *Modern Railway Track*, 2nd ed. Delft, The Netherlands: MRT-Productions.

Federal Railroad Administration (FRA) 2014. *Track and Rail and Infrastructure Integrity Compliance Manual Volume II Track Safety Standards Chapter 1 Track Safety Standards Classes 1 through 5* [WWW Document]. URL http://www.fra.dot.gov/eLib/details/L04404#p1_z50_gD_kcompliance manual (accessed 10.10.16).

INFORMS Railway Applications Section 2015. *2015 RAS Problem Solving Competition* [WWW Document]. URL https://www.informs.org/Community/RAS/Problem-Solving-Competition/2015-RAS-Problem-Solving-Competition (accessed 10.12.16).

Kleinbaum, D.G. & Klein, M. 2012. *Survival Analysis: A Self-Learning Text*, 3rd ed. New York NY, USA: Springer.

Liu, X., Lovett, A.H., Dick, C.T., Saat, M.R., & Barkan, C.P.L. 2014. Optimization of ultrasonic rail-defect inspection for improving railway transportation safety and efficiency. *Journal of Transportation Engineering* 140(10). doi:10.1061/(ASCE)TE.1943-5436.0000697.

Lovett, A.H., Dick, C.T., & Barkan, C.P.L. 2015a. Determining freight train delay costs on railroad lines in North America. In *Proceedings of the International Association of Railway Operations Research 6th International Conference on Railway Operations Modelling and Analysis (RailTokyo2015), Tokyo, Japan, 23-26 March 2015.*

Lovett, A.H., Dick, C.T., Ruppert, C.J., & Barkan, C.P.L. 2015b. Cost and delay of railroad timber and concrete crosstie maintenance and replacement. Transportation Research Record: Journal of the Transportation Research Board 2476: 37–44.

Lovett, A.H., Dick, C.T., Ruppert, C.J. & Barkan, C.P.L. 2015c. Evaluating opportunities to aggregate track maintenance activities in extended work windows on freight railroad lines. In *International Heavy Haul Association Conference Proceedings, Perth, Australia, 21-24 June 2015*: 1337.

Lovett, A.H., Dick, C.T., & Barkan, C.P.L. 2017 (Accepted). Predicting the cost and operational impacts of slow orders on rail lines in North America. In *Proceeding of the 7th International Conference on Railway Operations Modelling and Analysis (RailLille2017), Lille, France, 4-7 April 2017.*

MacLean, J.D. 1957. *Percentage Renewals and Average Life of Railway Ties: Information Review and Reaffirmed, No. 886*. Madison, WI, USA: United States Department of Agriculture Forest Service Forest Products Laboratory.

Mishalani, R.G. & Madanat, S.M. 2002. Computation of infrastructure transition probabilities using stochastic duration models. *Journal of Infrastructure Systems* 8(4): 139–148.

Oh, S.M., Lee, J.H., Park, B.H., Lee, H.U., & Hong, S.H. 2006. A study on a mathematical model of the track maintenance scheduling problem. In Allan, J., Brebbia, C.A., Rumsey, A.F., Sciutto, G., Sone, S. (eds.), *Computers in Railways X*: 88: 85–96.

Orringer, O. 1990. *Control of Rail Integrity by Self-Adaptive Scheduling of Rail Tests, DOT/FRA/ORD-90/05*. Washington, DC, USA: Federal Railroad Administration.

Riley, J.E. & Strong, J.C. 2003. Basic track. In *Practical Guide To Railway Engineering*. Lanham, MD, USA: American Railway Engineering and Maintenance-of-Way Association.

Selig, E.T. & Waters, J.M. 1994. *Track Geotechnology and Substructure Management*. London, England: Thomas Telford Services Ltd.

Shimatake, M. 1969. *A Track Maintenance Model for High-Speed Rail: A System Dynamics Approach*, Master's Thesis. Cambridge, MA, USA: Massachusetts Institute of Technology.

Sperry Rail Service 1999. *Rail Defect Manual*. Danbury, CT, USA: Sperry Rail Service.

Transport Canada 2011. *Rules respecting track safety* [WWW Document]. URL https://www.tc.gc.ca/eng/railsafety/rules-79.htm (accessed 10.19.16).

Volpe Center 2014. *Self-Adaptive Algorithm Calculator - 2014* [WWW Document]. URL http://www.fra.dot.gov/eLib/Details/L04918 (accessed 9.23.16).

Zeta-Tech Associates, Inc. 2006. *Development of Comparative Cross-Tie Unit Costs and Values*. Fayetteville, GA, USA: Railway Tie Association.

Zoeteman, A. 2004. *Railway Design and Maintenance from a Life-Cycle Cost Perspective: A Decision-Support Approach*, Doctoral Dissertation. Delft, The Netherlands: Delft University of Technology.

A Parametric Model of the Train Delay Distribution to Improve Planning of Heavy Haul Cycle Times

Mei-Cheng Shih, C. Tyler Dick, P.E. and Christopher P.L. Barkan
University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT: Efficient heavy haul operations require accurate predictions of terminal arrival times and equipment cycle times. Most analytical and parametric approaches for railway capacity evaluation and operations planning focus on predicting average train delay and not the performance of individual trains and their delay distribution. In practice, train delay varies about this average according to some distribution and the cycle-time performance of certain heavy haul trains may be far from average. Using average values of train delay and cycle time during the creation of a heavy haul operations plan may lead to erroneous conclusions that impact the stability of train operations. To address this shortcoming, this research developed a parametric model for predicting the distributions of train delay on single-track mainlines. The new train delay distribution model is based on a set of indices developed to measure both the amount of traffic and the degree of traffic heterogeneity (differences in train speed and priority) present on the route under study. A quantile regression approach was used to build the model since existing statistical distributions could not adequately represent typical train delay distributions within heterogeneous railway traffic. The developed model can be used to assess the impact of changes in traffic mixture (number of high, medium and low priority trains) and train parameters (speed and priority) on the train delay distribution.

1 INTRODUCTION

The duration of train cycles and arrival times at loading and unloading terminals are highly relevant to the planning of heavy haul operations. Many heavy haul operations use single-track routes that are shared with other types of trains, potentially leading to unplanned train meets that lengthen the cycle time and delay arrival at terminals. The amount of delay may fluctuate with each train run, with very few trains achieving the minimum run time between terminals. A precise estimate of the train delay distribution can be used to predict the reliability of a planned train cycle or terminal arrival time relative to the minimum running time over the route. Better planning-level estimates of the reliability of heavy haul operations can improve both the efficiency and robustness of the system.

Most analytical and parametric approaches to rail traffic performance focus on predicting average train delay (Burdett and Kozan, 2006; Mitra and Tolliver, 2010; Murali et al., 2010) and not the performance of individual trains or their delay distribution. In practice, train delay fluctuates according to some distribution and the performance of an individual train may be far from the average values. Using average values of train delay for planning heavy haul operations without knowledge of the train delay distribution could lead to erroneous conclusions if a system is particularly sensitive to the occurrence of extended train delays.

However, tools currently available to practitioners are not designed to directly estimate the distribution of train delay for a given train without the need for extensive simulation. While detailed simulation models can estimate the distribution of train delay across multiple days of simulated train operations, the models are computationally-intensive. Detailed information on the route infrastructure and signal system is required to develop the simulation model, consuming scarce railway planning resources. Practitioners could be better served by a model that combines the delay distribution output of simulation with the computational efficiency of a parametric approach.

This study develops a parametric model for the distribution of train delay on a single-track line. A quantile regression approach (Koenker and Bassett, 1978; Machado and Mata, 2001; Nielson and Rosholm, 2001) is used to build the parametric model since existing statistical parametric distributions (like Gaussian, Weibull, or Poisson distributions) do not adequately represent the delay distribution of heterogeneous railway traffic.

2 BACKGROUND

Most heavy haul operations that do not operate on dedicated lines involve trains with different length, horsepower, weight, speed and priority characteristics. Some parametric models of average train delay consider variation in train speed and priority. Previous research has proposed factors to better quantify the characteristics of rail traffic with multiple train types (Chen and Harker, 1990; Landex, 2008; Gorman, 2009; Lai et al., 2010; Dingler et al. 2013). The factors proposed by other researchers are either empirical, lack generality or do not have a direct physical meaning that translates to a train delay mechanism. More generalized factors with direct links to the mechanics of train delay may better quantify rail traffic heterogeneity and the causal relationships leading to the distribution of train delay observed on a route.

Additionally, previous parametric models typically do not consider the impact of flexible train operations common in North America. Freight trains in North America do not adhere to a fixed timetable with pre-planned train meets at specific passing sidings (passing loops). Instead, each train operates with a certain amount of schedule flexibility to achieve a desired level of service defined by a maximum allowable train delay. Variation in schedule flexibility and level of service adds another dimension of traffic heterogeneity that needs to be considered by parametric models of train delay.

This study defines schedule flexibility as the departure and trip time flexibility of a single train (Figure 1). Schedule flexibility is a parameter associated with an individual train and different trains on a route can exhibit varying degrees of schedule flexibility depending on their business objectives and level-of-service requirements. The variation of schedule flexibility across all trains operating on a route during a certain dispatching period defines the operating style.

Time-distance diagrams can be used to compare "structured operation" on a fixed timetable (Figure 2a) to the "flexible operation" more common on heavy haul lines in North America (Figure 2b), Under the structured operating style, trains follow predetermined timetables with precise departure times and pre-set meet locations. Under the flexible operating style, the business objectives of heavy haul service usually require dispatchers to dynamically adjust predefined train plans. As a consequence, the departure, and arrival time of trains under flexible operation are ranges instead of points, and the trip time is a band instead of a line. Also, the potential traffic conflict between two trains is a zone instead of a precise point on the structured timetable. These differences need to be considered when evaluating the impact of flexible operation and traffic heterogeneity on train delay.

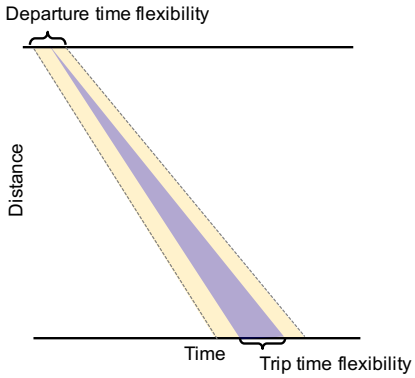


Figure 1. Departure time, and trip time flexibility

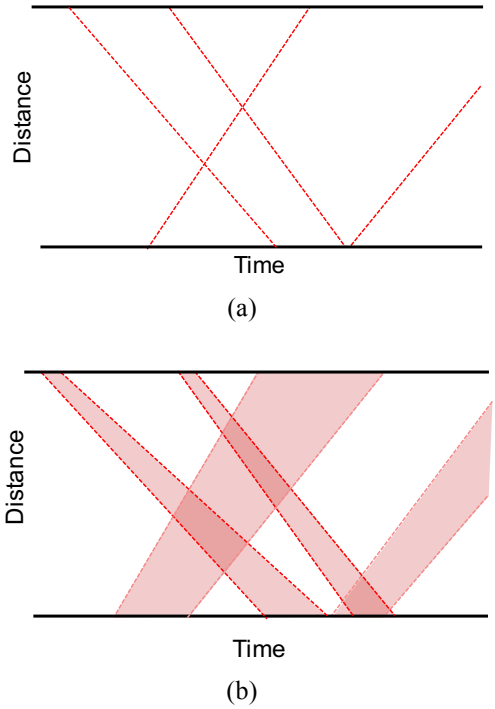


Figure 2. Time and distance diagrams under (a) fixed timetable (b) flexible operation

This study attempts to identify potential parametric model factors to explain the expected variation in train delay based on analysis of traffic conflicts under flexible operations.

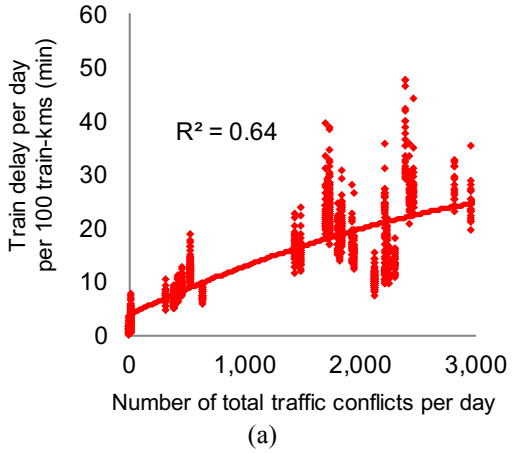
3 PARAMETRIC MODEL DEVELOPMENT

The proposed parametric model of the train delay distribution was developed in two stages. In the first stage, appropriate heterogeneity factors were developed based on the concept of traffic conflict analysis. In the second stage, the factors were used to construct a quantile regression model of the train delay distribution. A cross-validation-based process was used to assess the accuracy of the de-

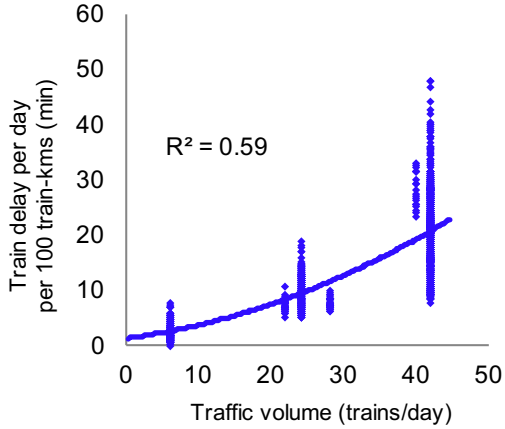
veloped quantile regression model.

Train delay is usually related to rail traffic volume through delay-volume curves. However, this approach only considers the amount of traffic on the route under study and not the degree of traffic heterogeneity. The concept of using train conflicts (also referred to as traffic conflicts) to predict train delay was presented by Gorman (2009). Gorman found that traffic conflicts, represented by the number of meets, passes and overtakes, impact train delay significantly. This study seeks to expand on this idea to determine if the number of traffic conflicts can be used to describe the relationship between the degree of traffic heterogeneity and the distribution of train delay.

As a preliminary study, various heterogeneous railroad traffic scenarios were simulated with Rail Traffic Controller (RTC) simulation software to investigate a potential relationship between the number of traffic conflicts and train delay on a representative North American shared corridor. Under these conditions, the expected number of traffic conflicts (calculated by counting the total traffic conflicts each train could encounter based on a Monte Carlo process) is more closely correlated with train delay than the total traffic volume (Figure 3). This suggests that the number of traffic conflicts captures both the impact of traffic volume and heterogeneity. By capturing additional information about the rail traffic, the number of rail traffic conflicts may be an alternative predictor of the train delay distribution than traffic volume alone.



(a)



(b)

Figure 3. Relationship between (a) traffic conflicts, (b) traffic volume, and the average train delay per 100 train-kilometers

Heterogeneity can arise from different combinations of speed variation, priority variation, and operating styles. Based on the concept of traffic conflict analysis, three train delay factors were defined to quantify these attributes of heterogeneous rail traffic:

- Total Conflicts (TC) considers all of the potential conflicts a train may encounter during its trip and that a larger number of traffic conflicts increases the difficulty of the train dispatching task. TC is calculated by examining a set of train departures, creating train paths at the train operating speed and counting the total number of conflicts between train paths. Conflicts between train paths are not resolved.
- Adjusted Train Priority (ATP) quantifies the actual priority of a train within the given traffic mixture on the route. ATP is calculated for a given target train by the summation of inferior conflicts (target train has inferior priority relative to the conflicting train) and half of equal conflicts (target train has equal priority to the conflicting train). The physical interpretation of ATP as a delay mechanic is the number of conflicts where the target train will need to stop and wait for the other conflicting train to pass.
- Inferior Pass (IP) represents the impact of train speed heterogeneity on train conflicts and delay. Variation in speed between trains creates additional delay when one train is required to pass another. IP calculates the expected number of inferior passes (target train has inferior priority to passing train). The physical meaning behind IP is the expected number of passes that will cause the target train to stop or encounter delay.

With the three train delay factors defined, the second step of the study is to apply a quantile regres-

sion approach to build the parametric model. Sogin (2013) fit a Weibull distribution to delay data for homogeneous unit train traffic. However, a preliminary test conducted for this study showed that a Weibull distribution, along with other common parametric distributions, did not adequately model the delay distribution of a train in heterogeneous rail traffic. For this reason, a quantile regression approach is used.

Quantile regression is a statistical method used by researchers in the area of macroeconomics to model the interaction between variables and output distributions (Arias et al., 2001; Koenker, 2005). The quantile regression model creates multiple regression lines that each represent a quantile boundary (Figure 4). For example, the 97.5th percentile line is the best fit such that 97.5 percent of the data points are below the line and 2.5 percent are above the line. In this case, the slope of the curve represents the sensitivity of the 97.5th percentile of train delay to the TC index defined above.

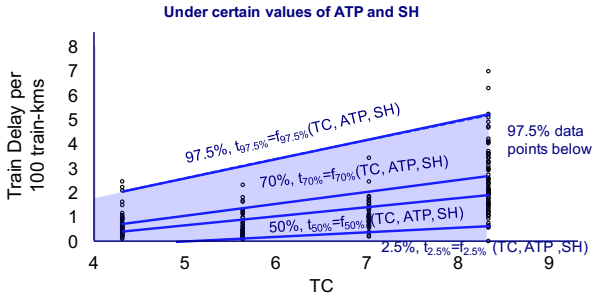


Figure 4. Example of a set of quantile lines

Based on the quantile regression technique, the mentioned train delay factors are used as possible variables to consider when building a regression model to predict train delay distributions. The constructed model can be used to analyze the response of the train delay distribution to changes in traffic (as reflected by changes to the three train delay factors).

To develop the quantile regression model, a train plan with associated departure flexibility is required as an input to calculate the values of the train delay factors for each individual train in the train plan (Figure 5).

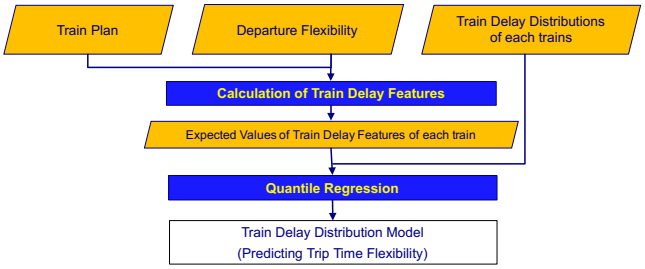


Figure 5. Framework of the development of parametric model

For a given train, each factor is calculated through a Monte-Carlo process that considers the departure time flexibility of each individual train (Figure 6). For each iteration of the process, a set of train paths is selected from within the theoretical train band space defined by the departure time flexibility. From the unresolved train conflicts between these train paths, the corresponding values of the three factors are calculated for each train. The final calculated factor levels for each train are taken as the average values of the factor levels of each train after a certain number of iterations. These values can be used as close approximations of the real factor values.

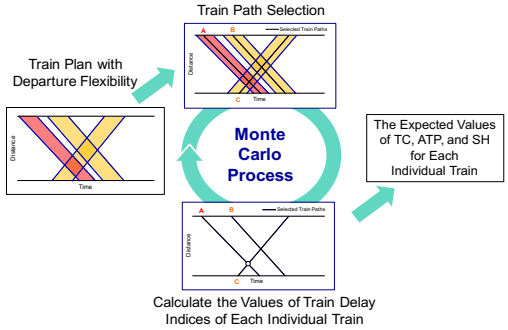


Figure 6. A Monte-Carlo-based process for factor level calculation

The calculated factor levels for each train and the corresponding distribution of train delays for that train (determined from historical data or simulation of the flexible operations) are then used by the quantile regression approach to construct a Train Delay Distribution Model. The Train Delay Distribution Model inputs are the train delay factor levels associated with a given train path. The model output is the quantile boundaries predicting the distribution of delay for that train.

4 MODEL VALIDATION

The Mean Absolute Deviation (MAD) (Ohashi et al. 2010) is used to assess the performance of the developed train delay distribution model. The function (1) shows the calculation of MAD, where $\hat{Q}_\alpha$ is the estimate of a certain quantile α, and Q_α is the real value of the quantile α.

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|\hat{Q}_\alpha - Q_\alpha| \quad (1)$$

For this study, cross validation (Kohavi, 1995) was used to obtain a more robust MAD (Figure 7). The first step in cross validation is to divide the available dataset into k different subsets with equal

size. The second step is a looping process where each round a subset is selected as the test data set and other subsets are combined as the training data set. The training data set is used to construct a regression model and the test data set is used to evaluate the performance of the model. Each round of this process generates a MAD for each test set, and the final MAD is the average of all MADs. This final MAD score represents the potential error in predictions of specific quantiles of the train delay distribution made by the constructed model. As will be demonstrated in the case study, the MAD scores for a model may vary between each quantile of the delay distribution. This indicates that the predictive ability of the model is not consistent across the entire delay distribution.

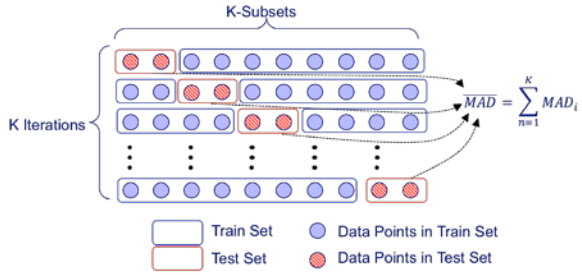


Figure 7. Relationship of cross validation and MAD in this study

5 CASE STUDY

To demonstrate the construction and predictive ability of the train delay distribution model, it was applied to a 242-mile single-track route representative of heavy haul operations in North America (Figure 8). The traffic tested on this infrastructure is assumed to have the following characteristics:

- Traffic volume varies from 8 to 26 trains per day
- Three train types with high. medium, and low priority exist within the traffic mixture with the percentage of each train type ranging from 25 to 75 percent
- For each potential combination of traffic volume and mixture, there exist three different patterns of train departures from the terminals at either end of the route
- The departure time flexibility for high, medium, and low train types are 0.5, 1.5, and 3 hours, respectively

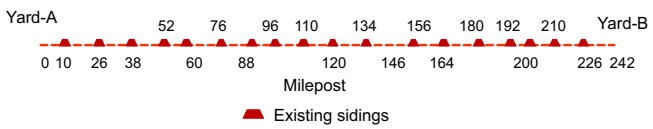


Figure 8. Tested infrastructure layout in RTC

The traffic and infrastructure parameter values above were used to construct an experiment matrix based on the concept of full factorial design. All scenarios in this matrix were simulated with RTC over multiple days of operation to obtain the delay distribution of each train.

The traffic parameters for each scenario in the experiment design were also input to the Monte Carlo process to obtain the factor levels associated with each train. The calculated train delay factors and their associated train delay distributions for each train were then used to construct the train delay distribution model for this route.

To assess the performance of the case study model, ten subsets were used in the cross validation process. The mean of delay corresponding to each quantile and calculated MADs show the potential prediction error for different quantiles (Table 1). For example, the mean of real delay value corresponding to the 5th percentile is 4.4 minutes. On average, five percent of trains experience less than 4.4 minutes of delay per 100 train-miles. The MAD for the 5th percentile is 1.6 minutes suggesting that the true value of the 5th percentile of train delay falls within the range of 4.4 +/- 1.6 minutes.

For the 5th to 70th quantiles, the MADs range from one to three minutes of train delay. Above the 70th quantile, the MAD increases quickly from three to eight minutes for the 95th quantile. The ability of the model to precisely predict delays corresponding to higher quantiles decreases. This result indicates that the current factors used to construct the model cannot completely explain the extreme train delays observed in the RTC simulation output. Additional research is needed to identify factors that can explain these high-delay cases and incorporate them into the model.

Table 1. MADs from cross validation and the corresponding delays of each quantile

Percentile	Mean of MADs	Mean of delay from raw data
5%	1.6	4.4
10%	1.7	5.9
20%	1.7	8.3
30%	1.8	10.5
40%	1.8	12.6
50%	2.0	14.6
60%	2.3	17.1
70%	2.8	19.8
80%	3.8	23.5
90%	4.8	29.1
95%	8.1	34.5

6 CONCLUSION AND FUTURE STUDY

Equipment cycle and terminal arrival times are important considerations for heavy haul operations

planning. Estimating the reliability of these times under different operating scenarios requires knowledge of both the average and distribution of train delay. An accurate prediction of the train delay distribution can improve the efficiency and robustness of heavy haul operation plans. However, previous parametric models used for predicting train delays estimate the average delay and not the entire train delay distribution.

This study proposed a train delay distribution model based on the concept of quantile regression and newly developed factors that capture the combined impact of speed, priority variation and operating style (with its associated schedule flexibility). The cross validation result in the case study suggests that the model can accurately predict the delay corresponding to different quantiles below the 70^{th} percentile.

To improve the performance of the model at higher percentiles, regularized quantile regression (i.e. Lasso) could potentially be used to replace the standard quantile regression approach implemented in this study. Additionally, conducting root cause analysis of extreme train delays under flexible operation may identify new factors to better characterize heterogeneous train operations. Adding these factors to the model could also improve its performance.

A next step for this research is to test the developed model on other rail corridors and generalize the train delay distribution model. The model developed in the case study was only validated on the same route and under similar traffic conditions used to construct the model. While it can accurately predict the train delay distribution under a wide range of rail traffic conditions on this corridor, it is an open research question to determine how well the model will translate to other single-track routes with different siding spacing. Additional factors describing the route infrastructure could be included in the model to potentially generalize it to any single-track corridor. However, even before this step is taken, this study still provides a general quantile regression framework that can be used by practitioners and researchers to develop their own route-specific versions of the train delay distribution model.

7 ACKNOWLEDGEMENTS

Partial support for this research was provided by the Association of American Railroads (AAR) and the National University Rail (NURail) Center, a USDOT-OST Tier 1 University Transportation Center. The lead author was also supported by the CN Research Fellowship in Railroad Engineering.

8 REFERENCES

Arias, O., Hallock, K.F., & Sosa-Escudero, W. 2001. Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twin data. *Empirical Economics* 2: 7-40.

Burdett, R.L. & Kozan, E. 2006. Techniques for absolute capacity determination in railways. *Transportation Research Part B: Methodological* 40(8): 616-632.

Chen, B. & Harker, P.T. 1990. Two moments estimation of the delay on single-track rail lines with schedule traffic. *Transportation Science* 24(4): 261-275.

Dingler, M.H., Lai, Y.C., & Barkan, C.P.L. 2013. Mitigating train-type heterogeneity on a single-track line. In *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(2): 140-147.

Gorman, M.F., 2009. Statistical estimation of railroad congestion delay. *Transportation Research Part E: Logistics and Transportation Review* 45(3): 446-456.

Koenker, R. & Bassett, G. 1978. Regression quantiles. *Econometrica* 46: 33-50.

Kohavi, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, San Mateo, California: Morgan Kaufmann, 2(12): 1137-1143.

Koenker, R. 2005. *Quantile regression*. Cambridge university press. Cambridge, UK.

Lai, Y.C., Dingler, M., Hsu, C.E., & Chiang P.C. 2010. Optimizing train network routing with heterogeneous traffic. *Transportation Research Record: Journal of the Transportation Research Board*, 2159: 69-76.

Landex, A. 2008. Methods to Estimate Railway Capacity and Passenger Delays. Doctoral Thesis. Technical University of Denmark, Department of Transport. Lyngby, Denmark.

Machado, J.A.F. & Mata, J. 2001. Earning functions in Portugal 1982-1994: evidence from quantile regressions. *Empirical Economics* 26: 115-134.

Mitra, S. & Tolliver, D. 2010. Estimation of railroad capacity using parametric methods. *Journal of the Transportation Research Forum* 49(2): 111-126.

Murali, P., Dessouky, M., Ordonez, F., & Palmer, K. 2010. A delay estimation technique for single and double-track railroads. *Transportation Research Part E: Logistics and Transportation Review* 46: 483-495.

Nielson, H.S. & Rosholm, M. 2001. The public–private sector wage gap in Zambia in the 1990s: a quantile regression approach. *Empirical Economics* 26: 169-182.

Ohashi, O., Torgo, L., & Ribeiro, R.P. 2010. Interval Forecast of Water Quality Parameters. In *Proceedings of the (ECAI) European Conference on Artificial Intelligence 2010*: 283-288.

Sogin, S.L. 2013. Simulations of Mixed Use Rail Corridors: How Infrastructure Affects Interactions Among Train Types. *Master's Thesis*, University of Illinois at Urbana-Champaign, Department of Civil and Environmental Engineering. Urbana, IL, USA.

Building Capacity through Structured Heavy Haul Operations on Single-Track Shared Corridors in North America

Darkhan Mussanov, Nao Nishio and C.Tyler Dick, P.E.
University of Illinois at Urbana-Champaign, Urbana, IL, USA

ABSTRACT: North American freight rail traffic reached a peak in 2006 on the strength of heavy haul transportation of bulk commodities and double-stack intermodal containers in international trade. However, the composition and geographic distribution of this traffic has substantially changed. Coal traffic has declined by over 20 percent since 2006 while intermodal traffic has reached record levels, with particularly strong growth in domestic intermodal traffic requiring predictable service on precise schedules. On the predominantly single-track North American rail network, allowing for schedule flexibility results in continually changing train meet and pass requirements that can increase train delay and constrain capacity, but also decrease utilization of capital-intensive infrastructure. This research uses simulation to investigate the relationship among line capacity, level of schedule flexibility, allowable train delay and the mixture of scheduled and flexible trains operating on a corridor. Rail Traffic Controller software is used to simulate different combinations of these factors. Based on the results of these simulations, practitioners can consider introducing more structured schedules for heavy haul traffic as one potential approach to building the capacity required to accommodate rising intermodal traffic on emerging shared corridors.

1 INTRODUCTION

Heavy haul railway operations provide for the safe, efficient, economical and reliable transportation of freight and bulk commodities in particular. A key to the efficiency of heavy haul operations is maximizing the amount of freight hauled by each train and continuously cycling rolling stock between loading and unloading points. To achieve these objectives, the departure and arrival times of heavy haul freight trains may be dictated by production timelines and vessel sailing schedules instead of a pre-planned railway timetable. Many heavy haul operations exhibit schedule flexibility with trains departing terminals at different times each day. Without a fixed timetable and pre-planned locations for each train meet on single track, flexible operations require train dispatchers to resolve train conflicts in real time. As this paper will demonstrate, for a given track infrastructure layout, these flexible operations consume more capacity than structured operations where trains depart within shorter time windows.

In North America, many corridors that were once dominated by heavy haul operations on flexible schedules are seeing increased numbers of intermodal trains that require more structured operations to meet business objectives. Although freight traffic volume in the United States peaked in 2006, following three years of traffic declines due to economic recession, US freight rail traffic has slowly returned to 2006 levels. However, the composition and geographic distribution of this traffic has substantially changed. Coal traffic has declined by over 20 percent since 2006 while intermodal traffic has reached record levels, with particularly strong growth in domestic intermodal traffic requiring predictable service on precise schedules (AAR, 2016).

Heavy haul and intermodal trains have different train speed, priority, length, weight and level-of-service requirements. Maintaining the efficiency of heavy haul trains through schedule flexibility while simultaneously providing the consistent and reliable level of service required by intermodal trains presents a substantial operational challenge.

While infrastructure expansion may provide the required capacity for these two operations on a shared single-track corridor, it may be possible to maximize the capacity and level of service of these corridors by striking a balance between flexible and structured operation of heavy haul traffic.

This research investigates the relationship among line capacity, level of schedule flexibility, allowable train delay and the mixture of scheduled and flexible trains operating on a corridor. Simulation experiments are used to determine the capacity of a representative North American single-track rail corridor

as it transitions from purely flexible operation with random train departures to a highly-structured operation on a precise timetable designed to minimize train delay. The results of this study can be used to develop a better understanding of the interaction between line capacity and schedule flexibility. Ultimately this knowledge can help railway practitioners make trade-off decisions involving capacity and the level of flexibility allowed in their heavy haul operating plans.

2 BACKGROUND

2.1 *Structured and flexible operations*

As described in the introduction, most North American freight trains do not operate according to a prescribed timetable. Although the same general pattern of trains may be operated on a given day of the week from week-to-week, a specific train may depart a terminal over a range of times. Heavy haul freight trains that cycle between loading and unloading terminals may almost enter the network at random according to production and shipping schedules of bulk commodities. This type of operation is referred to as "improvised" or "flexible operation" (Martland 2010).

Flexible operations are in contrast to "structured operations" where a pre-planned timetable specifies exact departure times, locations of meets and passes with other trains, and arrival times at the destination terminal and intermediate points along the route (Martland 2010). Structured operation is common on rail networks in Europe and Asia and even some transit and commuter railways in North America.

In this paper, the term "schedule flexibility" is used to describe the amount of variation in departure times relative to a baseline train operating plan with target departure times. Operations with low schedule flexibility are more structured and will have all trains departing relatively close to their planned departure times. A completely structured operation will exhibit no schedule flexibility and all trains will depart at their precise scheduled time. Operations with high schedule flexibility will have train departures distributed over a wider range around the planned departure time. A completely flexible operation will have trains departing randomly during each day of operation.

2.2 *Previous research*

In the context of North American heavy haul operations, railway line capacity is defined by the largest traffic volume that can be operated on a route segment while maintaining a minimum level-of-service standard (Krueger 1999). The level-of-service standard corresponds to a maximum allowable average train delay. Train delay is calculated as the difference between the actual running time of a train over a route segment and its minimum running time without interference from meets or passes with other trains. Capacity can be defined by the average train delay for all trains or may also consider level-of-service requirements that are specific to certain types of trains (Shih et al. 2015).

The Canadian National Parametric Capacity Model (Krueger 1999) describes relationships between train delay, traffic volume and various parameters describing track infrastructure, traffic and operating conditions. However, the model parameters do not explicitly consider the amount of schedule flexibility.

Martland (2003) modelled the link between heavy haul line capacity and terminal operations. Martland observed that capacity plans must account for disruptions and variability in operations. Martland also suggested that there may be ways to increase capacity through changes to operations that do not require infrastructure expansion. This view is shared by Ede and et al. (2007) who considered how to reach the most cost-effective heavy haul operation by balancing parameters involving infrastructure investment, train characteristics, operations and maintenance scheduling. It was suggested that variability in train departures throughout the day reduces capacity compared to an operation with evenly-spaced departures. Ede et al. (2007) also indicated that sharing the track with other types of trains presented a challenge for heavy haul operations capacity. Neither Martland or Ede at al. quantified their qualitative observations of the capacity effects of departure variability and interaction between types of trains.

Research on the capacity of a single-track corridor with heterogeneous train operations was conducted by Dingler who examined the interaction of higher-speed intermodal trains with lower-speed heavy haul bulk unit trains (Dingler 2010). The train delay response was highest on the simulated corridor when equal numbers of intermodal and unit trains were present on the corridor. However, the research did not consider different departure time flexibility for the two types of trains. Subsequent research confirmed that train delay increases as heterogeneity increases for randomized train departures (Sogin et al. 2013; Sogin et al. 2016).

Boysen (2012) built a capacity model that examined ways network capacity can be influenced by stakeholder needs, train characteristics and operating parameters. This model suggested that capacity could be increased by decreasing heterogeneity in the system but did not address schedule flexibility. Subsequent analysis of a heavy haul iron ore line in Sweden and Norway documented the heavy haul capacity consumed by dedicating timetable slots to passenger trains and maintenance activities (Boysen 2013). To operate passenger trains on a fixed timetable, the heavy haul iron ore trains must adhere to pre-planned timetable slots. To account for the

flexible departure of the iron ore trains according to production and shipment demand, the freight operator purchases more schedule slots than are needed to operate the average daily traffic volume (Lindfeldt 2010). Trains are held at terminals until their scheduled departure time and many slots go unused during times of low iron ore demand.

A case study of a Dutch railway line analysed the capacity of scheduled and flexible operations using two different train control systems (Goverde et al. 2013). The standard UIC compression method of capacity analysis was used to model scheduled trains but flexible trains required Monte Carlo simulation and rescheduling algorithms. The research focused more on the ability of each control system to handle unscheduled operations and less on the specific capacity effects of varying amounts of flexible operations.

Dick & Mussanov (2016) directly investigated the effect of varying amounts of schedule flexibility on train delay for fixed volumes of homogeneous rail traffic on representative North American single-track and partial-double-track lines. When starting from a fixed timetable with little train delay, introducing schedule flexibility in the form of random variation about the schedule departure time caused an increase in average train delay. Small amounts of schedule flexibility created rapid increases in train delay but beyond a certain level of schedule flexibility, further increases in train delay were not observed. Additional work investigated the relative delay performance of different combinations of scheduled and flexible trains operating on the same corridor across a range of schedule flexibility (Mussanov et al. 2017). The results suggest it is difficult for scheduled trains to have a high level of service in the presence of flexible trains and the effect is magnified as the line nears capacity.

2.3 *Research hypotheses*

This paper examines the relationships between schedule flexibility, train-type specific levels of service (allowable train delay) and the capacity of representative North American single-track corridors under combinations of flexible and scheduled freight trains.

Past study by Dick & Mussanov (2016) emphasized that for a given volume of trains operating with schedule flexibility, the incremental train delay response became increasingly insensitive to additional schedule flexibility (Figure 1a). It is also accepted that North American freight train delays increase exponentially with increasing train volume (Figure 1b) (Krueger 1999 and Dingler 2010). The shape of the relationship between delay and schedule flexibility for a constant volume in Figure 1a leads to the hypothesized family of delay-volume curves in Figure 1b. At high levels of schedule flexibility, the delay-volume curves in Figure 1b are more closely spaced to correspond to the decreasing sensitivity of delay to schedule flexibility illustrated in Figure 1a. The simulation experiments described in this paper investigate if this hypothetical relationship can be observed on representative single-track corridors.

Figure 1b suggests a hypothesis that a higher traffic volume with lower schedule flexibility may exhibit the same average train delay as a scenario with a lower traffic volume, but a greater schedule flexibility. By setting a maximum allowable delay to serve as the level of service that defines line capacity, the corresponding combinations of volume and schedule flexibility yield a hypothetical relationship between schedule flexibility and capacity (Figure 1c). As schedule flexibility increases, line capacity is expected to decrease. The decrease in capacity from pure structure operation (zero schedule flexibility) represents the capacity penalty for allowing some trains to operate on flexible schedules. Similarly, moving from flexible schedules to more structured operations is expected to increase line capacity.

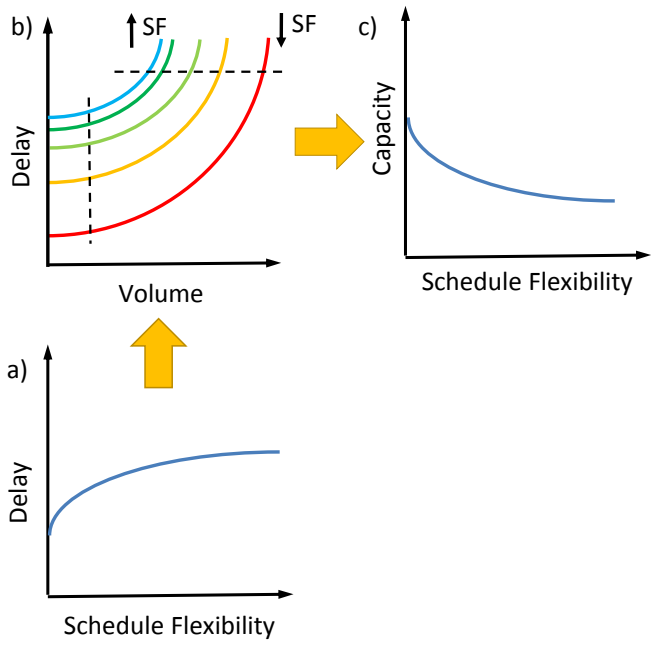


Figure 1. Combination of (a) Previously established relationship between schedule flexibility and train delay and (b) shape of the train delay-volume curve can be used to create a (c) hypothetical relationship between schedule flexibility and line capacity

3 METHODOLOGY

3.1 *Rail Traffic Controller*

This research develops train delay and capacity metrics with the use of Rail Traffic Controller (RTC), the industry-leading rail traffic simulation software in the United States. RTC is used by a wide range of public and private organizations, including most

Class I railroads, Amtrak, and rail operations consultants. Specially developed for the flexible North American railway operating environment, RTC emulates dispatcher decisions in resolving train meet and pass conflicts while simulating the movement of trains over rail lines subject to specific route characteristics. RTC allows users to alter different infrastructure, train and control system parameters in the simulation and analyse the train delay response.

RTC can dispatch trains according to a timetable for structured operations and also randomly for flexible operations. RTC dispatches scheduled trains from originating terminals at the specified time. Flexible trains depart terminals within a user-specified range before and after the planned departure time. For example, if the user assigns a schedule flexibility of 60 minutes to one train, RTC will randomly dispatch that train as early as 60 minutes before the planned departure or 60 minutes after the planned departure time. Within the interval of +/- 60 minutes, the departure time will vary for each simulated day according to a uniform distribution. RTC allows the user to specify a different amount of schedule flexibility for each train to be dispatched in the simulation. With this capability, RTC can simulate a representative route segment of the North American freight network carrying both heavy haul bulk commodity trains with schedule flexibility and high-priority intermodal trains with scheduled departure times.

The main output of the RTC simulation of interest to this research is the train delay response. Train delay is averaged by train type across all simulation days and normalized by the total train-miles (or train-km).

Since this research incorporates schedule flexibility, the exact train departure schedule is different for each simulated day and replication is required to achieve a stable average train delay response. To estimate the required number of replicates, one scenario in the experiment design was replicated 100 times with different initial random seeds for each five-day simulation. After seven replications, the average train delay stabilized. For this research, each scenario in the experiment design is replicated ten times with each simulation considering five days of train operations. For each scenario in the experiment design, this simulation plan yields train delay data for 50 days of train operations that are then averaged into a single train delay data point for those experimental conditions.

3.2 *Baseline schedule*

Trial-and-error was used to develop a combination of baseline schedule and infrastructure that minimized train delay for a traffic volume of 36 trains per day on a single-track corridor (Table 1). The baseline schedule follows the "return-grid" operating model where trains alternately depart from each end terminal on even intervals and all train meets are designed to occur at evenly-spaced passing sidings. As described in the next section, this baseline schedule is perturbed through the introduction of schedule flexibility and changes in traffic volume to generate a range of experimental conditions. The route is 386 km in length with 23 passing sidings (passing loops) that are 3.22 km long and placed every 16 km on-center. All trains have the same 125-railcar consist and are representative of North American freight operations.

Table 1. Baseline schedule, route and train parameters

Parameter	Values
Length of route	386 km
Siding length	3.22 km
Siding spacing	16 km
Number of sidings	23
Traffic volume	36
Scheduled departure interval	2 hours
Maximum speed	37 km/hr
Locomotive type	SD70 3206 kW, 3 locomotives per train
Train consist	115 railcars at 125 tons each. 2.07 km total length
Operating protocol	CTC 2-block, 3-aspect

3.3 *Experiment Design*

Three variable factors were used to generate the different simulation scenarios in the experiment design: traffic volume, schedule flexibility and traffic composition (Table 2). While the route infrastructure was held constant, the baseline pattern of train operations was altered according to obtain the factor levels in Table 2.

Table 2. Experiment design factors and factor levels

Factor	Factor levels
Traffic volume	24, 28, 32, 36, 40, 44 trains per day
Schedule flexibility	+/- 0, 10, 60, 120, 720 minutes
Traffic composition (Percent of trains with flexible schedules)	0%, 25%, 50%, 75%, 100%

To vary the traffic volume, trains were added or removed from the initial baseline schedule of 36 train departures (Table 3). Trains were removed and added in pairs to maintain directional balance. Removing four trains from the ideal schedule for 36 trains provides the initial departure times for the scenarios with 32 trains per day. The remaining trains are not re-spaced to even intervals but remain in their original departure slots to preserve the ideal "return-grid" schedule. To increase traffic volume from 36 to 38 trains per day, four trains were dispatched in a time slot used for two trains in the baseline schedule. For 44 trains per day, the extra trains required doubling traffic in four of the original slots (Table 3).

The traffic volume on the line is comprised of two types of trains: scheduled and flexible. Scheduled trains follow the exact departure times specified in the baseline schedule regardless of the schedule flexibility factor level. Flexible trains randomly depart each terminal over the range of departure times relative to the baseline schedule specified by the schedule flexibility factor. As described in the previous section, ten minutes of schedule flexibility allows the flexible trains to depart any time within a 20-minute window extending ten minutes before the scheduled departure and ten minutes after. This research considers five levels of schedule flexibility: 0, 10, 60, 120 and 720 minutes. Zero schedule flexibility corresponds to a structured operation on the baseline schedule while 720 minutes corresponds to completely random departures over the entire day. Since the previous work of Dick & Mussanov (2016) found little incremental change in train delay for high levels of schedule flexibility, additional levels of schedule flexibility between 120 and 720 minutes were not simulated.

Table 3. Comparison of train slots for 32, 36, & 44 trains when the traffic compositions of 50% and 75%

Scheduled time (% flexible trains)	32 trains (50%)	36 trains (50%)	36 trains (75%)	44 trains (75%)
0:00:00	---	SS	SS	SS
1:20:00	FF	FF	FF	FF
2:40:00	SS	SS	FF	FF
4:00:00	FF	FF	FF	FFSS
5:20:00	SS	SS	FF	FF
6:40:00	FF	FF	FF	FF
8:00:00	SS	SS	SS	FF
9:20:00	FF	FF	FF	FFSS
10:40:00	SS	SS	SF	FF
12:00:00	---	FF	FF	FF
13:20:00	SS	SS	SS	FF
14:40:00	FF	FF	FF	FFSS
16:00:00	SS	SS	FF	FF
17:20:00	FF	FF	FF	FF
18:40:00	SS	SS	FF	FF
20:00:00	FF	FF	FF	FFSS
21:20:00	SS	SS	SS	SF
22:40:00	FF	FF	FF	FF

To determine if the relationship between schedule flexibility and capacity depends on the combination of scheduled and flexible trains on the route, different traffic compositions were included in the experiment design. This study considered traffic compositions where 0%, 25%, 50%, 75% and 100% of the trains operate with flexible schedules. For example, the scenario with a volume of 40 trains per day and 25% flexible trains will include 10 flexible trains operating with schedule flexibility and 30 scheduled trains following scheduled departures. Within the RTC simulation, higher priority is assigned to scheduled trains.

The combination of factors in the experiment design yields 150 different scenarios for simulation with RTC.

4 RESULTS

4.1 *Capacity versus Schedule Flexibility*

Train delays for the simulation scenarios in the experiment design were extracted to study the relationship between volume, schedule flexibility and traffic composition. To estimate line capacity, normalized average train delay values were plotted for each combination of traffic volume and schedule flexibility under a given traffic composition (Figure 2a). By setting a maximum allowable average train delay (minimum level of service for all trains) of 40 minutes per 160 train-kilometres, it is possible to estimate the line capacity as the maximum volume that can be supported at that level of service. In this manner, the trend lines in Figure 2a are used to construct a relationship between line capacity and schedule flexibility for a given traffic composition (Figure 2b).

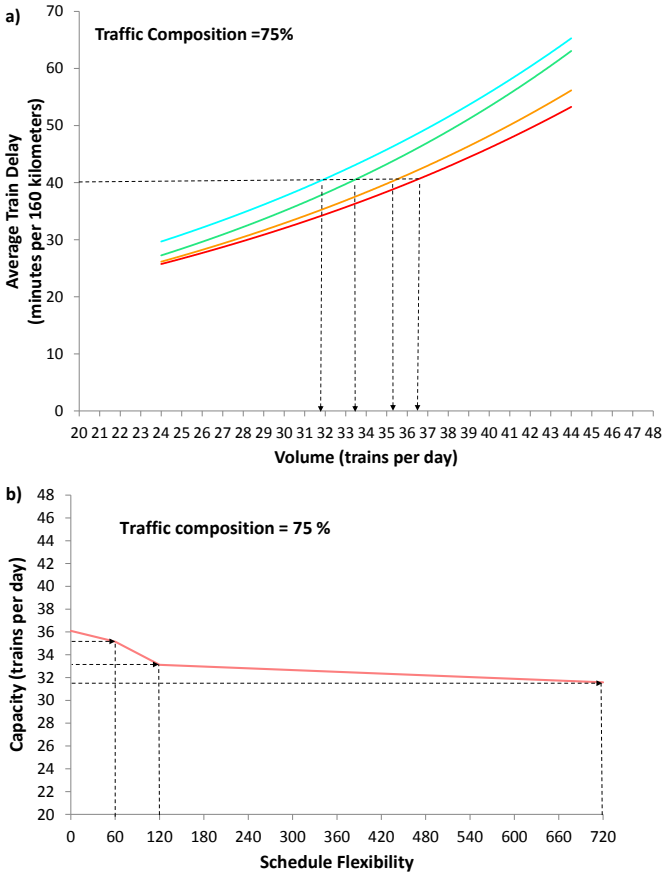


Figure 2. Example of capacity evaluation process a) Relationship between the average delay and train volume for different levels of schedule flexibility b) Capacity for a given schedule flexibility as defined by a maximum allowable train delay

The general appearance of Figures 2a and 2b follow the hypothesized relationships presented earlier (Figures 1b and 1c). For the specific case in Figure 2b, line capacity increases by approximately four trains per day when the flexible trains operating on the line transition from purely flexible to structured operations. This estimate of capacity only considers average train delay and not any train-type specific level-of-service requirements.

4.2 *Regression analysis of train-type performance*

To provide a comprehensive model of line capacity on the route under study, the general approach of Shih et al. (2015) was followed. A regression model with volume, traffic composition, and schedule flexibility as inputs and train delay specific to scheduled and flexible trains as an output was constructed (Equation 1). The model has an R-squared of 0.94 and significant interactions with p-values below 0.01. Thus the model should be a good predictor of the train delay associated with each type of train for the range of simulated factor levels.

$$D_t = f_t(c, SF) * V^2 + g_t(c, SF) * V + h_t(c, SF) \quad (1)$$

where c = traffic composition; SF = schedule flexibility; V = volume; g_t, f_t, h_t are functions representing first and second order functions of delay-volume function of a train type t; and D_t is the average normalized train delay for train type t.

To create an expression for line capacity, Equation 1 is set to equal a maximum allowable delay for each train type and then solved for the corresponding traffic volume using the quadratic formula (Equation 2). This volume corresponds to the capacity of the line as defined by the level of service required for a particular type of train.

$$V^* = \frac{-g_t(c, SF) + \sqrt{g_t^2(c, SF) - 4 * f_t(c, SF) * (h_t(c, SF) - D_t)}}{2 * f_t(c, SF)} \quad (2)$$

where D_t^* is the maximum allowable average normalized train delay for train type t; and V* is the line capacity as defined by the level of service of train type t.

Using Equation 2 and given train-type specific level-of-service requirements of 21 minutes of delay per 160 train-km for scheduled trains and 38 minutes of delay per 160 train-km for flexible trains, capacity can be estimated for a range of schedule flexibility and traffic composition values (Figure 3a and 3b). Since the transformation process has been applied using the maximum allowable delay values for each type of train, there will be two different capacity curves for each traffic composition. The final capacity for a particular level of schedule flexibility is the lowest of the two capacity values for that combination of schedule flexibility and traffic composition.

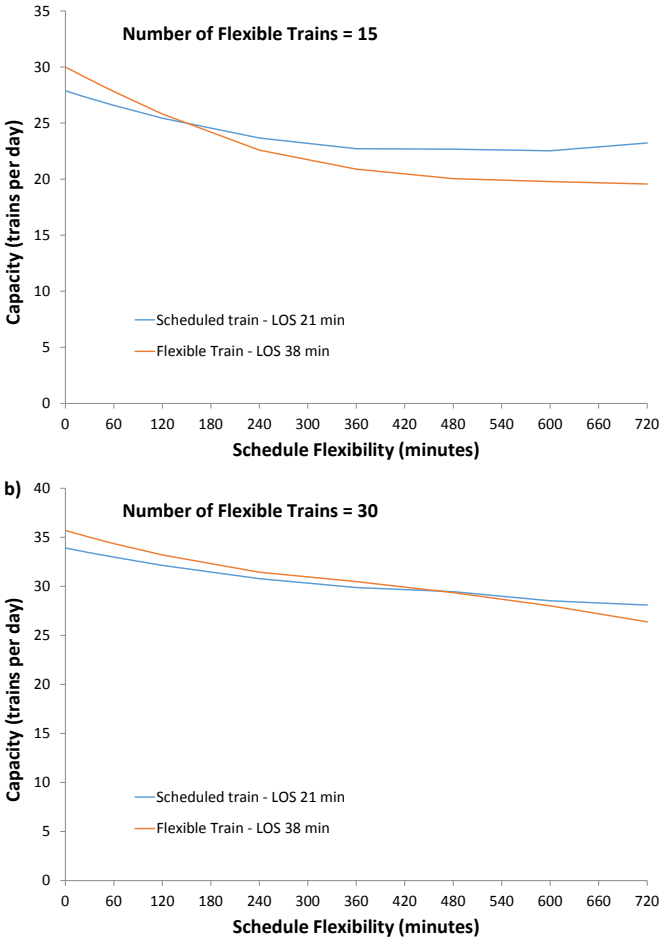


Figure 3. Final Capacity contours for scheduled and flexible trains a) 15 flexible trains b) 30 flexible trains

In general, from the relative positions of the two capacity contours (Figure 3a and 3b), the level-of-service requirements for scheduled trains defines the final contour at lower values of schedule flexibility. As schedule flexibility increases the final capacity is determined by the level of service of the flexible trains. Across the various traffic compositions, for the given train-specific levels of service, scheduled trains are less sensitive to changes in schedule flexibility compared to flexible trains.

From the perspective of the railway capacity planner, these results suggest that capacity is limited by the level of service of flexible trains if externalities and disruptions force the operations to become more flexible. However, if the operation moves toward structured operations, line capacity increases and the capacity is governed by the level of service of the scheduled trains.

4.3 *Comparison of traffic compositions*

Capacity values for different traffic compositions can be compared to determine the relative capacity

penalty for operating various levels of schedule flexibility on different traffic mixtures (Figure 4a and 4b). Each data series represents the capacity-schedule flexibility curve for a given traffic composition as defined by the level-of-service standard for flexible and scheduled trains.

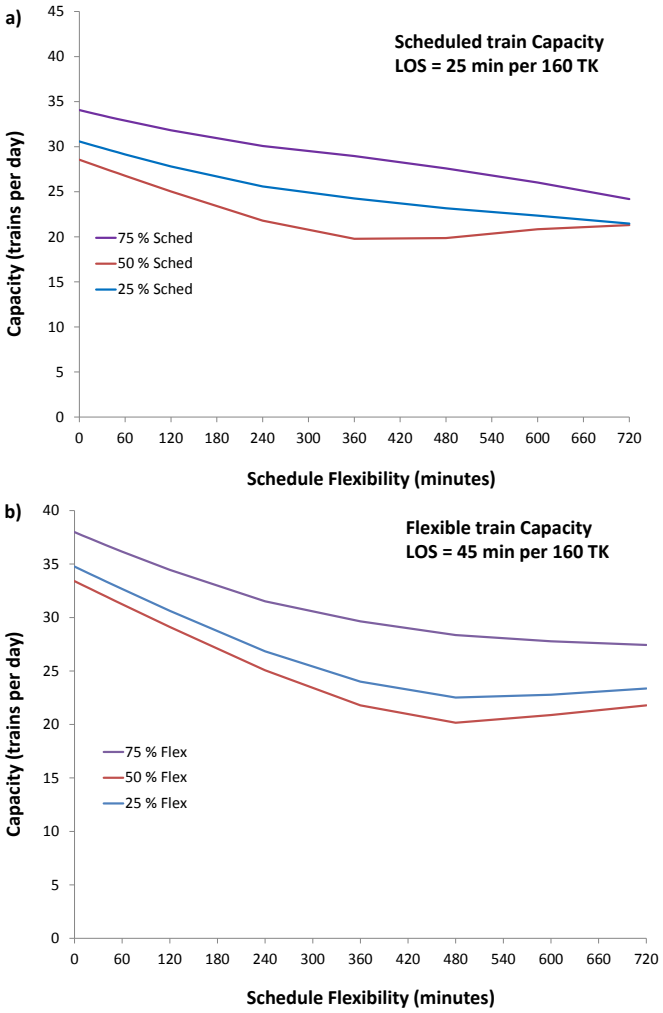


Figure 4. Capacity contours for various traffic compositions as defined by a) flexible train level of service and b) scheduled train level of service

Capacity contours defined by the level of service of flexible trains illustrate an inverse function with a decreasing slope that levels out at high values of schedule flexibility (Figure 4a). Capacity is sensitive to schedule flexibility values between 0 and 420 minutes; a move from structured operation to seven hours of schedule flexibility reduced capacity by approximately one third for all traffic mixtures. As schedule flexibility increases above 420 minutes, capacity becomes insensitive to increase in schedule flexibility.

Comparing traffic composition, the scenario with 75% flexible trains is operating at higher capacity than the 50% and 25% traffic compositions by about five trains per day. This result may be attributed to the smaller number of high-priority scheduled trains on the line that normally cause large amounts of delay to flexible trains. The lowest capacity is obtained for the case where the traffic composition has equal numbers of scheduled and flexible trains, consistent with the previous work on train heterogeneity described in Section 2.

Capacity contours defined by the level of service of scheduled trains (Figure 4b) follow a more linear trend with less overall sensitivity compared to those defined by flexible trains (Figure 4a). For the 25% and 75% traffic compositions, consistently increasing the schedule flexibility of flexible trains makes it more difficult to sustain the scheduled train level of service, forcing capacity to continually decline.

Overall, the results consistently indicate that decreases in schedule flexibility lead to increases in capacity for various traffic mixtures.

4.4 *Influence of level of service requirements*

Capacity for a traffic composition of 50% was plotted across a range of schedule flexibility for various train-type-specific levels of service (Figure 5). Equivalent capacity can be obtained by certain combinations of scheduled and flexible train levels of service. For instance, the capacity curve defined by a 65-minute level of service for flexible trains and the capacity curve defined by 45 minutes for scheduled trains belong to the same population with p = 0.05. At this level of capacity, the scheduled trains will have, on average, 20 minutes less delay per 160 train-km compared to the flexible trains. If an operator desires an equal 45-minute level of service for both train types, the capacity would decrease substantially to that defined by the 45-minute scheduled level of service contour.

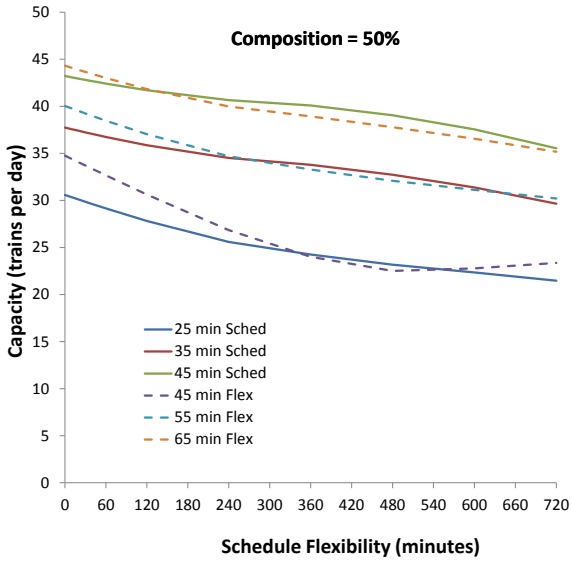


Figure 5. Capacity contours for various train-type-specific levels of service as defined by 50% composition

This change in capacity highlights the need to consider the level-of-service requirements of individual trains types when evaluating line capacity. If flexible heavy haul trains are more tolerant to delays than premium scheduled services, the specific level of service provided to scheduled trains will be a better metric for establishing line capacity. Average delay across all train types does not guarantee the performance of any particular type of trains and there may be a capacity penalty for one train type at the expense of another.

4.5 *Volume of flexible trains*

The previous sections have examined changes in capacity due to schedule flexibility under the assumption that traffic composition remains constant. In practice, a heavy haul operator can reduce both the schedule flexibility and the number of flexible trains when transitioning to structured operations. By examining the combinations of schedule flexibility and flexible train volumes that correspond to a given average train delay (level of service), the simulation data can be transformed to illustrate the relationship between number of flexible trains and capacity (defined by a given level of service) for various levels of schedule flexibility (Figure 6).

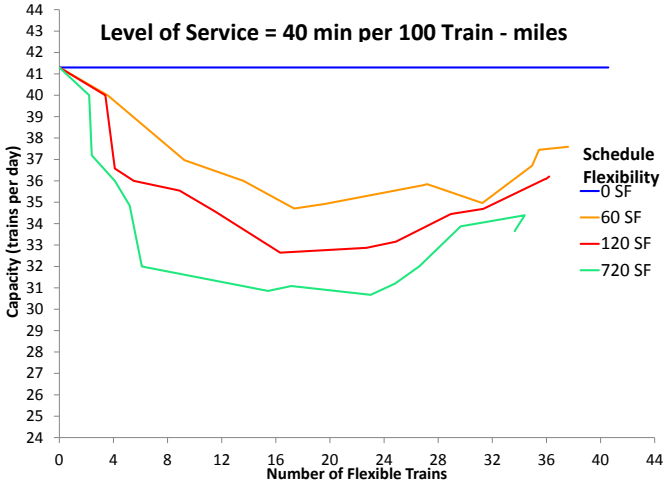


Figure 6. Capacity for various levels of schedule flexibility and number of flexible trains for a level of service of 40 minutes of train delay per 160 train-km.

Starting from a structured operation, capacity is sensitive to initial increases in the number of flexible trains. To maintain 40 minutes of train delay per 160 train-kilometres, the single-track route under study has a capacity of 41 scheduled trains per day if no flexible trains are operated. To replace twelve of the scheduled trains with flexible trains operating with 120 minutes of schedule flexibility, the capacity of the line must be reduced to approximately 34 trains per day to maintain the desired level of service. Alternatively, if the schedule flexibility of those twelve trains can be limited to 60 minutes, the capacity only drops to 36 trains per day. The capacity of 36 trains per day can also be achieved with a high schedule flexibility of 720 minutes but only if a maximum of four of these highly flexible trains are operated on the route. It is possible to increase capacity by moving to structured operations for most trains but still operating a small number of trains with high schedule flexibility. Adjusting both the number of flexible trains and schedule flexibility gives practitioners more options for maximizing the capacity of a line through structured operations while still accommodating the flexible schedule needs of selected heavy haul trains.

Figure 6 further illustrates the increasing magnitude of capacity reduction for increasing levels of schedule flexibility. Operators can see benefits in capacity from decreasing the number of flexible trains on the line. However, to achieve a specific increase in capacity, fewer flexible trains must become scheduled when operating at lower schedule flexibility compared to operations with higher schedule flexibility.

5 CONCLUSIONS

The research presented in this study used RTC simulation to analyse the relationship between traffic volume, schedule flexibility and traffic composition. For a given constant level of service and infrastructure, different traffic compositions follow a similar trend of increasing capacity with decreasing schedule flexibility. The largest capacity gains are made when moving from low levels of schedule flexibility to completely structured operation. Routes operating with a high degree of schedule flexibility might see little improvement in capacity until schedule flexibility is substantially decreased. Practitioners may adjust both the number of flexible trains and schedule flexibility to maximize the capacity of a line through structured operations while still accommodating the flexible schedule needs of selected heavy haul trains. While schedule trains will still experience train delay, it will typically be lower than the delay experienced by flexible trains. In defining capacity, it is important to consider the specific level-of-service requirements for each train type and not just the average delay over all scheduled and flexible trains

Future work will vary the infrastructure configuration and initial timetable to better understand relationships between infrastructure, schedule flexibility, initial timetable and line capacity. Operations that temporally separate scheduled and flexible trains may exhibit different capacity relationships.

6 ACKNOWLEDGMENTS

This research was supported by the National University Rail Center (NURail), a US DOT OST Tier 1 University Transportation Center, and the Association of American Railroads. The authors thank Eric Wilson and Berkeley Simulation Software, LLC for the use of Rail Traffic Controller simulation software. The authors also thank Mei-Cheng Shih, Graduate Research Assistant at the University of Illinois at Urbana-Champaign, for technical insight.

7 REFERENCES

Association of American Railroads. 2016. Freight Railroad Capacity and Investment. Washington DC.

Boysen, H.E. 2012. General model of railway transportation capacity. *WIT Transactions on The Built Environment, Vol 127 Computers in Railways XIII*: 335-346.

Boysen, H.E. 2013. Quicker Meets, Heavier Loads and Faster Empties – Effects on Transportation Capacity and Cycle Time. *Proceedings of 10th International Heavy Haul Association, Vol. 2*: 838-844.

Dingler, M.H., Koenig, A., Sogin, S.L., & Barkan, C.P.L. 2010. Determining the Causes of Train Delay. *Proceedings of the Annual AREMA Conference*. Orlando, FL, USA.

Dick, C.T. & Mussanov, D., 2016. Operational schedule flexibility and infrastructure investment: capacity trade-off on single-track railways. *Transportation Research Record: Journal of the Transportation Research Board*, 2546: 1-8.

Ede, B. M., Polivka, A., & Tunna, J. 2007. Optimizing Capacity by Trade-offs among Train Control, Infrastructure, Mechanical, and Engineering Considerations. *IHHA Specialist Technical Session (STS)*. Kiruna, Sweden.

Goverde, R. M. P., Corman, F, & D'Ariano, A. 2013. Railway line capacity consumption of different railway signaling systems under scheduled and disturbed conditions. *Journal of Rail Transport Planning & Management* Volume 3, Issue 3: 78–94

Kruger, H. 1999. Parametric Modeling in Rail Capacity Planning. *Proceedings of the 1999 Winter Simulation Conference*.

Lindfeldt, O. 2010. *Railway Operation Analysis: Evaluation of Quality, Infrastructure, and Timetable on Single and Double-Track Lines with Analytical Models and Simulation*. KTH Royal Institute of Technology, Department of Transport, Stockholm, Sweden. (Doctoral Dissertation).

Martland, C. 2003. Modeling the Effects of Heavy Haul Freight Service on Terminal Productivity. *Proceedings of International Heavy Haul Association, Implementation of Heavy Haul Technology for Network Efficiency, A Special Technical Session*. Dallas, TX, USA.

Martland, C.D. 2010. Improving on-time performance for long-distance passenger trains operating on freight routes. *Journal of the Transportation Research Forum*, 47(4): 63-80.

Mussanov, D., Nishio, N. & Dick, C.T. 2017. Delay Performance of Different Train Types Under Combinations of Structured and Flexible Operations on Single-Track Railway Lines in North America. *Proceedings of the International Association of Railway Operations Research (IAROR) 7th International Seminar on Railway Operations Modelling and Analysis*, Lille, France, April 2017.

Shih, M.-C., Dick, C.T., & Barkan, C.P.L. 2015. Impact of passenger train capacity and level of service on shared rail corridors with multiple types of freight trains. *Transportation Research Record: Journal of the Transportation Research Board*. 2475: 63-71.

Sogin, S., Lai, Y-C., Dick C.T. & Barkan, C.P.L. 2013. Comparison of capacity of single- and double-track rail lines. *Transportation Research Record: Journal of the Transportation Research Board*. 2374: 111-118.

Sogin, S., Lai, Y-C., Dick C.T. & Barkan, C.P.L. 2016. Analyzing the transition from single- to double-track railway lines with nonlinear regression analysis. *Journal of Rail and Rapid Transit*. 230 (8): 1877-1889

Effect of track conditions on the flexural performance of concrete sleepers on heavy-haul freight railroads

Z. Gao, M. S. Dersch, Y. Qian, & J. R. Edwards
University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT: Concrete sleepers have been widely used throughout the world as an alternative for timber sleepers. In North America, heavy-haul railroads have increased their use of concrete sleepers in recent years for a variety of factors. According to an international survey conducted by researchers at the University of Illinois at Urbana-Champaign (UIUC), railroad industry representatives consider center cracking to be one of the most common concrete sleeper failure mechanisms. Having a better understanding of sleepers' flexural behaviour can potentially reduce the occurrences of center cracking by ensuring both designs and maintenance practices are adequate for the field conditions. To measure the bending moments experienced in North American heavy-haul freight service, field experiments were conducted at three Class I railroad sites in North America under different track conditions. Concrete surface strain gauges were installed on concrete sleepers at each location to record bending strains experienced by the sleepers under the passage of trains. These strains were converted into moments using calibration factors determined either by calculations based on the sleepers' cross-sectional and material properties or by laboratory experimentation. This paper compares the measured moments from three test sites to analyse the effect of track conditions on the flexural performance of concrete sleepers.

1 INTRODUCTION

Throughout the world, the majority of railroad track infrastructure is supported by ballast. A ballasted track system typically consists of rail, fastening systems, sleepers, ballast, sub-ballast, and subgrade. The most commonly used material for sleepers in the United States is timber, which is used for approximately 90-95% of the sleepers in revenue service (Anonymous 2008). Concrete is the second most common material for sleepers, making up most of the remaining 5-10%. Typically, concrete sleepers are used in the most demanding service conditions (e.g. high curvature, steep grades, heavy tonnage, high speed passenger traffic, etc.).

The primary purpose of the sleeper in the overall track infrastructure system is to maintain track geometry (e.g. gauge, cross level, etc.) and to transfer applied wheel loads to the track substructure (Hay 1982). When a concrete sleeper supported on ballast is loaded vertically, the load is transferred from the wheel to the rail, fastening system, sleeper, ballast, sub-ballast, and subgrade, sequentially. The ballast support conditions play a critical role in the type and severity of bending that the sleeper will experience under loading from a passing train (Wolf et al. 2014). The ballast support is affected by a variety of factors that include loading during train operations, maintenance activities (i.e. tamping), fouling (i.e. intrusion of fine particles), and voids or gaps between the sleeper and ballast (Kaewunruen & Remennikov 2007).

Railroad operators, concrete sleeper manufacturers, and researchers from around the world participated in a survey and rated sleeper cracking from centre binding as the third most critical problem facing concrete sleepers (Van Dyk 2013). North American respondents considered centre cracking to be slightly less critical than their international counterparts, ranking it as the fifth most critical issue associated with concrete sleepers. However, North American respondents ranked cracking from dynamic loads as the third most critical issue, one place ahead of international respondents. This survey shows that sleeper cracking is a challenge experi-

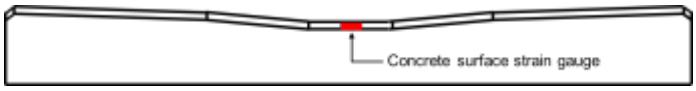
enced both domestically and internationally and thus an important issue and for research.

2 FIELD EXPERIMENTATION PLAN

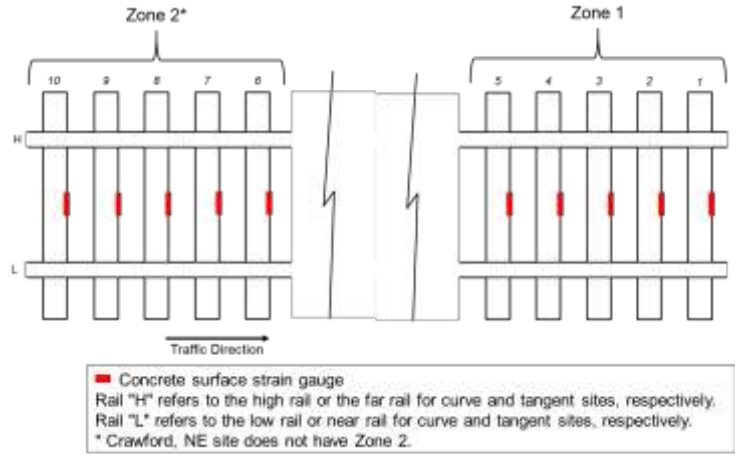
To measure the bending moments and support conditions experienced in North American heavy-haul freight service, field experimentation was conducted at three different locations of Class I heavy-haul freight railroads in the United States. The first field site was located on a tangent track in Ogallala, Nebraska (NE), on the Union Pacific Railroad's South Morrill Subdivision. The annual tonnage recorded in 2014 on this line was approximately 200 million gross tonnes (MGT). The second field site was located on a curved track in Norden, California (CA), on the Union Pacific Railroad's Roseville Subdivision. The location chosen consists of a curve with a 5°52' curvature on a 1.8% grade. Annual tonnage on this line was approximately 14 MGT at the time of installation (Holder et al. 2016). The final site was chosen on a curved track near Crawford, NE on the BNSF Railway's Butte Subdivision. This line is considered as one of the most demanding railroad lines in the United States due to its high curvature and high tonnage (Holder et al. 2016). The annual tonnage recorded in 2015 on this line was nearly 161 MGT. The field testing site consists of a 8° curvature on a 1.31% grade. Figure 1 shows the locations of the three field sites.

Figure 1. Field experimentation site locations

Researchers in the Rail Transportation and Engineering Center (RailTEC) at the University of Illinois at Urbana-Champaign (UIUC) have selected surface-mounted strain gauges to measure the bending strains experienced by concrete sleepers under revenue service heavy-haul freight train loads. Since centre cracking is considered as one of the most important issues with concrete sleepers, concrete surface strain gauges were installed at the centre of the sleepers within the field testing sites (Fig. 2a). For the Ogallala, NE and Norden, CA sites, the instrumentation was divided into two zones, with each zone consisted of five adjacent instrumented sleepers. For the Crawford, NE site, strain gauges were only installed on a total of five adjacent sleepers. Calibration factors, determined either by calculations based on the sleepers' cross-sectional and material properties or by laboratory experimentation, were applied to the recorded bending strains which converted the strains into bending moments. Loading configurations used for calibration tests were adapted from tests specified in Chapter 30, Section 4.9 in the American Railway Engineering and Maintenance-of-Way Association (AREMA) Manual for Railway Engineering (MRE) (AREMA 2016).



(a)



(b)

Figure 2. (a) Illustration of an instrumented sleeper (profile view), (b) experimentation layout for the Ogallala, NE and Norden, CA sites (Zone 1 was used at the Crawford, NE site)

3 PRILIMINARY RESULTS

The installation in Ogallala, NE occurred on March 27th, 2015. Since, 10 visits to the site were made with approximately 6 to 8 weeks between each visit. In total, 78 unit coal trains consisting of 43,284 loaded axles were collected. The installation in Norden, NE occurred on September 23rd, 2015 and data from 20 trains and 6,394 axles were collected between September 23rd and September 26th. Intermodal trains, passenger trains, mixed manifest trains, and empty trains were among the recorded data. The installation in Crawford, NE occurred on March 22nd, 2016, and similar to the Norden, NE site, only one site visit was made in which time a total of 11 train passes and 4,584 axles were collected. Among those train passes, there were 5 unit coal freight trains, 3 mixed manifest trains, and 3 empty trains. Overall, the instrumentation of all three sites were proven to be robust, as no surface strain gauge was damaged over the data collection period.

3.1 *Variations among three sites*

Measured bending moments were plotted versus their percentile exceeding (Figs. 3-5). For a given curve shown in these figures, each point represents the percentage of loaded axles that would cause a bending moment greater than or equal to a certain magnitude.

Figure 3 shows the bending moment distributions for different types of trains that were collected at the Norden, CA site. Intermodal trains and manifest trains shared comparable distributions, as at any magnitude of percent exceeding, the bending moment difference between those two train types was less than 1.0 kNm (8.9 kip-in). Since the majority of train axles passing through the site during the data collection period were from the manifest trains, the bending moment distribution of all the collected train passes was similar to that of the manifest trains. Overall, the measured bending moments never exceeded 22.7 kNm (201.0 kip-in), the AREMA recommended design limit for the centre of a concrete sleeper, which is shown as the vertical dash line on the graph.

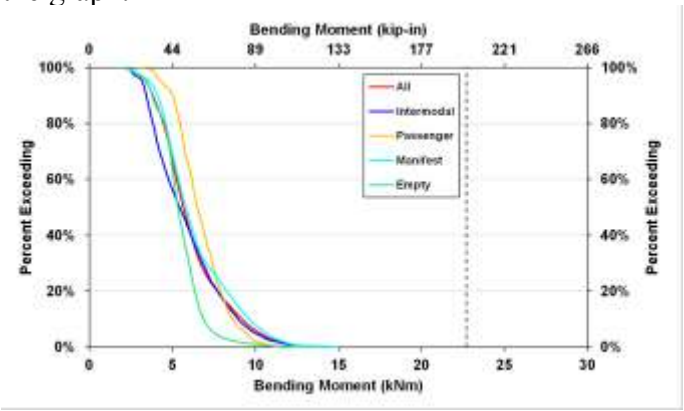


Figure 3. Bending moment variation with train type at the Norden, CA site

Bending moment distributions for all train types at the Crawford, NE site can be seen in Figure 4. Due to the limited amount of trains recorded at this site, the curves shown in Figure 4 are not as smooth as the curves shown in Figure 3. Among all train types, unit coal trains had the highest bending moment at any percent exceeding level, given that the axle load of a coal unit freight car was greater than that of a manifest or an empty car. In total, 4.5% of the measured bending moments exceeded the AREMA recommended design limit.

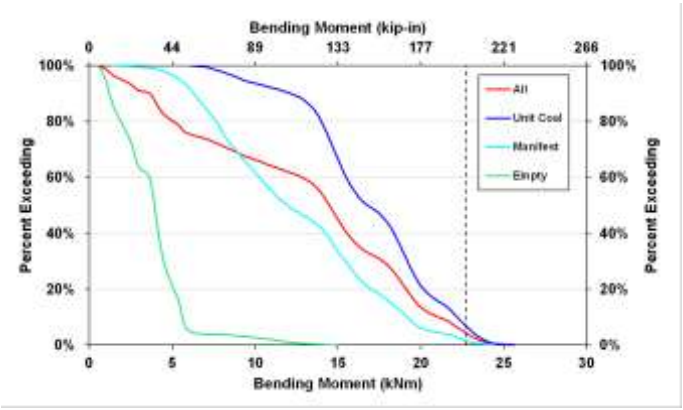


Figure 4. Bending moment variation with train type at the Crawford, NE site

Figure 5 shows the distributions of all the measured bending moments from all three sites. Bending moments experienced by sleepers at the Norden, CA site was the lowest among all three sites, indicating that the substructure (i.e. ballast, subballast, subgrade) of this site led to improved contact at the sleeper-ballast interface and thus lowered the bending moments. However, this indication might not be accurate considering the fact that the sleeper's bending moment is affected not only by the support condition, but also by the axle loads imposed onto the sleeper. Therefore, the low magnitude of bending moments recorded at the Norden, CA site could also be contributed by the low axle loads of the manifest trains that constituted the majority of trains passing through the site. In addition, the train speed at the Norden, CA site was considerably lower than those recorded at the other two sites; a few trains even stopped on top of the instrumented sleepers (Holder et al. 2016). This lower speed could potentially reduce the occurrence of high impact loads due to wheel defects, which would ultimately lead to sleepers experiencing lower bending moments.

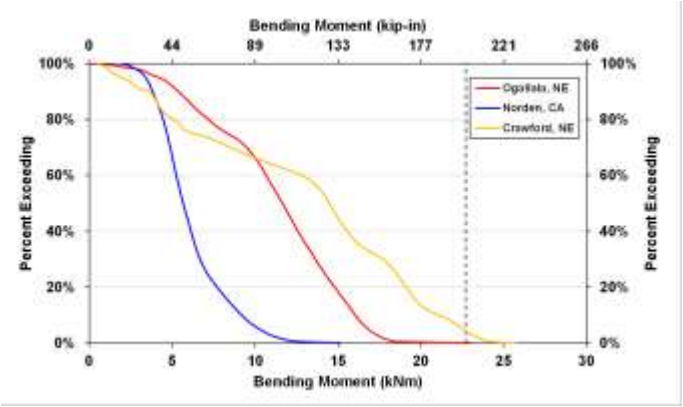


Figure 5. Bending moment variation with field experimentation site location

The bending moments measured at the Ogallala, NE site were lower than those measured at the Crawford, NE site between 0 and 65 percent exceeding. This is interesting to note given all trains recorded at

the Ogallala, NE site and the majority of trains recorded at the Crawford, NE site were unit coal trains, with a consistent nominal axle load of 320 kN (72 kips). The bending moment difference between these two sites under the same train type might suggest that the Ogallala, NE site had a more evenly distributed ballast at the sleeper-ballast interface. One possible explanation for the difference in the bending moments is that the high curvature at the Crawford, NE site could likely cause ballast underneath the sleepers to migrate towards the low rail. This redistribution of ballast would create an asymmetric centre-binding support condition for the sleepers, where the ballast was concentrated underneath the centre portion of the sleepers as well as the intermediate portion between the sleeper centre and the low rail.

Table 1 summarizes the key magnitudes of measured sleeper bending moments of all train types from all three sites. The maximum bending moment recorded from the unit coal trains at the Crawford, NE site was 2.6 kNm (23.0 kip-in) greater than the maximum recorded moment at the Ogallala, NE site, and the average bending moment was 5.7 kNm (50.4 kip-in) greater. Whether the high curvature characteristic of a railroad track could lead to these amounts of change in sleeper centre bending moment needs to be further investigated.

Table 1. Distribution of measured sleeper bending moments

Experimentation Site	Train Type	Sleeper Bending Moment kNm (kip-in)		
		Min	Mean	Max
Ogallala, NE	Unit Coal	0.5 (4.4)	11.2 (99.1)	23.1 (204.5)
Norden, CA	Intermodal	2.0 (17.7)	5.1 (45.1)	12.6 (111.5)
	Passenger	2.6 (23.0)	5.8 (51.3)	10.0 (88.5)
	Manifest	1.7 (15.0)	5.5 (48.7)	13.1 (115.9)
	Empty	1.7 (15.0)	4.7 (41.6)	11.7 (103.6)
Crawford, NE	Unit Coal	6.1 (54.0)	16.9 (149.6)	25.7 (227.5)
	Manifest	1.4 (12.4)	12.4 (109.7)	24.4 (216.0)
	Empty	0.5 (4.4)	3.8 (33.6)	15.0 (132.8)

3.2 *Variations with each site*

It is hypothesized that the varying ballast support conditions lead to the primary source of difference in the bending strains between adjacent sleepers. Therefore, bending moments among adjacent sleepers could be used to understand the variations in support conditions along the longitudinal direction of a railroad line. To visualize the distribution of measured bending moments of each sleeper within each site, box-and-whisker plots were developed (Figs. 6-8). The top line of the box represents the 75th percentile bending moment (Q3). The middle line is the median bending moment. The bottom line of the box represents the 25th percentile bending moment (Q1). The interquartile range (IQR), found as Q3 minus Q1, can provide an estimate of the variability of the data set – the greater the IQR, the higher the variability. The upper whisker shown on the graphs is the limit for upper outliers, which are defined as data points greater than Q3 plus 1.5 times the IQR (or (Q3 + 1.5*IQR)) (Ott & Longnecker 2001).

Figure 6 demonstrates the bending moment distribution for each of the ten instrumented sleepers at the Norden, CA site. In general, each sleeper exhibited a different distribution, but the differences among them were insignificant. On average, the mean bending moment difference between adjacent sleepers was 0.6 kNm (5.3 kip-in), while the maximum bending moment difference was 1.1 kNm (9.7 kip-in). The consistent moment distributions for all instrumented sleepers indicate that the ballast support condition along the longitudinal direction of the Norden, CA site was consistent as well. This uniform ballast distribution would allow the axle loads to be transmitted down into the substructure evenly along the longitudinal direction.

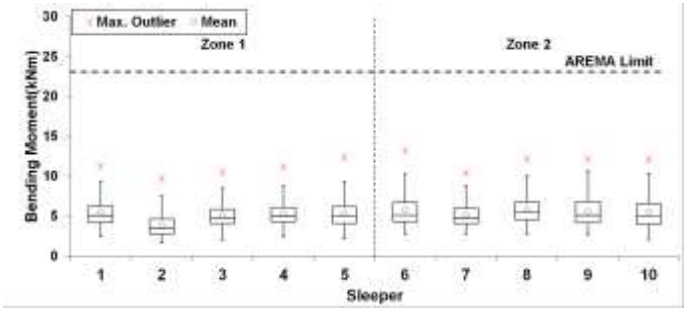


Figure 6. Box-and-whisker plot of measured bending moments at the Norden, CA site

Figure 7 illustrates the bending moment distribution for each instrumented sleeper at the Crawford, NE site. Unlike the Norden, CA site, the IQRs of bending moments at this site were more widely spread out, possibly due to the fact that there were fewer trains recorded at the site and 27% of the collected data were from empty trains. As mentioned in the previous sub-section, 4.5% of the total measured bending moments exceeded the AREMA limit, but based on Figure 7, it could be confirmed that all those exceedances occurred on a single sleeper (Sleeper #5), meaning that this sleeper was experiencing greater support at the centre. If no proper maintenance activity (e.g. tamping) was conducted around this sleeper, centre cracking might happen to

the sleeper which would eventually endanger the railroad safe operations.

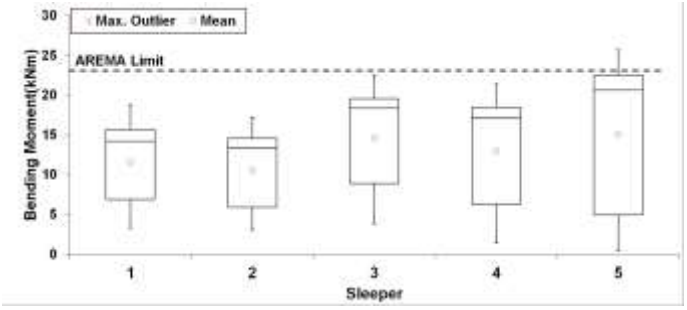


Figure 7. Box-and-whisker plot of measured bending moments at the Crawford, NE site

The support condition variations among instrumented sleepers at the Ogallala, NE site can be seen in Figure 8. The variability at this site was more significant than the other two sites. For instance, although Sleeper #9 and #10 were adjacent to one another, the centre support varied to the extent that Sleeper #9 experienced a bending moment that was nearly 6 kNm (53 kip-in) higher than Sleeper #10. The magnitude of the variation was over 26% of the AREMA recommended design capacity for the centre of concrete sleepers. It should be noted that Sleeper #4's bending moment outliers exceeded the AREMA recommended design limit value 3 times over the data collection period. That is, of the 43,284 loaded axles that passed over the sleeper, only 3 axles induced centre bending moments over the AREMA recommendation of 22.7 kNm (201.0 kip-in). The probability of exceedance was calculated to be 0.007% for Sleeper #4 and 0.0007% for all ten instrumented sleepers. Both of the probabilities were considered to be insignificant, thus indicating that the bending of the sleepers at this location would only cause centre moments to exceed the recommended values under very rare circumstances or if the support conditions were to vary more.

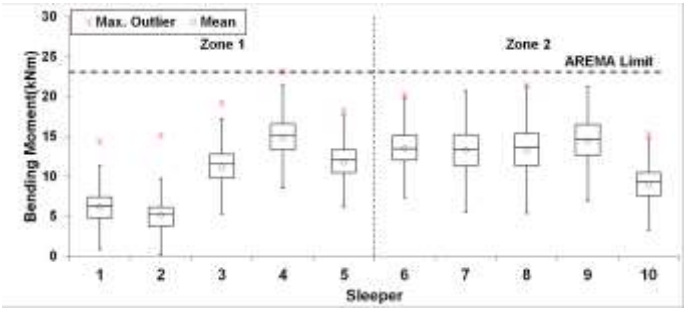


Figure 8. Box-and-whisker plot of measured bending moments at the Ogallala, NE site

4 CONCLUSIONS

Overall, bending strains, and subsequent moments, were successfully measured at three Class I heavy-haul freight railroad lines in North America. The effectiveness of surface-mounted concrete strain gauges in measuring sleeper bending behaviour was demonstrated. From this work, several conclusions were drawn relating to the flexural behaviour as well as support conditions of concrete sleepers at three field experimentation sites under revenue heavy-haul freight services:

- Measured center bending moments were highly variable among all three sites.
- Traffic differences at each site could partially account for the variations in bending moments. For the Norden, CA site, the majority of the recorded trains were manifest trains. The smaller magnitude of bending moments measured from this site could be caused by the axle loads of manifest trains being lower than those experienced from the other two sites.
- The bending moment variations among the three sites could also be a result of track conditions. The high tonnage and high curvature characteristics of the Crawford, NE site probably led to 4.5% of the measured bending moments exceeding the AREMA recommended design limit, posing a potential for increased deterioration of the concrete sleepers and the track substructure.
- Variations of support conditions existed along the longitudinal direction at each site. The Ogallala, NE site experienced significant variation in the bending moments along the longitudinal direction of track, in that the variation could be as much as 6 kNm (53 kip-in), or 26% of the AREMA recommended design capacity, between adjacent sleepers.

5 ACKNOWLEDGMENTS

Portions of this research effort were funded by U.S. National University Rail Center (NURail Center). Portions of the field experimentation were funded by Vossloh Fastening Systems. The authors would like to acknowledge the following industry partners: Union Pacific Railroad; BNSF Railway; National Railway Passenger Corporation (Amtrak); Progress Rail Services Corporation, a Caterpillar company; GIC; Hanson Professional Services, Inc.; CXT Concrete Ties, Inc., and LB Foster Company. The authors would also like to formally thank Steve Mattson of voestalpine Nortrak, Prof. Bill Spencer and Dr. Sihang Wei of UIUC, Prof. Dan Kuchma of Tufts University, and Prof. Fernando Moreu of the University of New Mexico for their knowledge in instrumentation, and Matt Csenge, Phanuwat Kaewpanya, and Álamo DiTarso for their assistance in the collection and processing of this data. The authors are also grateful for the advice and assistance provided by students and staff from RailTEC. J. Riley Edwards has been supported in part by grants to the

UIUC Railroad Engineering Program from CN and Hanson Professional Services.

REFERENCES

American Railway Engineering and Maintenance-of-Way Association (AREMA). 2014. Manual for Railway Engineering. Lancaster, MD.

Anonymous. 2008. M/W Budgets to Climb in 2008. Railway Track & Structures 104(1): 18–25.

Hay, W.W. 1982. Railroad Engineering, 2nd ed., John Wiley and Sons, New York, NY.

Holder, D.E., X. Lin, M.V. Csenge, Y. Qian, M.S. Dersch, and J.R. Edwards. 2016. Lateral Load Performance of SkI-Style Fastening System: Laboratory and Field Results. *Presentation at 2016 International Crosstie and Fastening System Symposium*, Urbana, IL, USA.

Holder, D.E., Y. Qian, M.V. Csenge, J.R. Edwards, M.S. Dersch, and B.J. Van Dyk. 2016. Comparison of Lateral Load Performance of Concrete Crosstie Heavy Haul Fastening Systems. *Presentation at 2016 AREMA Annual Conference and Exposition*, Orlando, FL, USA.

Kaewunruen, S. and A.M. Remennikov. 2007. Investigation of free vibrations of voided concrete sleepers in railway track system. Journal of Rail and Rapid Transit 221: 495-507.

Ott, R.L. and M. Longnecker. 2001. An Introduction to Statistical Methods and Data Analysis, 5th ed., Duxbury Press, Pacific Grove, CA.

Van Dyk, B. J. 2013. Characterization of Loading Environment for Shared-Use Railway Superstructure in North America. M.S. Thesis. University of Illinois at Urbana-Champaign, Urbana, IL, USA.

Wolf, H.E., S. Mattson, J.R. Edwards, M.S. Dersch, and C.P.L. Barkan. 2014. Flexural Analysis of Prestressed Concrete Monoblock Crossties: Comparison of Current Methodologies and Sensitivity to Support Conditions. In: Proceedings of the Transportation Research Board 94th Annual Meeting. Washington, DC, USA.

Quantitative prediction of the risk of heavy haul freight train derailments due to collisions at level crossings

S.G. Chadwick & C.P.L. Barkan
University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

M.R. Saat
Association of American Railroads, Washington, D.C., USA

ABSTRACT: The current methodology for prioritizing level crossing (LC) warning system upgrades and elimination in the United States (U.S.) focuses on the likelihood of collisions between highway and rail vehicles as well as highway user fatalities. However, these metrics do not encompass all LC risks. In particular, they do not consider the risk of derailment that LCs pose to trains, crews, and cargo (especially dangerous goods). Little previous research has considered this aspect of LC risk, although its impact is potentially severe. LCs have caused a number of train accidents in the U.S., including several that resulted in dangerous goods releases leading to injuries and fatalities. In this paper, we present a multi-factor statistical model that predicts the likelihood of a train derailment as a result of various LC parameters. The model was developed based on extensive data from the U.S. Department of Transportation's Federal Railroad Administration. It extends and formalizes previous work that identified factors leading to increased derailment likelihood for freight trains in LC collisions such as involvement of heavy highway vehicles (e.g. trucks/lorries) and higher train and motor vehicle speeds. The new model accounts for train and locomotive weight as additional factors to quantify derailment likelihood. The goal is development of a comprehensive understanding of the risk that level crossings pose to railroads and train operations.

1 INTRODUCTION

Highway-rail level crossing safety has been a topic of concern to railroads and the general public since the beginning of railroad construction. From 1991 to 2010, approximately 71,000 collisions occurred at public highway-rail level crossings in the United States (U.S.), including about 57,000 at publicly-accessible level crossings on mainline railroad tracks (FRA 2011a). Each collision has the potential to cause not only casualties to highway users, but also train passenger and crew casualties, property damage, or the release of dangerous goods. A number of serious level crossing collisions have occurred in recent years, resulting in casualties (NTSB 2015a, b, Associated Press 2015) and dangerous goods releases (NTSB 2014).

Since resources for level crossing improvements are finite, it is important to identify crossings that pose the greatest risk. Researchers and practitioners have devoted significant effort and resources to reducing risk to highway users. A variety of methods for modeling collision likelihood at level crossings have been developed, focusing on the risk trains pose to highway vehicles and their occupants, including the U.S. Department of Transportation Accident Prediction Model (FRA 1987; Ogden & Korve Engineering 2007) that is widely used in the U.S., and a variety of models developed to address limitations of that model (Benekohal & Elzohairy 2001, Austin & Carson 2002, Saccomanno et al. 2004, Oh et al. 2006, Washington & Oh 2006, Saccomanno et al. 2007). The results of these and other studies have led to improved level crossing warning systems, integration of level crossing operations with highway traffic signaling, public education programs such as Operation Lifesaver, and numerous other improvements in engineering and education (Mok & Savage 2005). These technologies and programs aim to reduce the number of casualties due to train-highway vehicle collisions, and the result has been a steady decline in the number of incidents and casualties over the past several decades.

Although the focus on level crossing safety has led to considerable improvements, one aspect has been largely overlooked – the risk that highway-rail level crossings pose to trains. Each year, 0.5 to 1% of level crossing collisions result in a train derailment. Even if a train does not derail, casualties can occur to passengers and crew aboard the train, and damage to the railroad track can result in lost service time and financial impacts. If the train does derail, there is additional potential for casualties among passengers and crew,

as well as the risk of a release if the train is carrying dangerous goods. With increased interest in passenger rail transportation and the growth in transportation of hazardous materials such as crude oil, the importance of comprehensive understanding of the risk of level crossing collisions is more critical than ever.

This paper presents a statistical model that enables quantitative assessment of the relative risk of different crossings to cause a derailment. Such a model enables more informal allocation of safety resources to minimize risk due to level crossings. This model could ultimately be integrated into an overarching risk analysis framework that would consider all sources of risk at a level crossing.

2 METHODOLOGY

2.1 *Dataset*

The U.S. Department of Transportation's Federal Railroad Administration (FRA) maintains two databases that were used to build the dataset for this study: the Rail Equipment Accident/Incident (REA) database, and the Highway Rail Accident (HRA) database. Data for all U.S. mainline railroads (both freight and passenger) during the 20-year period 1991 through 2010 were used to develop the model. It was validated using data from 2011 through 2014.

The REA database collects data on any damage sustained by a train consist that exceeds a reporting threshold set by the FRA. This threshold periodically changes to account for inflation and other adjustments; as of 2011 it was set at $9,400. These data are reported to the FRA using the FRA F 6180.54 form, filed by railroads that experienced an incident meeting this criterion. It provides useful information about incidents, such as incident cause, number of cars or locomotives derailed, length of consist, type of track involved, and a number of other variables of interest.

The HRA database collects data concerning "any impact, regardless of severity, between a railroad on-track equipment consist and any user of a public or private crossing site" (FRA 2011b). All level crossing collisions are reported to the FRA regardless of the monetary value of damage caused. The data are reported using form FRA F 6180.57. The database contains a variety of information including data about the type of highway vehicle involved, speed of the train at collision, and environmental factors such as time of day and weather conditions.

2.2 *Statistical method*

The statistical model presented in this paper was developed using the LOGISTIC procedure in the Statistical Analysis Software (SAS) computer package. This procedure uses the method of maximum likelihood to fit a linear logistic regression model to binary response data (SAS Institute 2013). In this way, the relationship between explanatory variables and the outcome responses can be analyzed. For the case of level crossing incidents, for each incident record the output of the model is a value between 0 and 1 representing the probability of a derailment occurring. Logistic regression is generally discussed in terms of "events" and "non-events"; in this paper, a derailment is an event, and an incident in which no derailment occurs is a non-event.

When logistic regression is used on data that has many more non-events than events, the regression will produce a poor fit even though there are indications of strong statistical relationships in the data. The models predict non-events correctly at the expense of predicting events, since this reduces the error rate. In this way, the model predicts a large percentage of all events correctly, but has poor fit because it fails to predict most derailment events. This problem can be remedied using a modified form known as "rare events logistic regression" (RELR) (King & Zeng 2001, van den Eeckhaut et al. 2006). RELR corrects for the disproportionate number of non-events by selecting a random subset of non-events equal to 1 to 5 times the number of events. In this case, a dataset was created containing a number of randomly-selected, non-derailment events equal to twice the number of derailment events.

Dick et al. (2001) define this as a "retrospective" model, as opposed to a "prospective" model. The retrospective model makes predictions about past events using a subset of the data, consisting of some number of events and some number of non-events. The output of this retrospective model must be calibrated to more accurately represent the probability of a derailment occurring in the overall population. While the factor coefficients from the small data set are equally valid for the large data set, the intercept term needs to be adjusted in the prospective model to account for the average rate of events in the actual population (Scott and Wild 1986). This adjusted "prospective" model can then be used to make predictions about unseen data.

2.3 *Model variables*

Six variables were selected as part of the modelling process (Table 1). Vehicle speed (VS) is the speed the highway vehicle was traveling at the time of collision, while train speed (TS) is the speed of the train at collision. Highway vehicle size (LV) differentiates between large highway vehicles such as semi-trucks (lorries) and small highway vehicles such as automobiles.

The FRA databases differentiate between level crossings of different "incident type" (IT). There are two defined incident types: incidents where the train strikes the vehicle (TSV) and incidents where the vehicle strikes the train (VST). Due to factors including the interaction between the train's wheels and the rail, the effects of the other model factors differ signifi-

cantly by incident type. This paper further distinguishes between incidents where the train strikes a stopped vehicle (TSV-S) and incidents where the train strikes a moving vehicle (TSV-M). About 43% of TSV incidents involve vehicles that are stopped on the crossing. It seemed plausible that this might mask the true effect of highway vehicle speed when conducting statistical regression. The same problem did not exist for VST incidents, since very few (less than five out of the whole database) trains involved in VST incidents are stopped at the time of collision.

Equipment class (EC) describes the type of rail vehicle that was struck by the highway vehicle. There are four types: freight car (FC), freight locomotive (FL), passenger car (PC), and passenger locomotive (PL).

Train length (TL) is the length of the train in number of rail vehicles. Rail vehicles include both locomotives and railcars.

Table 1. Definition of model variables

Variable Name	Definition	Variable Type	Range
VS	Highway vehicle Speed (mph)	Continuous	0-79
TS	Train speed (mph)	Continuous	0-106
LV	Was a large highway vehicle involved?	Binary	N if no; Y if yes
IT	Incident type	Categorical	VST, TSV-S, TSV-M
EC	Rail equipment class	Categorical	FC, FL, PC, PL
TL	Train length (rail vehicles)	Continuous	1-161

3 RESULTS

3.1 *VST incidents*

Of the derailment incidents reported in the REA database, 97 involved incidents in which the highway vehicle struck the train. To use RELR, 194 non-derailment incidents were randomly selected from the portion of the HRA database involving VST incidents. Combining 97 derailment and 194 non-derailment incidents creates a model dataset with a ratio of 1:2 events to non-events.

Initially, selection within the set of VST non-derailment incidents was done completely randomly. However, this resulted in selections that did not represent the true ratio of different rail vehicle types in the population because incidents involving passenger rail vehicles are so rare. Of the VST records, 30% involved a freight car, 64% involved a freight locomotive, 1% involved a passenger car, and 5% involved a passenger locomotive. Thus, 59 freight car incidents, 124 freight locomotive incidents, 2 passenger car incidents, and 9 passenger locomotive incidents were randomly selected to compile the model dataset of 194 non-derailment incidents. Repeating this process generated four different model datasets. Regression on each of them developed four models that performed similarly well and selected the same factors for the model, but one had the best fit statistics. This "best model" is:

$$p = \frac{1}{e^{-x}+1} \tag{1}$$

$$x_{VST} = -2.0204 + 0.0607\,VS + \begin{cases} 0, & LV = Y \\ -1.5459, & LV = N \end{cases} + \begin{cases} 1.8213, & EC = PC \\ 0.0648, & EC = FC \\ 0, & EC = PL \\ -1.3087, & EC = FL \end{cases} \tag{2}$$

This model provided the best fit to the data, with a Hosmer-Lemeshow (HL) goodness-of-fit test result of 0.7222. Values closer to 1 indicate good model fit, and values closer to 0 indicate poor fit. This model also has the ability to discriminate between derailment and non-derailment events, as measured by the area under the ROC curve. Generally, a model is considered to provide good discrimination if the ROC value is greater than 0.8. The area under the ROC curve for this result is 0.9011.

Additional performance statistics for this model are given in Table 2. For these values, the threshold value for predicting a derailment was a p value of 0.3. If the calculated value of p for a data point was greater than 0.3, it was classified as predicting a derailment, and if it was less than 0.3, it was classified as predicting no derailment.

Table 2. Performance statistics for retrospective model

Statistic	Value
Percent correct	81.8
Sensitivity	86.6
Specificity	79.5

In this model, the intercept term (b = -2.0204) is based on the average probability of a derailment for the RELR model dataset. This term needs to be adjusted in the prospective model to account for the average rate of derailment in the actual population of all level crossing collisions by altering the intercept term to account for the 20-year average likelihood of a VST derailment occurring. For the total VST population, the average derailment likelihood, $p_{all\ VST}$ can be calculated as:

$$p_{all\ VST} = \frac{97\ derailments}{7{,}040\ total\ events} = 0.0138 \tag{3}$$

The intercept term is modified to account for $p_{all\ VST}$ using the log-odds operator.

$$b_{VST} = b + \ln\left(\frac{p_{all\ VST}}{1-p_{all\ VST}}\right) \quad (4)$$

$$b_{VST} = -6.2912$$

Using the modified intercept term adjusts the probabilities predicted by the model to reflect the actual observed rate of derailments. Therefore, for all VST incidents, the final model is:

$$x_{all\ VST} = -6.2912 + 0.0607\,VS + \begin{cases} 0, & LV = Y \\ -1.5459, & LV = N \end{cases} + \begin{cases} 1.8213, & EC = PC \\ 0.0648, & EC = FC \\ 0, & EC = PL \\ -1.3087, & EC = FL \end{cases} \quad (5)$$

An ROC curve was generated by analysing the total population dataset with Equation 5 (Figure 1). The area under the ROC curve was equal to 0.9056. Additionally, model performance was quantified using the Brier score. This model had a Brier score of 0.0809; Brier scores closer to zero indicate better fit.

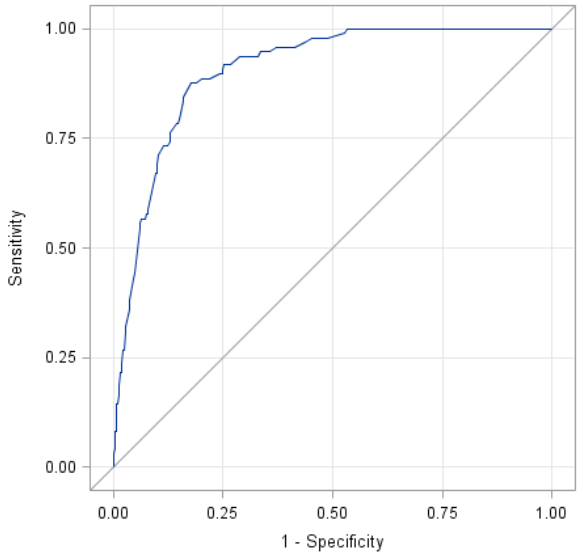


Figure 1. ROC Curve for dataset VST. Area under the ROC curve is equal to 0.9056.

In addition to these traditional techniques, the model was tested to see how it performed at ranking incidents by derailment likelihood, and whether this ranking corresponded to whether a derailment actually occurred. This technique has the advantage of being independent of the selected threshold value. To do this, all VST incidents in the HRA database were ranked by their $p_{all\ VST}$ value as calculated by the model, from least likely to most likely to derail. The dataset was divided into quintiles and the number of derailments in each quintile were counted (Table 3).

Table 3. Performance of VST model based on ranking

Quintile	Assigned Rank	Actual Derailments	Percent of Derailments
1	0 – 1,408	0	0
2	1,409 – 2,816	1	1.03
3	2,817 – 4,334	7	7.22
4	4,225 – 5,632	12	12.37
5	5,633 – 7,040	77	79.38

* Incidents in quintile 1 are least likely to derail, while incidents in quintile 5 are most likely to derail

Since approximately 80% of actual derailment incidents were ranked in the 5th quintile, the model does a good job of identifying derailment incidents. If, for example, level crossing decision makers ranked all crossings by derailment likelihood and chose to focus their efforts on the top 20%, they would likely capture 80% of all derailments.

3.2 *TSV-S incidents*

Of the derailment incidents reported in the REA database, 60 involved incidents where the train struck a stationary (VS = 0) highway vehicle. To use RELR, 120 non-derailment incidents were randomly selected from the portion of the HRA database involving TSV-S incidents. Combining 60 derailment and 120 non-derailment incidents results in a model dataset with a ratio of 1:2 events to non-events.

Selection within the set of TSV-S non-derailment incidents was random. The ratio of incidents involving freight and passenger rail vehicles was the same in the randomly selected development dataset as in the overall population. Approximately 11% of TSV-S incidents involved passenger trains. Unlike the VST case, it is not critical (and not possible) to differentiate between locomotives and railcars, because in TSV incidents less than a tenth of a percent (0.07%) involved a railcar. This is to be expected given that the vast majority of freight trains have a locomotive in the lead position. The dataset generation process was repeated to yield four different model datasets. Then a regression was run on each of them to develop four models. The four models performed similarly well and all selected the same factors, but one had the best fit statistics. This "best model" is:

$$x_{TSV-S} = -5.2729 + 0.0893\ \mathrm{TS} + 0.0362\ \mathrm{TL} - 0.00075\ \mathrm{TS} \times \mathrm{TL} + \begin{cases} 0, & LV = Y \\ -1.5459, & LV = N \end{cases} \quad (6)$$

This model had an HL test result of 0.8535 and an area under the ROC curve of 0.8688. Additional performance statistics for this model are given in Table 4 for a p value of 0.3.

Table 4. Performance statistics for retrospective model

Statistic	Value
Percent correct	78.2
Sensitivity	88.1
Specificity	73.3

The intercept term (b = *-5.2729*) needs to be adjusted in the prospective model to account for the average rate of derailment in the population of all grade crossing collisions. For the overall TSV-S population, the average derailment likelihood, $p_{all\ TSV\text{-}S}$ can be calculated as:

$$p_{all\ VST-S} = \frac{60\ derailments}{11{,}248\ total\ events} = 0.0053 \quad (7)$$

The intercept term is modified to account for $p_{all\ TSV\text{-}S}$ using the log-odds operator.

$$b_{TSV-S} = b + \ln\left(\frac{p_{all\ TSV-S}}{1-p_{all\ \mathrm{TSV-S}}}\right) \quad (8)$$

$$b_{TSV-S} = -10.5065$$

For all TSV-S incidents, the model is:

$$x_{all\ TSV-S} = -10.5065 + 0.0893\ \mathrm{TS} + 0.0362\ \mathrm{TL} - 0.00075\ \mathrm{TS} \times \mathrm{TL} + \begin{cases} 0, & LV = Y \\ -1.5459, & LV = N \end{cases} \quad (9)$$

The area under the ROC curve for the total population dataset was 0.8790, which is considered good discrimination (Figure 2). The model has a Brier score of 0.0892 indicating good fit.

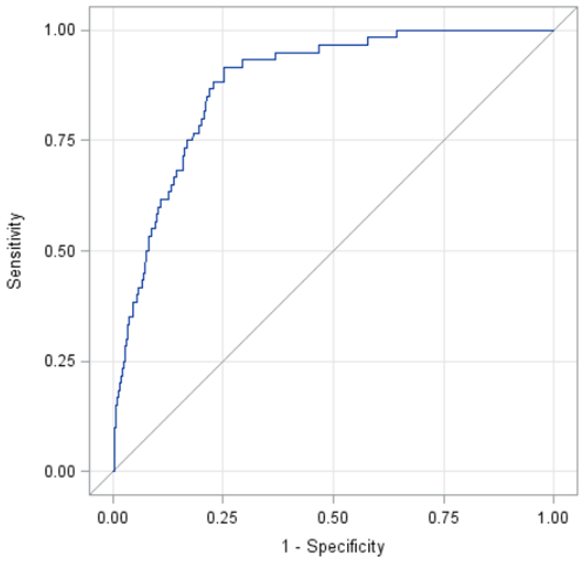


Figure 2. ROC Curve for dataset TSV-S. Area under the ROC curve is equal to 0.8790.

All TSV-S incidents were ranked by their $p_{all\ TSV\text{-}S}$ value as calculated by the model, from least likely to most likely to derail. The dataset was divided into quintiles to determine how many actual derailments occurred in each quintile (Table 5).

Table 5. Performance of TSV-S model based on ranking

Quintile	Assigned Rank	Actual Derailments	Percent of Derailments
1	0 – 2,261	0	0
2	2,262 – 4,522	1	1.67
3	4,523 – 6,783	2	3.33
4	6,784 – 9,044	10	16.67
5	9,045 – 11,305	47	78.33

* Incidents in quintile 1 are least likely to derail, while incidents in quintile 5 are most likely to derail

Since approximately 80% of actual derailment incidents were ranked in the 5th quintile, the model does a good job of identifying derailment incidents.

3.3 *TSV-M incidents*

Of the derailment incidents reported in the REA database, 114 involved incidents where the train struck a moving (VS > 0) highway vehicle. To use RELR, 228 non-derailment incidents were selected from the portion of the HRA database involving TSV-M incidents. Combining 114 derailment and 228 non-derailment incidents gave a model dataset with a ratio of 1:2 events to non-events.

Selection within the set of TSV-M non-derailment incidents was random. The ratio of incidents involving freight and passenger rail vehicles was the same in the randomly selected development dataset as the overall population. Approximately 11% of TSV-M incidents involved passenger trains. As with TSV-S incidents, it is not critical to differentiate between locomotives and railcars. The dataset generation process was repeated to generate four different model datasets, then a regression was run on each of them to develop four models. The four models performed similarly well and selected the same factors for the model, but one had the best fit statistics. This "best model" is:

$$x_{TSV-M} = -3.2144 + 0.0243\ \mathrm{VS} + 0.0233\ \mathrm{TS} + \begin{cases} 0, & LV = Y \\ -2.2628, & LV = N \end{cases} \quad (10)$$

This model provided the best fit to the data, with an HL goodness-of-fit test result of 0.9152 and an area under the ROC curve of 0.8438. Additional performance statistics for this model are given in Table 6 for a p value of 0.3.

Table 6. Performance statistics for retrospective model

Statistic	Value
Percent correct	74.1
Sensitivity	98.2
Specificity	61.5

The intercept term (b = *-3.2144*) needs to be adjusted in the prospective model to account for the average rate of derailment in the actual population of all

grade crossing collisions. For the total TSV-M population, the average derailment likelihood, $p_{all\ TSV\text{-}M}$ can be calculated as:

$$p_{all\ TSV-M} = \frac{114\ derailments}{15{,}027\ total\ events} = 0.0076 \quad (11)$$

The intercept term is modified to account for $p_{all\ TSV\text{-}M}$ using the log-odds operator.

$$b_{TSV-M} = b + \ln\left(\frac{p_{all\ TSV-M}}{1-p_{all\ TSV-M}}\right) \quad (12)$$

$$b_{TSV-M} = -8.0882$$

For all TSV-M incidents, the model is:

$$x_{all\ TSV-M} = -8.0882 + 0.0243\ VS + 0.0233\ TS + \begin{cases} 0, & LV = Y \\ -2.2628, & LV = N \end{cases} \quad (13)$$

The area under the ROC curve created by evaluating all records with $p_{all\ TSV\text{-}M}$ was 0.8625, considered good discrimination (Figure 3). The model has a Brier score of 0.1038 indicating good fit.

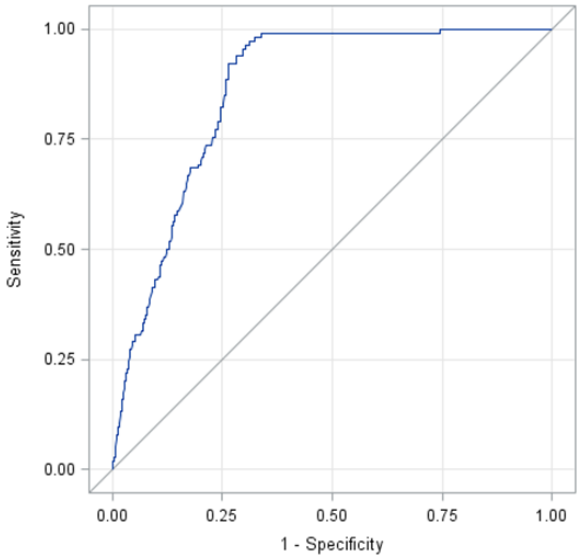


Figure 3. ROC Curve for TSV-M. Area under the ROC curve is equal to 0.8625.

TSV-M incidents in the model development dataset were ranked by their $p_{all\ TSV\text{-}M}$ value calculated by the model, from least likely to most likely to derail. The dataset was divided into quintiles and the actual derailments occurring in each quintile were counted (Table 6).

Table 7. Performance of TSV-M model based on ranking

Quintile	Assigned Rank	Actual Derailments	Percent of Derailments
1	0 – 3,006	0	0
2	3,007 – 6,012	1	0.88
3	6,013 – 9,018	0	0
4	9,019 – 12,024	0	0
5	12,025 – 15,027	113	99.12

* Incidents in quintile 1 are least likely to derail, while incidents in quintile 5 are most likely to derail

Since nearly 100% of actual derailment incidents were ranked in the 5th quintile, the model can be said to do a good job of identifying derailment incidents.

3.4 *Model validation*

To verify that the models developed based on data from 1991 to 2010 were valid for incidents outside the study period, data between 2011 and 2015 were tested to see where derailment incidents would be ranked. These results show that the model performs as well for more recent incidents as it did for incidents in the development dataset (Table 8).

4 DISCUSSION

4.1 *Interpretation of model terms*

Considered together, the models presented above contain five terms that indicate the effects of different vehicle and accident characteristics. A sixth characteristic, incident type, is accounted for by developing the three separate models. The SAS LOGISTIC procedure, using stepwise selection, chose three independent variables for the VST model, three independent variables and an interaction term for the TSV-S

Table 8: Performance metrics for validation dataset (2011-2015)

	VST (n = 1,150)		TSV-S (n = 2,578)		TSV-M (n = 2,553)	
AUC	0.9014		0.8935		0.8762	
Brier	0.0521		0.0855		0.0825	
Quintile	Actual Derailments	Percent Derailments	Actual Derailments	Percent Derailments	Actual Derailments	Percent Derailments
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	1	4.34	0	0	1	3.85
4	5	21.74	3	21.43	6	23.08
5	17	73.92	11	78.57	19	73.07

model, and three independent variables for the TSV-M model.

4.1.1 *VST incidents*

The first term in the VST model (Equation 5), $0.0607\,VS$, indicates that the speed of the vehicle at collision affects derailment likelihood. As vehicle speed increases, the probability of derailment also increases.

The second term in the model, $\begin{cases} 0, & LV = Y \\ -1.5458, & LV = N' \end{cases}$ indicates that the type of highway vehicle involved in the collision affects derailment likelihood. If the highway user is a small vehicle ($LGVEH = N$) then this term assumes a value of -1.5458; if the highway user is a large vehicle ($LGVEH = Y$) then the term disappears. This means that, all else equal, a collision where the highway user is a large vehicle is more likely to result in a derailment.

The third and final term in the model, $\begin{cases} 1.8213, & EC = PC \\ 0.0648, & EC = FC \\ 0, & EC = PL \\ -1.3087, & EC = FL \end{cases}$, shows the effect of equipment class on derailment likelihood. As the coefficients become more positive, derailment likelihood will increase. This means that incidents involving passenger cars are more likely to result in derailment than those involving freight rail, which in turn are more likely to result in derailment than those involving passenger locomotives, which in turn are more likely to result in derailment than those involving freight locomotives. This trend is what one would expect if lighter rail equipment is more likely to derail than heavier rail equipment.

It should be noted that the confidence intervals of the estimates for freight railcars and passenger locomotives overlap, meaning it is statistically uncertain if there is a difference between these two equipment classes. This overlap was observed in all four of the candidate models, suggesting it is not an artefact of the dataset. The overlap is probably explained by the fact that freight cars vary widely in weight. A loaded freight car can weigh five times as much as an empty one. When unloaded, the average freight car is lighter than the average passenger locomotive, but the opposite is true for loaded freight cars. Unfortunately, the HRA database does not track whether the railcar was loaded or empty. Therefore, it was not possible to distinguish between loaded and empty freight cars.

4.1.2 *TSV-S incidents*

The first term in the TSV-S model (Equation 9), $0.0893\,TS$, indicates that the speed of the train at collision affects derailment likelihood. As train speed increases, the probability of derailment also increases.

The second term in the model, $0.0362\,TL$, indicates that there is a relationship between train length and derailment likelihood. As the length of the train increases, so does derailment likelihood.

The third term in the model, $-0.00075\,TS \times TL$, is an interaction effect between train speed and train length. This indicates that at higher train speeds, the likelihood of derailment will decrease with longer train length; or, alternatively, for longer trains, the likelihood of derailment will decrease with higher speeds.

The final term in the model, $\begin{cases} 0, & LV = Y \\ -1.5733, & LV = N' \end{cases}$ indicates that the type of highway vehicle involved in the collision affects derailment likelihood. If the highway user is a small vehicle (LGVEH = N) then this term assumes a value of -1.5733; if the highway user is a large vehicle (LGVEH = Y) then the term disappears. *Ceteris paribus*, a collision where the highway user is a large vehicle is more likely to result in a derailment.

4.1.3 *TSV-M incidents*

The first term in the TSV-M model (Equation 13), $0.0243\,VS$, indicates that the speed of the vehicle at collision affects derailment likelihood. As vehicle speed increases, the probability of derailment also increases.

The second term in the model, 0.0233 TS, indicates that the speed of the train at collision affects derailment likelihood. As train speed increases, the probability of derailment also increases.

The final term in the model, $\begin{cases} 0, & LV = Y \\ -2.2628, & LV = N' \end{cases}$ indicates that the type of highway vehicle involved in the collision affects derailment likelihood. If the highway user is a small vehicle (LGVEH = N) then this term assumes a value of -2.2628; if the highway user is a large vehicle (LGVEH = Y) then the term disappears. This means that, *ceteris paribus*, a collision where the highway user is a large vehicle is more likely to result in a derailment.

4.2 *Model limitations*

As with any model, these findings are limited by the quantity and quality of data available. Derailments due to grade crossing collisions are uncommon events. Development of a reasonably-sized dataset of accidents required use of 20 years of data during which there were 399 verified derailments due to grade crossing incidents.

Overall, the quality of the data are good. There are some errors and inconsistencies between the REA and HRA databases, but in general it is a simple matter to identify and correct these errors. There are sufficient data that incomplete or internally-inconsistent records can be dropped if they cannot be corrected.

5 CONCLUSIONS

This paper explored the development of a set of models to predict derailment rates for both trains at highway-rail level crossings using logistic regression analysis. Three regression models were ultimately developed based on incident type – one each for incidents where the vehicle strikes the train, incidents where a train strikes a stopped vehicle, and incidents where a train strikes a moving vehicle. Results show that, other than incident type, five factors are important to derailment prediction: highway vehicle type, highway vehicle speed, train length, rail equipment type, and train speed. The key factors varied for each of the three regression models in ways that are consistent with expectations given the physical forces for each incident type. This model could ultimately be integrated into an overarching risk analysis framework that would consider all sources of risk at a level crossing.

6 ACKNOWLEDGMENTS

This research has been supported by ABS Consulting, the Association of American Railroads, and the National University Rail (NURail) Center, A U.S. DOT Tier-1 OST-R University Transportation Center. The first author has additionally been supported by the Eisenhower Transportation Fellowship Program, and a CN Graduate Research Fellowship in Railroad Engineering. The authors would like to thank Steve Laffey at the Illinois Commerce Commision for insight and feedback.

7 REFERENCES

Associated Press 2015. Passenger Train Collides With Truck in North Carolina. *The New York Times*. <http://www.nytimes.com/2015/03/10/us/passenger-train-collides-with-truck-in-north-carolina.html?_r=1> Accessed March 2016.

Austin, R.D. & Carson, J.L. 2002. An alternative accident prediction model for highway-rail interfaces. Accident Analysis & Prevention 34, 31-42.

Benekohal, R.F. & Elzohairy, Y.M. 2001. A New Formula for Prioritizing Railroad Crossings for Safety Improvement. In: Proceedings of the Institution of Transportation Engineers Annual Meeting, Chicago, IL.

Dick, C.T. 2001. *Factors Affecting the Frequency and Location of Broken Railway Rails and Broken Rail Derailments.* M.S. Thesis, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana.

Federal Railroad Administration (FRA) 1987. Rail-Highway Crossing Resource Allocation Procedure, 3rd ed. U.S. Department of Transportation, Washington, DC.

Federal Railroad Administration (FRA) 2011a. Download Accident Data On Demand. http://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx. Accessed March 2011.

Federal Railroad Administration (FRA) 2011b. *FRA Guide for Preparing Accident/Incident Reports*, FRA, U.S. Department of Transportation.

King, G. & Zeng, L. 2001. Logistic Regression in Rare Events Data. Society for Political Methodology.

Mok, S.C. & Savage, I. 2005. Why has safety improved at rail-highway grade crossings? Risk Analysis 25 (4), 867-881.

National Transportation Safety Board (NTSB) 2014. Highway Accident Report: Highway-Railroad Grade Crossing Collision: Rosedale, Maryland. < https://app.ntsb.gov/doclib/reports/2014/HAR1402.pdf> Accessed March 2016.

National Transportation Safety Board (NTSB) 2015a. Crash Summary Factual Report: Valhalla, NY. < http://dms.ntsb.gov/public/58000-58499/58036/583617.pdf> Accessed March 2016.

National Transportation Safety Board (NTSB) 2015b. Preliminary Report: HWY15MH006 (Grade Crossing Collision in Oxnard, California). < http://www.ntsb.gov/investigations/AccidentReports/Reports/HWY15MH006_preliminary.pdf> Accessed March 2016.

Ogden, B.D. & Korve Engineering 2007. Railroad-Highway Grade Crossing Handbook – Revised 2nd Edition. Office of Safety Design, Federal Highway Administration, Washington, DC.

Oh, J., Washington, S.P. & Nam, D. 2006. Accident prediction model for railway-highway interfaces. Accident Analysis & Prevention 38 (2), 346-356.

Saccomanno, F.F., Fu, L. & Miranda-Moreno, L.F. 2004. Risk-Based Model for Identifying Highway-Rail Grade Crossing Blackspots. Transportation Research Record 1862, 127-135.

Saccomanno, F.F., Park, P.Y.-J. & Fu, L. 2007. Estimating countermeasure effects for reducing collisions at highway-railway grade crossings. Accident Analysis & Prevention 39, 406-416.

SAS Institute Inc. 2013. The LOGISTIC Procedure. *SAS Help and Documentation*.

Scott, A.J. & Wild, C.J. 1986. Fitting logistic models under case-control or choice based sampling. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp.170-182.

Van Den Eeckhaut, M., T. Vanwalleghem, et al. 2006. Prediction of Landslide Susceptibility Using Rare Events Logistic Regression: A Case-Study in the Flemish Ardennes (Belgium). *Geomorphology* 76.

Washington, S. & Oh, J. 2006. Bayesian methodology incorporating expert judgment for ranking countermeasure effectiveness under uncertainty: Example applied to at-grade railroad crossings in Korea. Accident Analysis & Prevention 38, 234-247.

Semi-quantitative risk assessment of adjacent track accidents on shared-use rail corridors

Chen-Yu Lin, Christopher P. L. Barkan
University of Illinois at Urbana-Champaign, Urbana, Illinois, USA
Mohd Rapik Saat
Association of American Railroads, Washington D.C., USA

ABSTRACT: Adjacent track accidents (ATA) primarily refer to train accident scenarios in which derailed railroad equipment intrudes upon ("fouls") adjacent tracks and is potentially struck by another train on an adjacent track. ATA has been identified as one of the most important safety issues on shared passenger and freight railroad corridors. Various infrastructure, equipment and operational factors affect the probability and consequence of an ATA. The research described in this paper presents a comprehensive approach to identifying and evaluating various factors affecting the probability and consequence of ATAs. ATAs are divided into three sequential events: initial derailment, intrusion, and train presence on adjacent tracks. Each event is associated with a probability component. Factors affecting each probability component and consequence are identified and their effects are discussed. This research intends to depict a high-level overview of ATA risk and provides a basis for future quantitative risk analyses and risk mitigation measures.

1 INTRODUCTION

Adjacent track accidents (ATA) primarily refer to train accident scenarios in which derailed railroad equipment intrudes upon ("fouls") adjacent tracks and is struck by, or strikes another train on the adjacent track (Lin et al. 2016). Figure 1 depicts a typical sequence of events for an ATA. Under normal operation, the “equipment loading gauge” which defines the allowable height, width, and loads of rolling stock (referred to as the "clearance plate" in North America) of a train stays entirely within the clearance envelope (the clearance limits of civil infrastructure) of the track (Figure 1a). If a train derails, derailed equipment will generally exceed the clearance envelope of its own track (Figure 1b) and if the derailed equipment intrudes upon the adjacent track’s clearance envelope, it results in an "intrusion" (Figure 1c). When an intrusion occurs, another train on the adjacent track may be on, or approaching, the location of the intrusion location and potentially collide with the derailed equipment (Figure 1d).

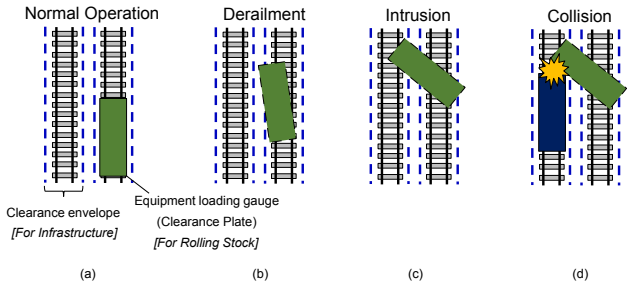


Figure 1. A Typical ATA Event Sequence (Lin et al., 2016)

A derailment without an intrusion may cause equipment and infrastructure damage, casualties, and operational disruption, while an intrusion may lead to more severe consequences due to the potential risk of subsequent collision. ATAs have been an emerging topic of railroad safety, especially due to actual or proposed growth in passenger rail services in the United States (Saat & Barkan 2013). As new passenger train services are introduced, overall train frequency will increase. Consequently, the probability of the presence of other trains on adjacent tracks increases. Furthermore, many existing or planned passenger train services involve shared-use of railroad tracks and/or right-of-way (ROW). A passenger train may derail, foul an adjacent track and then be struck by a freight train, or vice versa. Another concern is transportation of hazardous materials on a shared-use rail corridor because of the potential hazard if there is a release due to ATAs that may affect the safety of passengers in another train.

The research described in this paper presents a comprehensive approach to identifying and evaluating various factors affecting the probability and consequence of ATAs. ATAs are divided into three sequential events: initial derailment, intrusion and train presence on adjacent tracks. Each event has an associated set of probability components. Factors affecting each of these components and consequences are identified and their effects are discussed.

2 SEMI-QUANTITATIVE RISK ASSESSMENT

2.1 *Risk Model Development*

A common definition of risk is the multiplication of the probability or frequency of an event and the consequence of the event. It is commonly expressed as follows:

$$R=P\times C \quad (1)$$

where:
R: Risk
P: Probability
C: Consequence

The probability, P, is divided into three components corresponding to the three stages described above. ATA risk is thus defined as:

$$R=P(D)\times P(I|D)\times P(T|I|D)\times C \quad (2)$$

where:
R: The risk index for an ATA
P(D): The probability of an initial derailment on a multiple track section
P(I|D): Conditional probability of intrusion (CPI) given an initial derailment
P(T|I|D): Conditional probability of the presence of a train on adjacent track given an intrusion
C: The consequence of an ATA

There are three probability components and one consequence component in the model intended to calculate and compare the relative ATA risks on different track segments. A rating system was developed and each model component has five levels with corresponding values from 1 (the lowest) to 5 (the highest). To assess the risk for each track segment, infrastructure, rolling stock, train operating characteristics and any other relevant factors that affect the model components are evaluated to determine the probability and consequence levels. The overall ATA risk of a track segment can then be calculated using equation (2). In the following subsections, factors affecting each model component are introduced and discussed.

2.2 *Probability of Initial Derailment, P(D)*

The probability of an initial derailment can be estimated by analyzing previous train accident data. The United States Department of Transportation (U.S. DOT) Federal Railroad Administration (FRA) Rail Equipment Accident/Incident database contains train accident data as well as annual railroad traffic volume data in the United States (FRA 2011). Five factors affecting the probability of initial derailment are identified and discussed below: method of operation, track quality, traffic density, type of equipment, and rolling stock defect detection technology.

Method of operation

Method of operation indicates the presence of a wayside signal or automatic train control system. Previous research suggested that accident rate on signaled track segments is lower than non-signaled track segments (Liu et al 2017). In this study, track segments are classified as either signaled or non-signaled based on Liu et als' results (2017).

Track quality

The FRA classifies track quality into nine classes based on their construction and maintenance standards. Previous research has found an inverse relationship between track class and train derailment rate (Anderson & Barkan 2004, Liu et al. 2015). In this research, track classes are categorized into five groups. This categorization is based on their differences in train derailment rates (Liu et al. 2017). Track classes 6 and higher are grouped together because in general they are only used on lines that are primarily for passenger train operations. We are not aware of any quantitative analyses of derailment rates for these track classes but based on the research cited above, we presume that they are at least as low, and probably lower, than Class 5.

Traffic density

Traffic density is measured in annual gross tonnage in millions of gross tons (MGT) and is the total weight of all locomotives, rolling stock and lading operating on a particular segment of track. Higher traffic densities are correlated with lower derailment rates (Liu et al. 2017). The exact mechanism for this is not known but it may result from more frequent inspection, maintenance and frequency of wayside defect detection systems on high density rail lines. Dedicated passenger lines usually have lower derailment rates due to higher track maintenance standards and inspection frequency. In addition, the lighter axle loads of passenger equipment inflicts relatively less damage to the track structure, reducing the potential for accidents due to track defects. Thus, it is assumed that, *ceteris paribus*, dedicated passenger lines will have lower derailment rates.

Type of equipment

Failures of wheels, axles and other rolling stock components can cause derailments. Different compo-

nent designs may have differing failure rates. However, there is generally little quantitative data on how these may affect derailment rates. Further research is needed to address these potential effects. For the purposes of this research we identify it as an important factor but do not attempt to assign quantitative values.

Defect detectors and track inspections

Wayside defect detection technology is used to identify incipient flaws in various rolling stock components before they fail, thereby reducing the likelihood of a derailment. For example, Wheel Impact Load Detectors (WILD) are used to identify wheel defects that could lead to a mechanical failure (Van Dyk et al. 2013, Hajibabai et al. 2012). Similarly, various types of track inspections and technologies are used to identify incipient defects before they develop into a failure such as a broken rail, thereby reducing the likelihood of infrastructure-related derailments (Dick et al. 2003, Barkan et al. 2003, Liu et al. 2012, 2013a, 2013b, 2013c, 2014). Although it is well-accepted that these technologies and practices are effective at preventing derailments, the quantitative relationship between the use of a particular technology and its preventive effect has not be measured, so as is the case with type of equipment type above, we do not attempt to assign a quantitative value.

Levels of initial derailment probability are developed based on the aforementioned factors, except type of equipment, defect detectors and track inspections for the reasons discussed above. An Accident Factor Score (AFS) is assigned to each factor for a specific track segment (Table 1). The AFS ranges from 1 to 2 for each factor where the base value is 1. The higher the AFS, the higher the increase in initial derailment rate from that factor. For a track segment, all AFS are summed and based on the total AFS, a level of initial derailment probability is assigned to the track segment.

The effects of different factors on the initial derailment rate may vary. For instance, the effect of FRA track class on initial derailment rate may be different from the effect of traffic density. The effects of different levels within a factor may also differ. Take FRA track class as an example, the difference in train derailment rate between class 4 track and class 5 track is likely to differ from the difference of train derailment rate between class 5 track and class 6 track. Some of these relationships are addressed by quantitative analyses, while others are not fully understood. In our research, a linear approach is implemented for each affecting factor in which each factor has equal effect on the initial derailment rate, and each level within a factor also has equal impact on the initial derailment rate. For the purpose of consistency and simplicity, there are some underlying assumptions for the AFS: the effect of each factor is weighted equally, AFS for each factor is equally divided by the number of categories for the factor and the total AFS is equally divided into 5 levels.

Table 1. AFS and initial derailment probability

Initial Derailment Factors	**Criteria**	**AFS**
FRA Track Class	6 or above	1.00
	5	1.25
	4	1.50
	2, 3	1.75
	X, 1	2.00
Traffic Density	*Freight-Train only or Freight and Passenger Shared Lines*	
	More than 60 MGT	1.00
	40 - 60 MGT	1.33
	20 - 40 MGT	1.67
	Less than 20 MGT	2.00
	Passenger-Train only Lines	
	Passenger Line	1.00
Method of Operation	Signaled	1.00
	Non-Signaled	2.00
The highest score possible		**6.00**
The lowest score possible		**3.00**

Total AFS	**Level of initial derailment probability**
AFS ≤ 3.6	1
3.6 < AFS ≤ 4.2	2
4.2 < AFS ≤ 4.8	3
4.8 < AFS ≤ 5.4	4
AFS > 5.4	5

2.3 *Conditional Probability of Intrusion, P(I|D)*

Several factors affect intrusion probability including distance between track centers, track alignment and geometry, elevation differential, adjacent structures, containment, train speed, and point of derailment. In order to account for these factors in the model, an Intrusion Factor Score (IFS) is assigned to each factor for a track segment. The rationale and qualitative effect of each factor are discussed in the following sections.

Distance between track centers

Research by English et al. (2007) found an inverse relationship between the distance between track centers and probability of intrusion. Both developed the distribution of maximum lateral distance traveled by derailed rolling stock in accidents (Figure 2). In the model described in this paper, IFS for distance between track centers is assigned based on the 25th,

50th, 60th and 80th percentile from the cumulative probability distribution of lateral displacement.

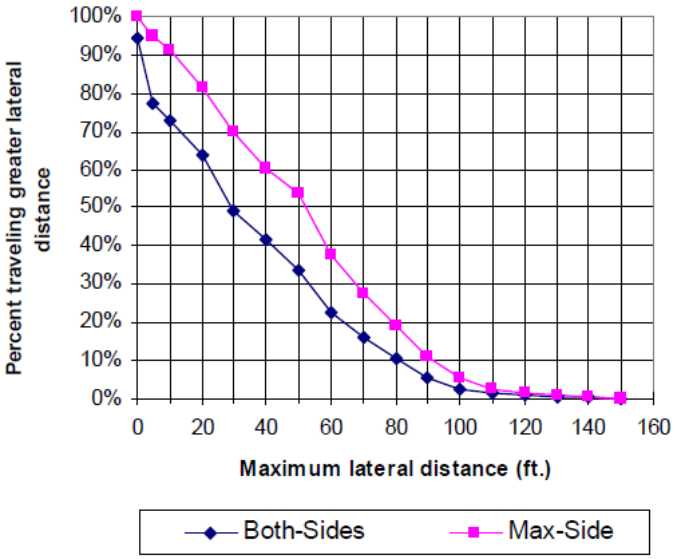


Figure 2. Distribution of maximum lateral displacement by derailed rolling stock (English et al. 2007)

Track alignment and geometry

Track alignment and geometry indicates whether a track segment is tangent, curved, level or on a grade. A level, tangent track segment is considered the base case scenario. If a derailment occurs on a curved track segment, additional lateral forces may be introduced that increase the intrusion probability. A derailment on a grade may affect longitudinal forces (draft or buff, respectively depending on whether the train is traveling up or down the grade at the time of the derailment) that indirectly affect intrusion probability. These longitudinal, in-train forces do not directly cause lateral movement of equipment; however, they may affect the extent that derailed rolling stock impacts other equipment in the train. These impacts may cause equipment to be moved laterally or rotate causing an intrusion on an adjacent track.

Elevation differential between adjacent tracks

If there is an elevational difference between two adjacent tracks then derailments occurring on the two tracks may have different intrusion rates. Specifically, derailments on the high track are more likely to intrude upon the lower track due to derailed equipment falling down the embankment (Figure 3a). Conversely, derailed equipment on the lower track, is less likely to intrude upon the higher adjacent track because of the constraining effects of the embankment (Figure 3b).

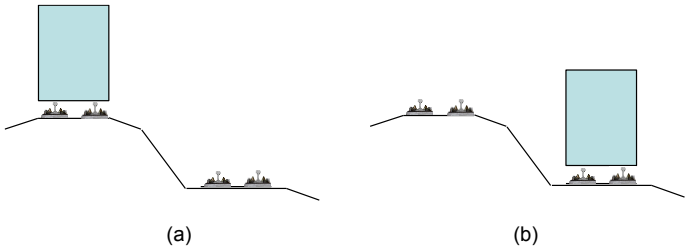


Figure 3. Effect of elevation differential

Adjacent structures

Adjacent structures refers to structure along the railroad segment that may have a "rebound" effect (Figure 4). If a structure is close enough to the railroad tracks and strong enough to potentially redirect the movement of derailed equipment from outside track back onto adjacent tracks opposite the track where the train derails, then its presence could affect intrusion probability. Adjacent structures, depending on their shape and density, are classified into single, discrete and continuous structures. A single structure is an independent, self-supported structure such as a bridge abutment or pier. Discontinuous structures could be multiple buildings located close to each other along a track segment, such as a group of grain elevators or silos. Examples of a continuous structure are noise barriers located alongside the track or residential buildings along the track in an urban area.

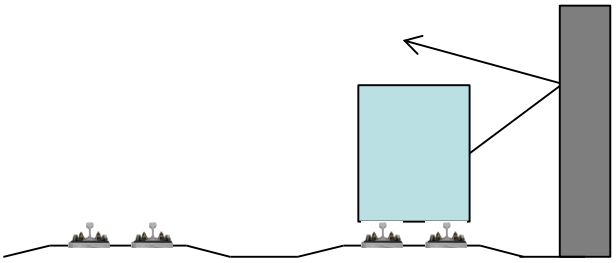


Figure 4 Effect of adjacent structure

Containment

Containment is located between adjacent tracks and its purpose is to prevent intrusions. Containment may also reduce the consequences by absorbing the energy from derailed equipment (further discussed in the Consequence subsection of this paper). Three types of containment are currently used in high-speed rail systems in Europe and Asia: guard rails, parapets, and physical barriers (Hadden et al. 1992, Moyer et al. 1994, Ullman and Bing, 1995, Rulens, 2008).

Guard rails or check rails are used for various related but distinct purposes. They are widely used in railroad turnouts and other special trackwork to ensure safe passage of rolling stock and minimize damage to track components. They are used in sharp curves to help keep equipment on the track. In the context of this study, guard rails are also used in trackage on, or leading up to, bridges and certain other special situations. In this case their purpose is to contain derailed equipment within or close to the

clearance envelope in order to prevent it from damaging the structural members of the bridge or adjacent tracks in case of a derailment (Figure 5). The latter type of guard rail is expected to reduce the intrusion rate.

Figure 5. An example of guard rail (Rulens 2008)

Parapets are reinforced railings mounted adjacent to the track (Figure 6) and physical barriers can be earth berms or concrete walls (Figure 7). Both are installed to absorb the impact of derailed equipment and prevent it from intruding upon adjacent tracks.

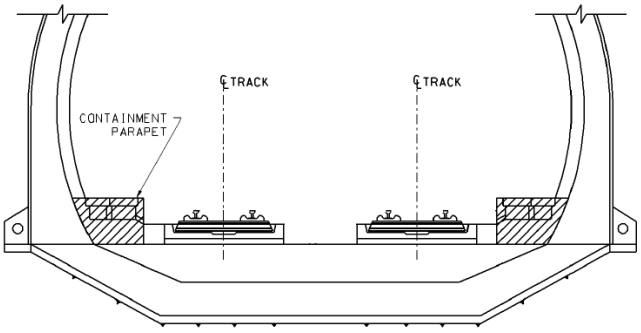


Figure 6. Example of a parapet (Rulens 2008)

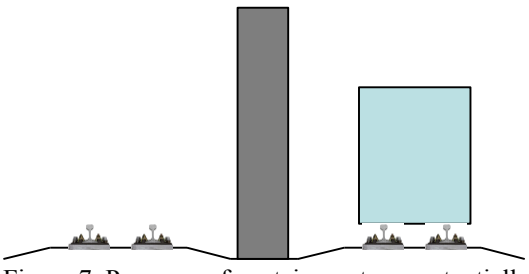


Figure 7. Presence of containment can potentially reduce the intrusion probability

Train speed

Speed of train may affect intrusion rate because the higher the speed, the more energy involved when a train derails, resulting in more opportunity for derailed equipment to move farther and foul adjacent tracks.

Train speed is assigned high, medium, or low to a track segment, based on the average speed on the segment. The average speed is affected by various factors, including type of traffic (bulk freight, intermodal, passenger, etc.), track alignment, track class, and so on. Two speeds are selected for categorization: 50 mph for key trains and 79 mph for maximum authorized speed of passenger trains on most U.S. passenger rail corridors. Key trains transport hazardous materials cars and their maximum authorized speed is 50 mph (AAR 2013).

Point of derailment

Point of derailment (POD) refers to the position-in-train of the first car derailed (Anderson, 2005; Liu et al., 2013a). The position of the first derailed car may affect intrusion rate due to reaction forces at the coupler. Also, because the first and the last car are only coupled at one end, they are less restrained with regard to lateral movement and might have more chance to rotate and foul adjacent tracks in a derailment. On the other hand, cars in the middle of the train consist are coupled at both ends, providing more restraining forces to the cars. However, the most common situation is when a single car in the middle of a train derails and causes other cars to derail, resulting in a larger derailment and intrusion.

Similar to AFS, IFS is assigned for each intrusion factor. The higher the IFS, the higher the increase in intrusion rate. Each factor has an IFS ranging from 1 to 2 where the base value is 1. For a track segment, IFS from all intrusion factors are summed. Based on the total IFS, a level of intrusion probability (from 1 to 5) is assigned to the specific track segment. The intrusion probability has the same assumption as the probability of initial derailment. Table 2 summarizes aforementioned intrusion factors except POD and associated IFS and the relationship between total IFS and corresponding levels of P(I|D). The higher the level, the more likely the occurrence of intrusion given an initial derailment.

Table 2. IFS and intrusion probability

Intrusion Factor	Criteria	IFS
Distance Between Track Centers, X, in feet (meters)	X > 75 (22.9)	1.00
	52 (15.8) < X ≤ 75 (22.9)	1.25
	40 (12.2) < X ≤ 52 (16.7)	1.50
	20 (6.1) < X ≤ 40 (12.2)	1.75
	X ≤ 20 (6.4)	2.00
Track Alignment	Tangent and level	1.00
	Tangent and on gradient when traveling upward	1.13

	Tangent and on gradient when traveling downward	1.25
	Curved on outside track and level	1.38
	Curved on inside track and level	1.50
	Curved on inside track and on gradient when traveling upward	1.63
	Curved on outside track and on gradient when traveling upward	1.75
	Curved on outside track and on gradient when traveling downward	1.88
	Curved on inside track and on gradient when traveling downward	2.00
Elevation Differential	The track where a train derail is 10 ft. lower than adjacent track	1.00
	The track where a train derail is level with adjacent track	1.50
	The track where a train derail is 10 ft. higher than adjacent track	2.00
Adjacent Structure	No adjacent structure	1.00
	Single structure	1.33
	Discrete structure	1.67
	Continuous structure	2.00
Containment	All containments installed	1.00
	Physical barrier and guard rail or parapet installed	1.20
	Physical barrier installed only	1.40
	Parapet and guard rail installed	1.60
	Parapet or guard rail installed only	1.80
	No containment installed	2.00
Train Speed	Low (less than 40 mph)	1.00
	Medium (40 mph to 70 mph)	1.50
	High (more than 70 mph)	2.00
The highest score possible		**12.00**
The lowest score possible		**6.00**

Total IFS	**Level of intrusion probability**
IFS ≤ 7.2	1
7.2 < IFS ≤ 8.4	2
8.4 < IFS ≤ 9.6	3
9.6 < IFS ≤ 10.8	4
IFS > 10.8	5

2.4 *Conditional probability of train presence on adjacent tracks, P(T|I|D)*

The third probability component of the ATA risk model is the presence of trains on adjacent tracks given an intrusion. There are two variants for the presence of train. One is that when an intrusion occurs, there is a train adjacent to the derailing equipment, and the other is that the train on the adjacent track is approaching the intrusion location. Factors affecting this probability include intrusion detection and warning systems, traffic density, method of operation, train speed, and shunting.

Intrusion detection and warning system (IDW)

The IDW system detects intruding rail equipment when it derails and breaks the fences installed with detectors between tracks, and changes the signal on either side of the adjacent track to stop (Hadden et al. 1992, Ullman & Bing 1995, Saat & Barkan 2013). Trains on adjacent tracks beyond the next block would have enough time to stop before striking the intruding equipment. However, IDW may not work if the train is already in the block where the intrusion occurs unless there are cab signals or other advanced train control system that transmits the information directly to the train and may allow it to stop before it encounters the intruding equipment.

Traffic density

The higher the traffic density, the more likely the presence of a train at the time of intrusion occurs. The traffic density of a track segment is measured using the gross tonnage of the track segment. The traffic density for dedicated passenger lines is assigned the highest level.

Method of operation

Different train control systems have different accuracy of train location as well as the ability to communicate the information between engineers (train drivers) and dispatchers. For example, the traditional track circuit system can only identify a train's location by "block" but does not provide the exact position of the train, whereas more advanced train control systems may be capable of identifying the trains' location more precisely. Representative systems include the European Rail Traffic Management System (ERTMS) in European countries and Advanced Train Administration & Communications System (ATACS) in Japan. Also, advanced train control systems communicate information more efficiently than traditional communication methods between dispatchers and engineers. IDW can also be integrated with advanced train control systems so that the intrusion warnings can be efficiently and instantly delivered to other trains in the proximity (Hadden et al. 1992, Ullman & Bing 1995).

In the model described in this study, train control systems are divided into three categories: advanced train control system, typical train control system and dark territory. Advanced train control systems refer to the track segment with these train control systems. Typical train control systems refer to track segments protected by track circuits. Dark territory refers to non-signaled track segments.

Train speed

If a train on an adjacent track is already in the block where initial accident and intrusion occur, typical train control systems may not be able to protect the train from striking the derailed equipment. If train speed is high or the distance is short, it may not be able to stop in time and will result in a collision. Train speed is assigned high, medium, or low to a track segment based on the average train speed of the adjacent track.

Based on engineering judgments, Train Presence Score (TPS) is assigned to train presence factors and are summarized in Table 3. Similar to the initial derailment probability and intrusion probability, each train presence factor has a TPS ranging from 1 to 2 where the base value is 1. The total TPS in a specific track segment is calculated by summing the TPS from all train presence factors. Total TPS is then converted into levels of train presence. The higher the level, the more likely the presence of a train given an intrusion. Although not all the combinations are considered, the selected factor combinations are assumed to be representative to account for most of the circumstances. TPS probability holds the same assumption as AFS and IFS.

Table 3. TPS and train presence probability

Train Presence Factors	Criteria	TPS
IDW	Presence	1.00
	Absence	2.00
Traffic Density	*Freight or Freight and Passenger Shared Lines*	
	Less than 20 MGT	1.00
	20 - 40 MGT	1.33
	40 - 60 MGT	1.67
	More than 60 MGT	2.00
	Passenger Lines	
	Dedicated Passenger Line	2.00
Method of Operation	Advanced train control	1.00
	Typical train control system	1.50
	Dark territory	2.00
Average Train Speed	Low (less than 50 mph)	1.00
	Medium (50 mph to 79 mph)	1.50
	High (more than 79 mph)	2.00
The highest score possible		**8.00**
The lowest score possible		**4.00**

Total TPS	Level of train presence probability
TPS ≤ 4.8	1
4.8 < TPS ≤ 5.6	2
5.6 < TPS ≤ 6.4	3
6.4 < TPS ≤ 7.2	4
TPS > 7.2	5

2.5 *Overall Probability, P*

The three probability levels are multiplied into a single score to represent the overall probability:

$$P=P(D)\times P(I|D)\times P(T|I|D) \quad (3)$$

Based on the values of P, a level of overall probability will be assigned. Table 4 shows the relation between the value of P and the level of overall probability.

Table 4. Overall probability level

Multiplication of P(D), P(I|D), and P(T|I|D)	Overall Probability Level, P
1 < P ≤ 10	1
10 < P ≤ 20	2
20 < P ≤ 30	3
30 < P ≤ 50	4
P > 50	5

2.6 *Consequence, C*

Consequence is the impacts from an ATA. The major concern is the consequence resulting from the collision between derailed equipment and trains on adjacent tracks. Previous research showed that the average casualties for passenger train collisions are higher than for passenger train derailments (Lin et al. 2013). The consequences of ATAs include multiple types of impact as follows:

- Casualties (injuries and fatalities)
- Equipment damage
- Infrastructure damage
- Non-railroad property damage
- System disturbance and delay
- Environmental impact
- Economic loss

Casualties refer to passenger and non-passenger fatalities or injuries from accident impact, and/or casualties due to exposure to hazardous materials release in an ATA involving a freight train transporting hazardous materials. Equipment damage is the cost required to repair rail cars. Infrastructure damage is the cost required to replace damaged track structure. Non-railroad property damage includes the non-railroad structure damaged by the impact of derailed equipment or explosion. System disturbance and delay resulted from the derailment is measured by system shutdown time and the number of trains affected. Environmental impact refers to environmental damage due to the release of fuel or hazardous materials. Economic loss refers to the damage or release of the lading being carried by freight cars. Several factors are identified to affect the severity of ATA accidents: speed of train, equipment strength, containment, and product being transported. These factors are discussed in the following subsections.

Equipment strength

Equipment strength is a key factor for reducing the potential casualties on board from the derailment and/or collision impact. Crashworthiness analyses have been conducted for higher-speed passenger trains (Tier I standard) (Carolan et al. 2011) to understand how reinforced equipment can withstand larger collision impact and thus result in lower consequences. Rolling stock is classified into two categories: reinforced equipment and traditional equipment. Reinforced equipment refers to passenger rail cars that meet the FRA Tier I or higher crashworthiness regulations, or tanks that are equipped with top fitting

protection, jacket, and couplers that prevent them from overriding other rail cars. Traditional equipment refers to railcars that do not meet the aforementioned standards.

Train speed
With higher train speed, more energy will be involved when a derailment or collision occurs. Research shows that train speed may affect the consequence of an accident (Liu et al., 2011). Therefore, more severe consequence are expected if the train speed is higher.

Containment
The presence of containment reduces not only the conditional probability of intrusion but also the consequence by absorbing the impact from the derailing equipment (Hadden et al., 1992; Moyer et al., 1994; Ullman and Bing, 1995).

Product Being Transported (Freight Train)
If the collision involves freight trains carrying hazardous material (or dangerous goods), it may release the hazardous material and result in more severe consequences.

Similar to the way probability components of ATA are calculated, the Consequence Factor Score (CFS) is assigned to different situations for each consequence factor (Table 5). The total CFS is calculated by summing the CFS from individual consequence factor together. The total CFS is then related to the level of consequences.

Table 5. CFS and consequence level

Consequence Factor	**Criteria**	**Consequence Factor Score (CFS)**
Equipment Strength	Reinforced equipment	1.00
	Traditional equipment	2.00
Train Speed	Low (less than 40 mph)	1.00
	Medium (40 mph to 70 mph)	1.50
	High (more than 70 mph)	2.00
Containment	Containment Present	1.00
	No Containment	2.00
Product being transported	No Hazardous material	1.00
	Hazardous material	2.00
The highest score possible		**8.00**
The lowest score possible		**4.00**

Total CFS	**Level of Consequence**
CFS ≤ 4	1
4 < CFS ≤ 5	2
5 < CFS ≤ 6	3
6 < CFS ≤ 7	4
CFS > 7	5

3 CONCLUSION

The research described in this paper presents a comprehensive approach to evaluating the ATA risk and identifying factors affecting the probability and consequence of ATA. Levels of probability and consequences are defined. Various factors affecting the initial accident, the intrusion, the presence of trains on adjacent tracks as well as the consequences are identified and investigated. This research intends to depict a high-level overview of ATA, and provides a basis for future quantitative risk analyses and risk mitigation implementations.

4 ACKNOWLEDGMENTS

This research was funded by the National University Rail (NURail) Center, a U.S. DOT University Transportation Center.

5 REFERENCES

American Association of Railroads (AAR) 2013. Recommended Railroad Operating Practices For Transportation of Hazardous Materials. Publication Circular OT-55-N. American Association of Railroads, Washington, D.C.

Anderson, R.T. & Barkan, C. P. L. 2004. Railroad accident rates for use in rail transportation risk analysis. *Transportation Research Record - Journal of the Transportation Research Board 1863*: 88 – 98.

Anderson, R.T. 2005. Quantitative Analysis of Factors Affecting Railroad Accident Probability and Severity. M.S. Thesis. University of Illinois at Urbana-Champaign, Urbana, Illinois.

Barkan, C.P.L., Dick, C.T. & Anderson, R. T., 2003. Analysis of railroad derailment factors affecting hazardous materials transportation risk. *Transportation Research Record - Journal of the Transportation Research Board* 1825, 64 – 74.

Carolan, M., Jacobsen, K., Llana, P., Severson, K., Perlman, B. & Tyrell, D. 2011. Technical Criteria and Procedure for Evaluating the Crashworthiness and Occupant Protection Performance of Alternatively Designed Passenger Rail Equipment for Use in Tier I Service. Publication DOT/FRA/ORD-11/22. U.S. Department of Transportation, Washington, D.C.

Dick, C.T., Barkan, C.P.L., Chapman, E.R. & Stehly, M.P. 2003. Multivariate statistical model for predicting occurrence and location of broken rails. *Transportation Research Record - Journal of the Transportation Research Board* 1825, 48 – 55.

English, G.W., Highan, G. & Bagheri, M. 2007. Evaluation of Risk Associated with Stationary Dangerous Goods Railroad Cars. TranSys Research Ltd.

Federal Railroad Administration (FRA). 2011. FRA Guide for Preparing Accident/Incident Reports, U.S. Department of Transportation, Washington, D.C.

Hadden, J., Lewalski, W., Kerr, D. & Ball, C. 1992. Safety of High-Speed Guided Ground Transportation Systems: Shared Right-of-Way Safety Issues. Publication DOT/FRA/ORD-92/13. Federal Railroad Administration, U.S. Department of Transportation, Washington, D.C.

Hajibabai, H., Saat, M.R., Ouyang, Y., Barkan, C.P.L., Yang, Z., Bowling, K., Somani, K., Lauro, D., & Li, X. 2012.

Wayside Defect Detector Data Mining to Predict Potential WILD Train Stops. Presented at the Annual Conference and Exposition of the American Railway Engineering and Maintenance-of-Way Association (AREMA), Chicago, Illinois.

Lin, C.Y., Saat, M.R. & Barkan, C.P.L. 2013. Causal Analysis of Passenger Train Accident on Shared-Use Rail Corridor. In Proceedings of the World Congress on Railway Research, Sydney, Australia.

Lin, C.Y., Saat, M.R. & Barkan, C.P.L. 2016. Fault tree analysis of adjacent track accidents on shared-use rail corridors. *Transportation Research Record - Journal of the Transportation Research Board* 2546: 129 – 136.

Liu, X., Barkan, C.P.L. & Saat, M.R. 2011. Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. *Transportation Research Record - Journal of the Transportation Research Board* 2261: 178 – 185.

Liu, X., Saat, M.R., Qin, X. & Barkan, C.P.L. 2013a. Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression. *Accident Analysis and Prevention* 59: 87 – 93.

Liu, X., Saat, M.R. & Barkan, C.P.L. 2013b. Integrated risk reduction framework to improve railway hazardous materials transportation safety. *Journal of Hazardous Materials* 260: 131 – 140.

Liu, X., Saat, M.R. & Barkan, C. P. L. 2013c. Safety effectiveness of integrated risk reduction strategies for rail transport of hazardous materials. *Transportation Research Record - Journal of the Transportation Research Board* 2374: 102 – 110.

Liu, X., Lovett, A., Barkan, C.P.L., Dick, C.T. & Saat, M. R. 2014. Optimization of rail defect inspection frequency for the improvement of railway transportation safety and efficiency. *ASCE Journal of Transportation Engineering* 140(10): 04014048.

Liu, X., Saat, M.R. & Barkan, C.P.L. 2017. Freight-train derailment rates for railroad safety and risk analysis. *Accident Analysis and Prevention* 98: 1 – 9.

Moyer, P.D., James, R.W., Bechara, C.H. & Chamberlain, K. L. 1994. Safety of High Speed Guided Ground Transportation Systems Intrusion Barrier Design Study. Publication DOT/FRA/ORD-95/04. Federal Railroad Administration, U.S. Department of Transportation, Washington, D.C.

Rulens, D. 2008. Rolling Stock and Vehicle Intrusion Protection for High-Speed Rail and Adjacent Transportation Systems TM 2.1.7. Parsons Brinckerhoff.

Saat, M.R. & Barkan, C.P.L. 2013. Investigating Technical Challenges and Research Needs Related to Shared Corridors for High-Speed Passenger and Railroad Freight Operations. Publication DOT/FRA/ORD-13/29. Federal Railroad Administration, U.S. Department of Transportation, Washington, D.C.

Ullman, K.B. & Bing, A.J. 1994. High Speed Passenger Trains in Freight Railroad Corridors: Operations and Safety Considerations. Publication DOT/FRA/ORD-95/05. Federal Railroad Administration, U.S. Department of Transportation, Washington, D.C.

Van Dyk, B.J., Dersch, M.S., Edwards, J.R., Ruppert Jr., C.R. & Barkan, C.P.L. 2013. Quantifying Shared Corridor Wheel Loading Variation Using Wheel Impact Load Detectors. Proceedings of the 2013 ASME Joint Rail Conference, Knoxville, Tennessee.

Principal factors contributing to heavy haul freight train safety improvements in North America: a quantitative analysis

B. Wang & C.P.L. Barkan
University of Illinois At Urbana-Champaign, Urbana, Illinois, USA
M.R. Saat
Association of American Railroads, Washington, D.C., USA

ABSTRACT: Heavy haul freight railways play an important role supporting North American economic and environmental sustainability. Train accidents have declined considerably over the past decade; however, when they do occur they can damage infrastructure and rolling stock, interfere with transportation, and may cause causalities and releases of dangerous goods. Consequently, continuous improvement in train safety is an ongoing priority of the rail industry. Extensive, detailed data from the US DOT Federal Railroad Administration (FRA) were used to quantify the most important factors contributing to the declining accident trend. The objective is quantitative understanding of where the greatest safety improvements have occurred, and what are the most important opportunities for further improvement in train safety. The rates of derailments, collisions, highway-rail grade crossing accidents and others combined, all declined, but the largest reduction was in mainline derailments. The derailment causes that showed the greatest reduction were broken rails and welds, track geometry and railcar bearing failure, followed by 16 other causes. However, the rate of buckled track derailments, and to a lesser degree roadbed failures and obstructions increased. Previous work investigating the correlation between derailment rate and FRA track class, method of operation and annual gross tonnage were analyzed and updated. The results indicate a general decline in accident rate across nearly all mainline conditions. The results can be used to guide priorities for research and investment of resources to further improve train safety in the most effective and efficient manner.

1 INTRODUCTION

North American railroads play a major role in freight transport, moving 42 percent of the intercity freight in the United States or approximately 40 tons of freight per person per year (FRA 2010). Due to this large volume of traffic, safe operation has broad implications. Freight train derailment accidents have declined over the past decade, from 2,197 in 2006 to 1,344 in 2015, a 39 % reduction (FRA 2016). Despite this improvement, derailments have the potential to damage infrastructure and rolling stock, disrupt services, cause casualties, and may result in release of hazardous materials.

Improvements in freight train safety and derailment reduction have been an ongoing priority for the rail industry and government. Some safety measures reduce risk more effectively than others, so understanding the relative importance of different causes is important if safety is to be improved in the most efficient manner. This paper uses freight train accident data from the U.S. Department of Transportation Federal Railroad Administration (FRA) to examine the contributing derailment causes over the past decade and quantify how they have changed. The objective is to provide insights to assist decision makers in choosing the most effective approaches to further reduce or eliminate accidents.

The principal objective of this paper is to identify the derailment causes having the greatest effect on train safety and risk, and to quantify and rank changes in the number and distribution of derailment causes. The study focus on mainline derailments on the major U.S. railroads (Class 1) and is separated into two time periods: 2006 to 2010 and 2011 to 2015. This approach provides a robust data set in terms of total gross ton-miles (tonne-kilometres). The two time periods were roughly equivalent in terms of total traffic, 16.7 trillion gross ton-miles (24.4 tonne-kilometres) during the first time period and 17.2 trillion ton-miles (25.1 tonne-kilometres) during the second.

2 DATA AND METHODOLOGY

The Rail Equipment Accident/Incident Report (REAIR) form is used by the U.S. railroads to report accidents that exceed a monetary threshold for damages to equipment or infrastructure and it is periodically adjusted for inflation (FRA 2011a). The REAIR database contains details for each accident

including date, railroad, weather, as well as types of track – mainline, siding, yard, or industry. The database also identifies thirteen types of accidents including derailment, collision, and various others. The FRA classification of accident types is based on the initial cause, but may include a secondary cause as well. For example, if a collision caused a derailment, the initial cause would be collision and the secondary cause would be derailment. In this study, such an accident would be classified as a collision. The principal information used in this paper includes date, track type, number of cars derailed, and accident cause.

The FRA lists approximately 400 different cause codes in five categories using single-letter codes as follows: (T) track, roadbed, and structure, (S) signal and communication, (H) train operation, (E) mechanical and electrical failures, and (M) miscellaneous (FRA 2011b). Each of the five categories contains more specific sub-categories with an additional level of detail. This is useful for many studies; however, identification of certain trends benefits from some modified aggregation of related causes. In the early 1990s, Arthur D. Little Inc. (ADL) worked with the Association of American Railroads (AAR) and developed an alternative grouping to the FRA accident cause categorization (ADL 1996; Schafer and Barkan 2008) that consolidated various causes into groups while enabling distinction between certain other cause groups (Anderson and Barkan 2004). For example, the FRA sub-categorization combines broken rails or welds, joint bars, and rail anchors; whereas the ADL method separates broken rails or welds, joint bar defects, and rail anchors into three separate groups. Another example is that FRA combines buckled track as a sub-group within track geometry while the ADL method separates those into two groups. These distinctions are important because the underlying causes and consequent solutions differ.

Estimation of accident rates requires traffic exposure data. FRA provides partial exposure data and the AAR provides additional railroad traffic data. The annual gross ton-miles (tonne-kilometres) for Class 1 railroad freight trains were used as a metric for traffic exposure for calculation of accident rates (AAR 2006 - 2015). Three candidates for traffic exposure were considered: car-miles (car-kilometres), train-miles (train-kilometres), and ton-miles (tonne-kilometres). Ton-miles were chosen because of its availability compared to the other two (Nayak and Palmer 1980).

3 TRAIN DERAILMENT CAUSE ANALYSIS

Developing the most effective derailment prevention strategies requires understanding the root causes of derailments. The most frequent causes included broken rails or welds and track geometry (excluding wide gauge) (Table 1). The five most frequent causes consist of almost 40% of the all the causes in terms of derailment frequency. Eight of the top ten causes were track, roadbed, and structure or mechanical and electrical category, labeled in red and blue, respectively in the table. Higher derailment rate does not necessarily correlate with higher derailment severity (average number of cars derailed). In the next section, the relationship between the frequency (number of derailments) and severity of accident causes simultaneously.

Table 1. Top Ten Frequent Derailment Causes and Respective Derailment Rates, 2006 – 2015

ADL Cause Group	Cause Type	Derailments	Average Number of Cars Derailed	Derailments per Trillion Tonne-Kilometre
Broken Rails or Welds	Track, roadbed and structure	397	12.9	17.0
Track Geometry (excl. Wide Gauge)	Track, roadbed and structure	188	6.6	8.1
Broken Wheels (Car)	Mechanical and electrical	166	7.6	7.1
Obstructions	Miscellaneous	166	13.2	7.1
Bearing Failure (Car)	Mechanical and electrical	155	4.6	6.6
Buckled Track	Track, roadbed and structure	131	12.8	5.6
Train Handling (excl. Brakes)	Train operation and human factors	121	7.1	5.2
Other Axle/Journal Defects (Car)	Mechanical and electrical	107	8.3	4.6
Coupler Defects (Car)	Mechanical and electrical	97	5.4	4.2
Other Wheel Defects (Car)	Mechanical and electrical	92	3.7	3.9

3.1 *Derailment Frequency and Severity*

Dick et al (2003) and Barkan et al (2003) introduced a graphical approach to illustrate the relationship between train derailment causes' frequency and severity. Individual derailment cause is plotted in terms of its average frequency and average severity (Figure 1) and the graph is divided into four quadrants. Data points to the right of the vertical line indicate above-average frequency and points above the horizontal line indicate above-average severity. The relative impact of different cause groups can be evaluated in terms of their respective quadrant. Previous research has found that in addition to derailment frequency, number of cars derailed is a good predictor variable for the derailment severity and risk of hazardous material release due to the higher amount of kinetic energy from more cars derailed (Barkan et al 2003).

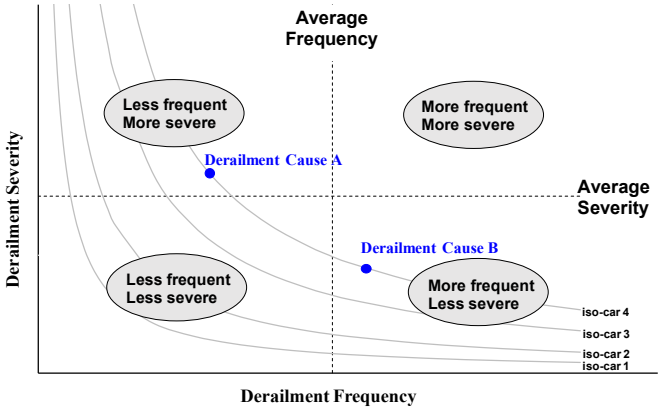


Figure 1 Example of a frequency-severity graph including iso-car lines

Derailment causes in the upper right quadrant occur more frequently and are more severe thereby posing the greatest risk in terms of number of cars derailed. Conversely, causes in the lower left quadrant are rare and tend to be lower severity derail-

ments. The causes in the upper left quadrant have more severe consequences, but their lower frequency makes it more difficult to make consistent predictions. The lower right quadrant includes less severe, but higher frequency derailment causes, which are of secondary interest.

A graphical technique is also introduced here referred to as "iso-car" lines; these represent equal levels of risk in terms of numbers of cars derailed. Iso-car lines are an inverse function of the average number of cars derailed and the number of derailments. Iso-cars provide additional resolution for comparison of derailment causes relative to one another. The distance from the origin represents risk level with higher iso-car numbers indicating higher risk levels. For example, in the absence of iso-car lines, derailment cause A cause would be classified as “less frequent and more severe” while derailment cause B is classified as “more frequent and less severe”. Use of iso-car lines indicates that these two causes pose the same risk level in terms of number of cars derailed.

This paper used the iso-car approach to evaluate frequency and severity in which the normalized derailment frequency and average severity per derailment were plotted. Mainline freight train derailment causes over the study period were compared using a frequency-severity plot such as described above (Figure 2). Throughout the study period, broken rails or welds were the most frequent derailment cause, with the highest iso-car level of 85. This finding is consistent with previous studies (Anderson 2005; Liu 2013). Other rail and joint defects were, on average, the most severe derailment cause, although they occurred much less frequently.

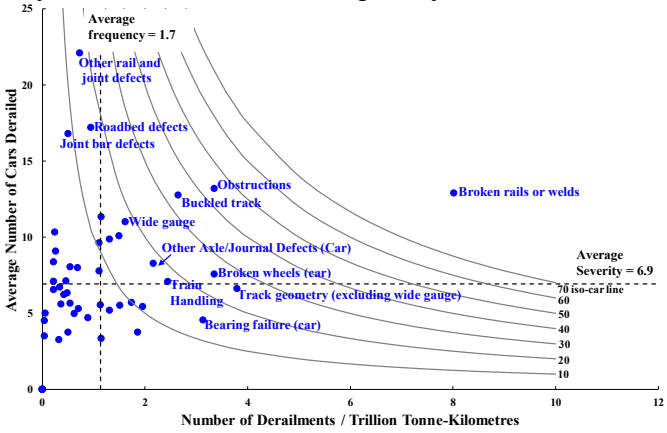


Figure 2 Frequency-severity graph from 2006 to 2015 (Causes with iso-car greater than 20 are labelled)

As discussed above, comparison of cause-specific derailment trends is the objective of this paper so the frequency-severity plot was adapted for this comparison by using different symbols to plot the two time-periods on the same graph (Figure 3). Considering broken rails or welds again, there was a slight increase in severity, but a substantial reduction in frequency when comparing 2006-2010 to 2011-2015. Despite this decrease, broken rails or welds remained the leading cause of large derailments and the highest iso-car level. Track geometry other than wide gauge also declined substantially. Other rail and joint defects were supplanted by joint bar defects as the cause with the highest severity.

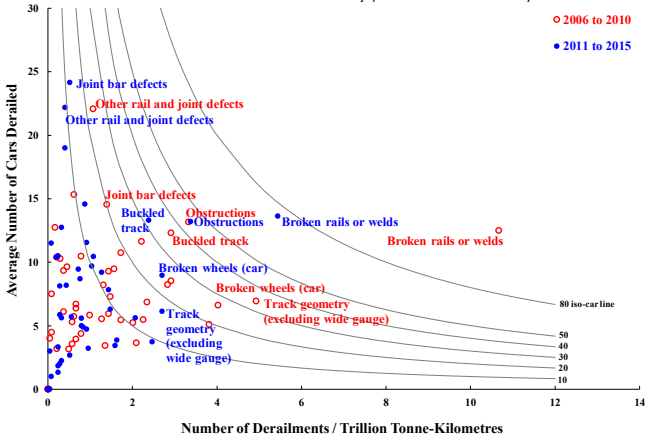


Figure 3 Frequency-severity graph of two time periods (Causes in the first time period with iso-car greater than 40 are labelled)

The graphical technique with iso-cars (Figures 2 and 3) offers a means to simultaneously compare the relative changes among causes' relative frequency and severity; however, quantitative comparison of changes is more difficult. To study the changes quantitatively, the change in the number of derailments and number of cars derailed per trillion tonne-kilometres between the two time periods were compared (Figures 4 and 5).

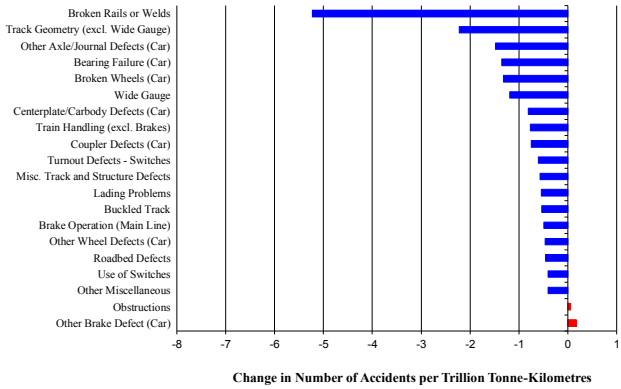


Figure 4 Change in derailment rate by cause group from 2006 – 2010 to 2011 – 2015

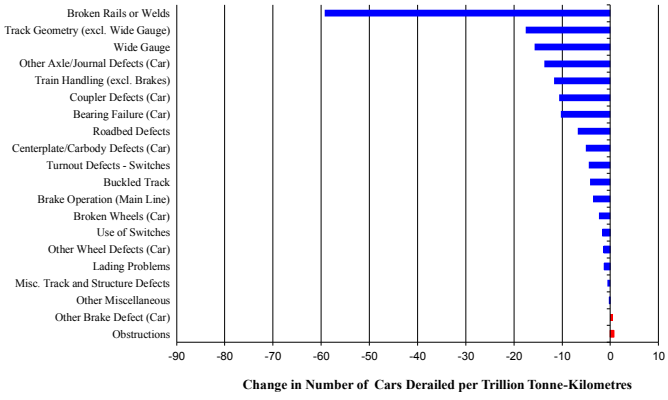


Figure 5 Change in number of cars derailed by cause group from 2006 – 2010 to 2011 – 2015

Comparison of the derailment frequency and severity for the two periods, showed that broken rails or welds were the most frequent and severe in both periods, track geometry (excluding wide gauge) was the second most frequent cause in 2006 to 2010, and obstructions ranked second in terms of severity for both time periods. Although broken rails or welds had the largest number of derailments; they also had the greatest reduction in both derailment frequency and overall number of cars derailed (Figure 4 and 5). Track geometry (excluding wide gauge) showed the second largest reduction in frequency and number of cars derailed. This reduction in derailment rate is consistent with recent studies (Liu 2015). Most derailment causes declined in occurrence rate between the two periods, but two causes, obstructions, and other brake defects (car), showed an increase in terms of derailment rate and number of cars derailed.

A final question considered was whether the decline in most accident causes was proportional to their relative frequency, or whether some declined disproportionately, i.e. more or less rapidly than average. This was addressed by comparing the magnitude of change of each cause group to its frequency in the first time period (Figure 6). The linear regression line represents the average change for all the cause groups. Causes above the regression line had relatively less change between the two periods, whereas causes below the regression line indicate disproportionately greater reduction in derailment frequency. The same approach was used for the number of cars derailed (Figure 7).

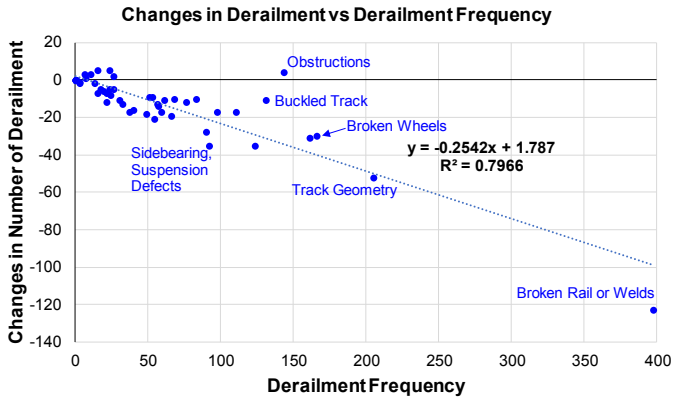


Figure 6 Change in Derailment Frequency vs Derailment Frequency, 2006 – 2010 to 2011 – 2015

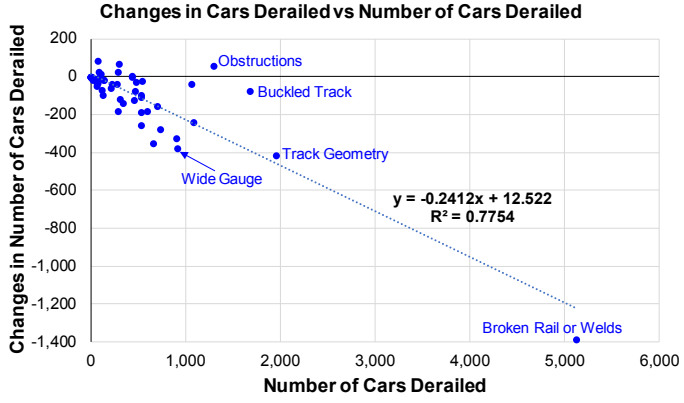


Figure 7 Change in Number of Cars Derailed vs Number of Cars Derailed, 2006 – 2010 to 2011 – 2015

Broken rails or welds, wide gauge, and other axle or journal defects showed greater reduction in terms of both derailment frequency and number of cars derailed, while buckled track, obstructions, and broken wheels had relatively less change in both derailment frequency and number of cars derailed.

4 CONCLUSIONS

U.S. Class I railroad mainline freight train derailment causes were analyzed and a generally decreasing trend was found in derailments caused by broken rails or welds, track geometry, and most other cause groups. The exceptions were obstructions and other brake defects that showed a modest increase in frequency and severity. Broken rails or welds, wide gauge, and other journal or axle defects showed a disproportionately greater reduction compared to their frequency in the earlier time period, while obstructions, buckled track, and broken wheels showed less reduction relative to their frequency. These results provide insights regarding progress in train safety and derailment reduction efforts.

5 ACKNOWLEDGEMENTS

Support for this research was provided by the Association of American Railroads, BNSF Railway, and the National University Rail Center, a US DOT OST Tier 1 University Transportation Center. This paper is solely the work of the authors and does not necessarily reflect the opinions of the sponsors.

6 REFERENCES

Arthur D. Little (ADL). 1996. Risk Assessment for the Transportation of Hazardous Materials by Rail, *Supplementary Report: Railroad Accident Rate and Risk Reduction Option Effectiveness Analysis and Data,* 2nd rev.

Association of American Railroads (AAR) 2006 - 2015. Analysis of Class I Railroads. *Association of American Railroads.*

Anderson, R.T. 2005. Quantitative Analysis of Factors Affecting Railroad Accident Probability and Severity. *M.S. Thesis. University of Illinois at Urbana-Champaign,* 2005.

Barkan, C.P.L., C.T. Dick & R. Anderson. 2003. Railroad derailment factors affecting hazardous materials transportation risk. *Transportation Research Record: Journal of the Transportation Research Board* 1825: 64-74

Federal Railroad Administration 2010. National rail plan progress report. *U.S. Department of Transportation*.

Federal Railroad Administration 2011a. Rail Equipment Accident/Incident Form F 6180.54 Accident Downloads on Demand Data File Structure. *U.S. Department of Transportation*.

Federal Railroad Administration 2011b. FRA Guide for Preparing Accident/Incident Reports. *U.S. Department of Transportation*.

Federal Railroad Administration 2014. U.S. Track Safety Standards; improving rail integrity: Final rule. U.S. *Department of Transportation.*

Federal Railroad Administration 2016. Accident Query. *FRA Office of Safety Analysis, U.S. Department of Transportation.*

Liu, X. Statistical Temporal Analysis of Freight-Train Derailment Rates in the United States : 2000 to 2012. 2015. *In Transportation Research Record: Journal of the Transportation Research Board* 2476: 119-125.

Nayak, P.R., & Palmer, D.W. 1980. Issues and Dimensions of Freight Car Size: A Compendium. *Report FRA/ORD-79/56, Federal Railroad Administration, U.S. Department of Transportation*.

Saccomanno, F.F., & El-Hage, S.M 1991. Establishing Derailment Profile by Position for Corridor Shipments of Dangerous Goods. *Canadian Journal of Civil Engineering* 18: 67-75.