IDEA

Innovations Deserving
Exploratory Analysis Programs

Rail Safety IDEA Program

Field Validation of Inspection Gauges for Wheel Climb Safety at Switch Points

Final Report for Rail
Safety IDEA Project 28

Prepared by:
Dr. Allan M. Zarembski
University of Delaware

November 2016

TRB TRANSPORTATION RESEARCH BOARD
The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Innovations Deserving Exploratory Analysis (IDEA) Programs Managed by the Transportation Research Board

This IDEA project was funded by the Rail Safety IDEA Program.

The TRB currently manages the following three IDEA programs:

- The NCHRP IDEA Program, which focuses on advances in the design, construction, and maintenance of highway systems, is funded by American Association of State Highway and Transportation Officials (AASHTO) as part of the National Cooperative Highway Research Program (NCHRP).
- The Rail Safety IDEA Program currently focuses on innovative approaches for improving railroad safety or performance. The program is currently funded by the Federal Railroad Administration (FRA). The program was previously jointly funded by the Federal Motor Carrier Safety Administration (FMCSA) and the FRA.
- The Transit IDEA Program, which supports development and testing of innovative concepts and methods for advancing transit practice, is funded by the Federal Transit Administration (FTA) as part of the Transit Cooperative Research Program (TCRP).

Management of the three IDEA programs is coordinated to promote the development and testing of innovative concepts, methods, and technologies.

For information on the IDEA programs, check the IDEA website (www.trb.org/idea). For questions, contact the IDEA programs office by telephone at (202) 334-3310.

IDEA Programs
Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001

Field Validation of Inspection Gauges for Wheel Climb Safety at Switch Points

Final Report

IDEA Project S-28

Prepared for

Safety IDEA Program

Transportation Research Board

National Research Council

Prepared by

Dr. Allan M Zarembski PE, FASME, Hon. Mbr. AREMA
Research Professor
Director of the Railroad Engineering and Safety Program
Department of Civil and Environmental Engineering
University of Delaware
Newark DE

November 2016

Acknowledgments

This research was supported by the National Academies' Transportation Research Board (TRB) IDEA Program. The project, "Field Validation of Inspection for Wheel Climb Safety at Switch Points," is part of the Rail Safety IDEA Program, which is funded by the Federal Railroad Administration. Jo Allen Gause was the TRB Project Manager.

Acknowledgment must be made to the members of the project's Expert Review Panel for their assistance and support:

Brad Kerchof (Chairman): Director Research & Tests, Norfolk Southern Railway
Anthony Bohara: Assistant Chief Engineer (retired), Southeastern Pennsylvania Transpiration Authority (SEPTA)
Melody A. Sheahan: Executive Director Engineering Policy, Association of American Railroads (AAR)
Joseph Smak: Senior Director of Track maintenance and Compliance, Amtrak
Yifeng Mao: Assistant Engineer Standards, Long Island Rail Road
Steven Abramopaulos: Engineer Standards, Long island rail Road

Special acknowledgment is extended to Ali Alsahli, Graduate Student, University of Delaware, for his able support in this project.

The authors would also like to acknowledge Norfolk Southern Railway for its active support in the design and fabrication of the prototype gauges and for providing access to their property for evaluation of the gauges.

Table of Contents

Executive Summary

Inspection of turnouts has not kept up with advances in track inspection technologies that have occurred over the last several decades. This is in spite of the fact that turnouts represent the third largest area of track-caused derailments, with almost half in the switch point area.

IDEA Project S-28 is aimed at developing practical and easily deployable measurement gauges for track inspectors to determine whether switch condition has degraded to the point that it could cause a derailment. This activity is a follow-up to the IDEA S-23 project, under which four switch point inspection gauges were developed based on concepts and gauges used in Europe and then modified and adapted for North American application. Each of the four gauges addressed a particular risk condition that might lead to wheel climb at the switch point. These were:

1. Chipped point gauge,
2. AAR 1B wheel contact gauge,
3. Severely worn wheel profile gauge, and
4. Gage-face wear angle gauge.

This project was an implementation and field assessment of the IDEA S-23 activity with the goal of providing railroads and transit systems with a set of practical and directly useable gauges for track inspectors to evaluate the condition of switch points. Field assessment of these gauges was performed in two steps: an initial assessment activity, after which modifications were made to the gauges to reflect the field experience, and a final assessment activity.

In the initial assessment, 272 switches were inspected using the four gauges, with these measurements taken by five participating railroads/consultants. The results were analyzed by comparing the inspector evaluation of the switch point condition with the results of the measurement gauges (pass/fail). The results showed agreement ranging from under 50% to almost 85%, depending on the specific railroad. This suggested several issues including a range of inspection standards among railroads and a range of experience among the inspectors. Analysis of the data suggested that some of the railroad inspectors may not be familiar with certain derailment categories for which specific gauges are intended.

Based on these results, modifications were made to several gauges and a new set of field assessments performed by six railroads or consultants using the modified gauges. This second phase of field assessments reviewed 74 switches and the results showed an average agreement of 82%, with individual railroad agreements varying from under 50% to 95%.

Feedback from the railroads and inspectors was very favorable, noting that the gauges addressed several conditions that are often overlooked, such as high gage face wear (Gauge 4), and in general made for a more detailed inspection of the switch point. As one inspector noted in his report “The Switch Point Inspection Gauges were a great tool for me in learning what to look for when inspecting a point. It increased my awareness on how the wheel actually interacted with the point. … This gauge caused to me to ask the proper questions during inspections, and helped my judgment with determining if a switch point is in correspondence.”

Based on the results of this program, the committee has recommended that these gauges be made available to the railroad industry, as well as to railroad standards organizations such as American Railway Engineering and Maintenance-of-Way Association (AREMA) for industry implementation.

Introduction

Inspection of railroad right-of-way has been an area of significant technological development and implementation, including significant advances in automated inspection technologies. However, the area of turnout inspection has not kept up with these advances in track inspection technologies, despite turnouts representing the third largest area of track caused derailments, with more than $4 million in FRA reported derailment costs in 2014 alone [*1*]. Of these derailments, almost half are in the switch point area.

IDEA Project S-28 is a field assessment and validation of the four switch point inspection gauges previously developed under IDEA Project S-23 for prevention of wheel climb derailments. The specific activity took the gauges developed under S-23 into a field environment and selected track inspectors on several railroads who used these gauges for a trial period to assess their practical usability as well as their value in identifying marginal or potentially dangerous switch points. The gauge designs were then re-evaluated in light of these field trials and modified accordingly. Inspectors then performed a second set of field evaluations on several railroads using the modified gauges. The results and conclusions from both evaluations are presented here.

The four gauges are listed here:

1. Chipped point gauge,
2. AAR 1B wheel contact gauge,
3. Severely worn wheel profile gauge, and
4. Gage-face wear angle gauge.

Each of these gauges addresses a particular risk condition that might lead to wheel climb at the switch point. This was presented in the earlier IDEA S-23 report [*2*]. This project, S-28, represents an implementation and field assessment of the IDEA S-23 activity, with the goal of providing railroads and transit systems with a set of practical and directly useable gauges for track inspectors to evaluate the condition of switch points.

This report presents the results of the IDEA S-28 activity, which includes the field assessment of these gauges, both before and after final modifications, and detailed analysis of the results. The field assessment participants, which included five railroads, TTCI, and one independent consultant, provided detailed inspection data and comments that are presented in this report. The results of this activity are a final set of switch point measurement gauges to be used by U.S. railroads in evaluating switch point conditions and safety.

IDEA S-28: Stage 1

Task 1. Initial Field Deployment of Prototype Gauges

The results of S-23 were a set of four handheld swtich point gauges modified to reflect U.S. conditions and practices, each of which focuses on a specific set of potentially dangerous switch point wear conditions. These initial gauges were as follows:

- Chipped point (CP) gauge: Addresses chipped or damaged switch points.
- Gage face angle gauge (GFA gauge): Checks the gage face angle and the potential for wheel climb, particularly for high L/V conditions.
- AAR 1B gauge: Addresses unsafe wheel and rail contact using an AAR 1B new wheel profile.
- Severe profile gauge (SP gauge): Severely worn profile gauge that addresses potential for wheel climb derailment for a severely worn wheel.

The first task of the S-28 project involved taking the recommended S-23 gauges and preparing them for actual production and field deployment. This included developing all of the manufacturing details required to physically produce the gauges and then employing them in the field. This included how the gauges were mounted and deployed so as to allow for their physical mounting in the field by inspectors so as to allow for easy and reliable measurements in a range of field conditions.

Based on the recommendations of the advisory team following early field deployment, the gauges were mounted on a bar or rod that spans both rails, similar to a gauge rod, in order to reference the adjacent rail and provide an accurate measurement. To minimize the number of gauge sets to be carried, two gauges are mounted on one measurement bar, thus requiring two bars to accommodate the four gauges as shown below.

	Gauges
Bar 1	• AAR 1B wheel contact gauge • Chipped point gauge
Bar 2	• Severely worn wheel profile gauge (severe profile gauge) • Gage face wear angle gauge

A set of user guides was developed for each of the four gauges (Appendix A presents the final user guide based on the final recommended versions of the four gauges).

The four gauges were initially implemented as follows:

Chipped point gauge (*Gauge 1*)
This gauge addresses potential for wheel climb at chipped or damaged switch points. The key dimensions of the gauge include a vertical dimension (height) of 0.70” and a gauge-face angle of 70° (Figure 1). The gauge indicates the switch fails if the bottom edge of the gauge is vertically above the top of a broken or worn point as shown in Figure 2.

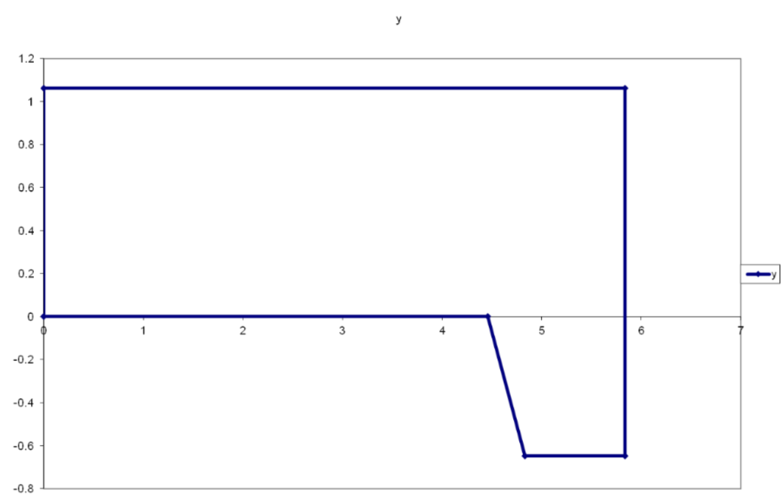
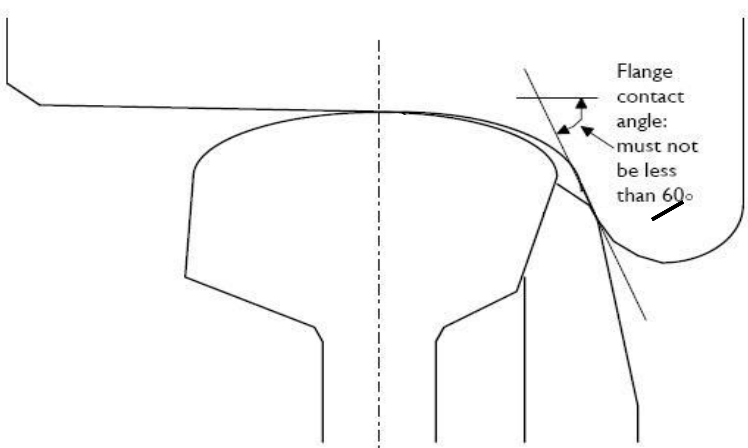


Figure 1: Diagram of chipped point gauge (Gauge 1) dimensions.

Figure 2: Damaged switch point (courtesy of NS).

AAR 1B wheel contact gauge (*Gauge 2*)

This gauge is the US version of the Network Rail TGP8 gauge, using the AAR 1B wheel profile with a 60^{o} contact mark. The gauge monitors inappropriate wheel/rail contact on switch points; specifically, to ensure there is no contact between the bottom of the wheel flange below the 60^{o} mark and the switch rail, where there may be a potential for the wheel climb. Figures 3 and 4 a show a diagram and photo, respectively, of the gauge illustrating the 60^{o} flange angle and contact mark. Note the contact point between the gauge and the switch rail in the two figures. Contact below the 60^{o} mark is considered to be potentially unsafe.



e.

Figure 3: AAR 1B checks for contact under 60^{o} mark.

Figure 4: AAR 1B wheel contact gauge where the contact point is above 60^{o} mark (courtesy of NS).

Severely worn wheel profile gauge (*Gauge 3*)

This gauge gives an indication of the potential for wheel climb derailment for severely worn (but non-condemnable) wheel profiles. It can be used to identify gapping or broken points or top-worn stock rails that can, when combined with a severely worn wheel flange, result in a wheel climb derailment. The gauge has a slider mechanism (see instructions in Appendix A), which is used to determine if the gauge will catch on the top of the switch point, indicating a potential wheel climb condition ("fail" condition as shown in Figure 5) or whether the slider on the gauge slides down the gage face of the switch point, indicating a "pass" condition (Figure 6). The gauge is designed with a 70° gage face angle that appears to correspond to unsafe severely worn wheel conditions.

Figure 5: Severely worn wheel profile gauge (severe profile gauge) (courtesy of LIRR).

Figure 6: Severely worn wheel profile gauge (severe profile gauge) (courtesy of NS).

Gage face wear angle gauge (*Gauge 4*)

This gauge checks the gage face wear angle of the switch point and the potential for wheel climb, particularly for high L/V conditions. The gage is designed with a 32º gage face angle (58º if the complementary angle is used); so that if the switch point's gage-face angle is greater than 32º and the switch condition is safe otherwise, the switch is unsafe. Figure 7 shows a safe condition using this gauge.

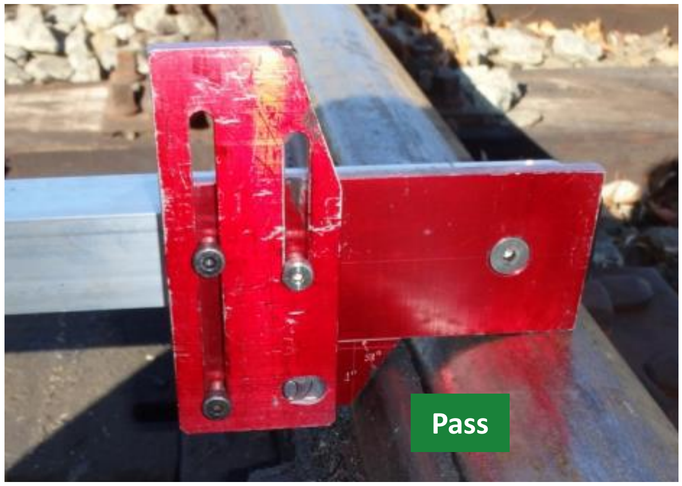


Figure 7: Gage face wear angle gauge (courtesy of LIRR).

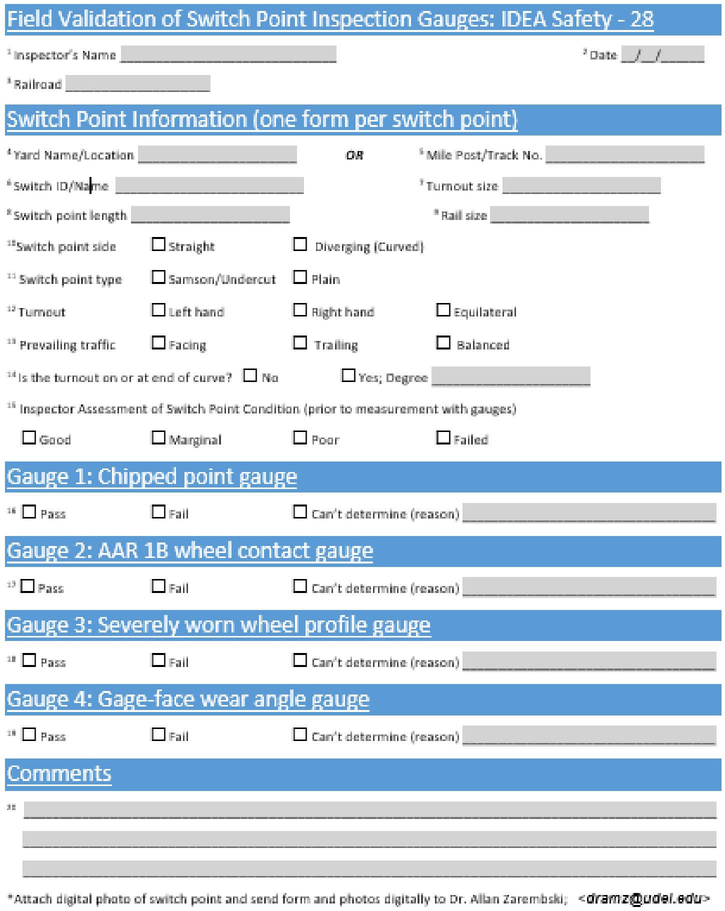
Implementation of Gauges in Task 1

As part of the Task 1 implementation, six sets of gauges were fabricated by Norfolk Southern and distributed to the following participating railroads: Norfolk Southern, Long Island Rail Road (LIRR), BNSF, Canadian National, and Gary Wolf (consultant). Preliminary assessment of the analysis requirements resulted in a recommended minimum of 200 switch point inspections for the process of field validation. Based on this, each railroad was asked to inspect 40 to 50 switches as part of Task 1. Each of the participating railroads was provided with a full set of gauges (four gauges mounted on two bars), a set of instructions, and an inspection form (see next section).

Inspection form

An inspection form was developed that included the following information (Figure 8):

- General railroad information:
  - Switch location
  - Switch point properties.
- Inspector's assessment:
  - The inspector was asked to choose one of the following four conditions: Good, Marginal, Poor, and Failed.
- Gauge inspection results:
  - Results from each of the four gauges separately; Pass, Fail, Can't determine.
- Railroad inspectors to also send photos of the switches inspected.


Field Validation of Switch Point Inspection Gauges: IDEA Safety - 28

1 Inspector's Name ____________ 2 Date __/__/____

3 Railroad ____________

Switch Point Information (one form per switch point)

4 Yard Name/Location ____________ OR 5 Mile Post/Track No. ____________

6 Switch ID/Name ____________ 7 Turnout size ____________

8 Switch point length ____________ 9 Rail size ____________

10 Switch point side ☐ Straight ☐ Diverging (Curved)

11 Switch point type ☐ Samson/Undercut ☐ Plain

12 Turnout ☐ Left hand ☐ Right hand ☐ Equilateral

13 Prevailing traffic ☐ Facing ☐ Trailing ☐ Balanced

14 Is the turnout on or at end of curve? ☐ No ☐ Yes; Degree ____________

15 Inspector Assessment of Switch Point Condition (prior to measurement with gauges)

☐ Good ☐ Marginal ☐ Poor ☐ Failed

Gauge 1: Chipped point gauge

16 ☐ Pass ☐ Fail ☐ Can't determine (reason) ____________

Gauge 2: AAR 1B wheel contact gauge

17 ☐ Pass ☐ Fail ☐ Can't determine (reason) ____________

Gauge 3: Severely worn wheel profile gauge

18 ☐ Pass ☐ Fail ☐ Can't determine (reason) ____________

Gauge 4: Gage-face wear angle gauge

19 ☐ Pass ☐ Fail ☐ Can't determine (reason) ____________

Comments

20 ____________

*Attach digital photo of switch point and send form and photos digitally to Dr. Allan Zarembski; <dramz@udel.edu>

Figure 8: Blank inspection form.

Task 2. Field Evaluation and Analysis of Results

Task two examined the results of the initial field validation activity with a focus on the effectiveness of the gauges and whether any of the gauges need to be modified or improved.

Field Evaluation (*Oct 2015–Feb 2016)*

A comprehensive initial field evaluation was conducted during the period from October 2015 to February 2016. Measurements were taken by four railways plus an independent consultant who inspected a total of 285 switch points of different design configurations and conditions using the four study gauges. The tests were taken in more than 20 train yards.

Table 1 shows the number of switches inspected by railroad.

	Switches inspected
NS	22
LIRR	45
CN	135
Wolf[1]	41
BNSF	42
Total	285

Table 1: Summary of Inspections by Railroads

[1] Gary Wolf is an independent railroad consultant who agreed to test the gauges on several local railroads.

Of the 285 measurements shown in Table 1, 13[2] were discarded due to the unclear or missing key answers and/or measurements, resulting in a working total of 272 measurements.

Data Analysis

The 272 sets of measurements taken from five participating railroads and consultant were initially analyzed by comparing the evaluation of the switch point condition with the results of the measurement gauges. Thus, this preliminary analysis must be used with caution since it is not clear if the different inspectors were familiar with all of the failure mechanisms addressed by the four gauges or were sufficiently experienced to be able to properly evaluate the switches. It should be noted that for several of the railroads multiple inspectors were used.

A second analysis used a more formal Decision Tree analysis approach to test the relationship between the gauges and the assessment of the inspectors. The study team chose the Decision Tree analysis based on WEKA, a collection of machine learning algorithms for data mining tasks developed by the University of Waikato [*6*].

Of the 272 switch points inspections used in the analysis, 137 had four categories of switch point conditions[3], and one railroad, CN, had 135 switch point inspections, modified into a three-category assessment of "Good," "Fair," and "Poor."

As a result, four different sets of analyses were performed representing:

1. Full data set (272 inspections).
2. CN data with three inspector categories (135 inspections).
3. All other data with four inspection categories (137 inspections).
4. Analysis by individual railroad.

For easier identification, the inspection results were also color coded as follows:

Red—indicating failed or near failed condition. For a gauge measurement, this represents a failure reading of the gauge.

Yellow—indicating degraded or poor condition. If the gauge is on the border, it is set to yellow.

"No color"—indicating good or marginal condition, but not near failing. If the gauge does not indicate failure, the gauge reading is set to "no color."

These analysis results are summarized in Table 2.

Case/Railroad	Number of inspections	% Agreement	% Disagreement
1. All inspections	272	58	42
2. CN only	135	49	51
3. All others	137	66	34
4A: BNSF	41	44	56
4B: LIRR	45	84	16
4C: NS	16	50	50
4D: Gary Wolf	35	77%	23%

Table 2: Summary of Comparison Between Inspectors and Gauges

As can be seen in this table, there is a wide range of agreement with the gauges, ranging from under 50% to almost 85%., which suggests several issues:

[2]One measurement is included for some of the analysis since it is a switch point assessed to be failed by the inspector and which failed three of the four gauges; the fourth gauge being not conclusive. As such it represents an inspector-assessed failed switch point that also failed several gauges.

[3]"Good," "Marginal," "Poor," and "Failed."

- There is a range of inspection standards between railroads. For example, LIRR is a passenger carrying railroad and as such is less tolerant of a marginal switch point conditions that can lead to a derailment.
- There is a range of experience among the inspectors. For example, Gary Wolf is an experienced inspector with an extensive range of derailment investigation experience. As such, he may be more sensitive to identifying switches that are at or near derailment potential.
- Different inspectors may be familiar with different types of derailments and switch conditions and not familiar with others. Preliminary analysis of the data suggests that some of the railroad inspectors may not be familiar with certain derailment categories for which specific gauges are intended, such as wheel climb (Gauge 2 or Gauge 3) and worn ear (Gauge 4). This is an area that requires more detailed investigation at the individual gauge level.

Table 3 summarizes the distribution of gauge failures.

Case/Railroad	Number of inspections	Gauge 1	Gauge 2	Gauge 3	Gauge 4
4. All inspections	272	44%	21%	28%	7%
5. CN only	135	60%	19%	11%	10%
6. All others	137	28%	24%	45%	4%
4A: BNSF	41	28%	32%	38%	2%
4B: LIRR	45	30%	10%	60%	0%
4C: NS	16	29%	21%	36%	14%
4D: Gary Wolf	35	29%	7%	64%	0%

Table 3: Summary of Comparison of Gauge “Failures”

As can be seen from Table 3, Gauges 1 and 3 generated the greatest number of failures. Of particular interest is that CN had 60% of its failures with Gauge 1 (chipped point gauge), whereas all the others appeared to have the greatest number of failures with Gauge 3 (severely worn wheel gauge). This would suggest a different method of switch point inspection and replacement based on railroad policy or practice, with CN perhaps having a lesser emphasis on broken points. The inspectors’ unfamiliarity with the severely worn wheel condition would lead to a high percentage of failures for this gauge.

Preliminary Railroad Comments

As part of the field validation activity, the study team asked for the participant railroads’ feedback regarding the gauges usability in different situations in the working environment. The feedback came in two forms:

1. Effectiveness of the different gauges, and
2. Suggestions for improving the usability of the gauges.

In the first area, Effectiveness, the responses were very positive:

- The gauges provided useful information on switch point condition by reinforcing the track inspector’s judgment on when maintenance or replacement is necessary.
- Internally, we like the gauges as developed. The local track supervisors like them. We aren’t using them yet to condemn switches, but we probably should be giving them a hard look.
- All of the gauges were equally easy to use; however, they were cumbersome and thus unlikely to be taken out of the truck during a regular inspection. If a measuring device could be stored in a pouch and the device modified to be lined up with the base of the rail it would be easier to use.
- The biggest concern about switch point inspections is that they are a judgment call. Under the current inspection requirements, it is up to a track inspector to know if a switch point poses a

hazard or not. If these gauges identify 100% of the dangerous points they would be one of the best things to be added to a track inspector's tools in a long time. However, two concerns:
    - Not convinced that these gauges would identify every dangerous point.
    - If widely implemented, inspectors will rely more heavily on the gauge than on their own knowledge of track/train dynamics and track structure. The gauges will be used as a crutch to help poor inspectors get by and either cause more money to be spent in areas where the point is actually acceptable or leaving questionable points in areas where they shouldn't be.
- One additional benefit found from Gauge No. 3, severely worn wheel profile gauge, is that it can also be used to identify the top head wear on the switch point. One also could say that an experienced track inspector could tell from looking at that too, but the gauge can reinforce the Inspector's judgement, as it is not subjective.
- One railroad indicated a preference of the chipped point gauge (Gauge 1) and AAR 1B gauge (Gauge 2) over the other two gauges.
- One railroad indicated that the AAR 1B (Gauge 2) and gauge face wear angle gauge (Gauge 4) were the most useful.
- One railroad expressed concern that the chipped point gauge (Gauge 1) used too small a chip point depth; their standards used 7/8" (0.875), instead of 0.7".
- Several inspectors reported difficulty using a severely worn wheel profile gauge in special weather conditions and to inspect a special type of switch points. For example, snow conditions and when there is a switch point protector.

In the second area, there were several very practical suggestions for improvement:

- Apply instruction on the side of the bar.
- Extend instructions to include locations on switch to be measured.
- Modify the handle on the gauges to be larger for hands with gloves.
- Color code the four gauges in four different colors to make it more user friendly.
- Put template on both sides of Gauge 2 and add color differences above and below the 60-degree mark (green/red to indicate pass/fail).
- Have the gauges interchangeable on a single rod.

Decision Tree Analysis

A second set of analyses was performed on the data making use of a statistical Decision Tree analysis approach. Decision tree analysis is a "data mining" technique that has gained traction in the statistical data analysis field. It uses a "divide-and-conquer" approach to the problem of learning from a set of independent data events, which are in this case switch points inspections. The specific analysis approach used here made use of WEKA Data Mining Software [*6*], and specifically the J48 Algorithm. The J48 algorithm chooses one attribute[4] of the data that most effectively splits its set of samples into subsets enriched in one class[5] or the other. As a result, the J48 will choose a "sufficient" gauge in the top of the tree followed by gauges that are less "sufficient" further down the Decision Tree. The algorithm excludes gauges that are not based on the data and assumptions of each case.

This section includes analyses of the gauge measurements and inspector evaluations using WEKA's J48 Algorithm [*6*] on the data sets presented previously. The purpose of this analysis is to study the available data set and bring out any relationships that can be identified using this data mining technique, and to further determine the effectiveness of the gauges based on this initial field validation. The outcome of the Decision Tree is a determination of which of the tested gauges are "sufficient."

[4]Each instance that provides the input to machine learning WEKA is characterized by its values on a fixed, predefined set of features or *attributes* (i.e., each switch that has been inspected has five attributes: gauge 1 result, gauge 2 result, gauge 3 result, gauge 4 result, and assessment).
[5]This outcome of an attribute is called the *class* (e.g., assessment of a switch will have four classes "Good/Marginal/Poor/Failed").

Three set of analyses were performed as follows:
1. Full data set (272 inspections)[6]
2. CN data with three inspector categories (135 inspections)
3. All other data with four inspection categories (137 inspections)[7].

Case 1: Analysis of full set of inspections

The results of WEKA Decision Tree simulation are illustrated in Figure 9, which presents a Decision Tree diagram for the full data set. The most relevant or "sufficient" gauge is shown on top with the remaining in descending order. The "P" refers to the "Pass" branch for that gauge; that is, the gauge passes the switch point, the "F" refers to the "Fail" branch; that is, the gauge fails the switch point. Case 1 was divided into two subcases, 1A and 1B, both with the full data set, but with different logic to try to combine the three-tiered CN evaluations with the four-tiered assessments made by the other railroads. The results are presented in Figures 9 and 10.

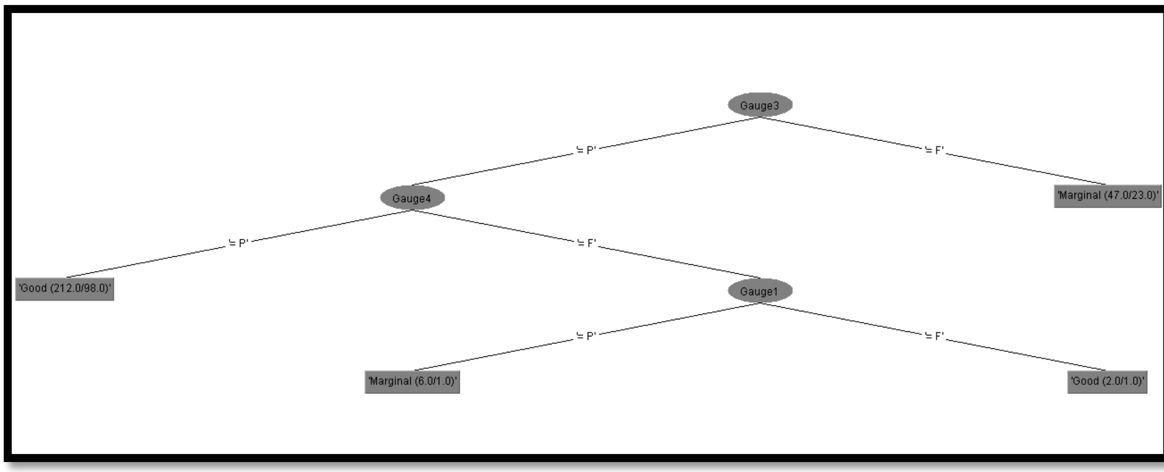


Figure 9: Decision Tree diagram for the full data set Case 1A.

The following observations can be made using the Decision Tree shown above:

- Severely worn gauge Gauge 3 is the most relevant based on analysis of case 1A followed by the Gauge-Face Angle gauge, Gauge 4, and the chipped point gauge, Gauge 1.
- AAR 1B gauge Gauge 2 was not considered in the tree as it has low relationship between the gauge results and the assessment based on the analysis. This suggests lack of knowledge of this type of derailment by the inspectors.
- The outcomes predicted by the software were "Good" or "Marginal" only, due to a low number of instances that have a "Poor" or "Failed" assessment associated with them.

The accuracy statistics for this simulation are shown in Table 4.

Correctly Classified Instances	Incorrectly Classified Instances	Kappa Statistic	Weighted avg. F-measure
144 (53.93%)	123 (46.07%)	0.1396	0.468

Table 4: WEKA Simulation Statistics for Case 1A

[6]In the WEKA analysis, the full data set includes 267 switch points.

[7]The WEKA analysis did not include the individual railroad inspections, except for CN, due to the relatively low number of inspections for the individual railroads.

As can be seen in Table 20, the accuracy is 53.9%, with 144 instances out of a total of 267 found to be "correctly" classified. This is very close to the 58% correlation found in the direct correlation analysis. The Kappa statistic is used to assess the accuracy of any particular measuring case to distinguish between the reliability of the data collected and their validity. $K = 1$ indicates perfect agreement, $K = 0$ indicates chance agreement.

Case 1B (Figure 10) also includes the full data set, but with a different logic applied to the three-tier CN condition assessments, as discussed previously.

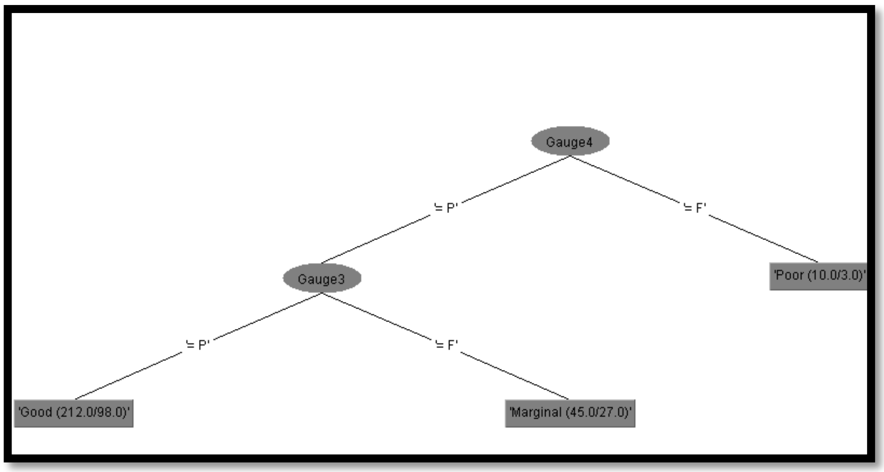


Figure 10: Decision Tree diagram for the full data set Case 1B.

Based on this diagram, the following observations can be made:

- Gauge-face angle gauge, Gauge 4, is the most relevant based on analysis of Case 1B followed by the severely worn gauge, Gauge 3.
- Chipped point gauge, Gauge 1, and AAR 1B gauge, Gauge 2, were not considered in the tree as they had a low relationship between the gauge results and the assessment based on the analysis.

The accuracy for this simulation was 52.1%, with 139 instances of "correct classification" out of a total of 267 (the direct correlation showed 54% agreement.) The Kappa statistic is 0.1709.

Case 2: Analysis of CN inspections

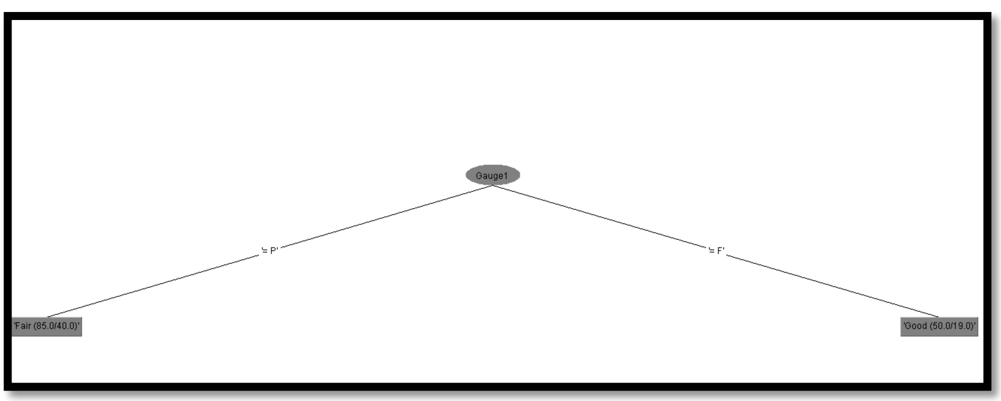


Figure 11: Decision Tree diagram Case 2 CN data only

The WEKA analysis of Case 2 showed that the chipped point gauge, Gauge 1, is the most relevant and there is no relationship between the outcomes “assessments” the other three gauges. The accuracy for this case is 56.3%, with 76 instances out of total of 135 found to be correctly classified. The direct correlation analysis showed only 49% agreement. The Kappa statistic is 0.2055.

Case 3: Analysis of Non-CN inspection

Figure 12 (Decision Tree diagram for Case 3) presents the results for the non-CN data (four inspector assessment categories).

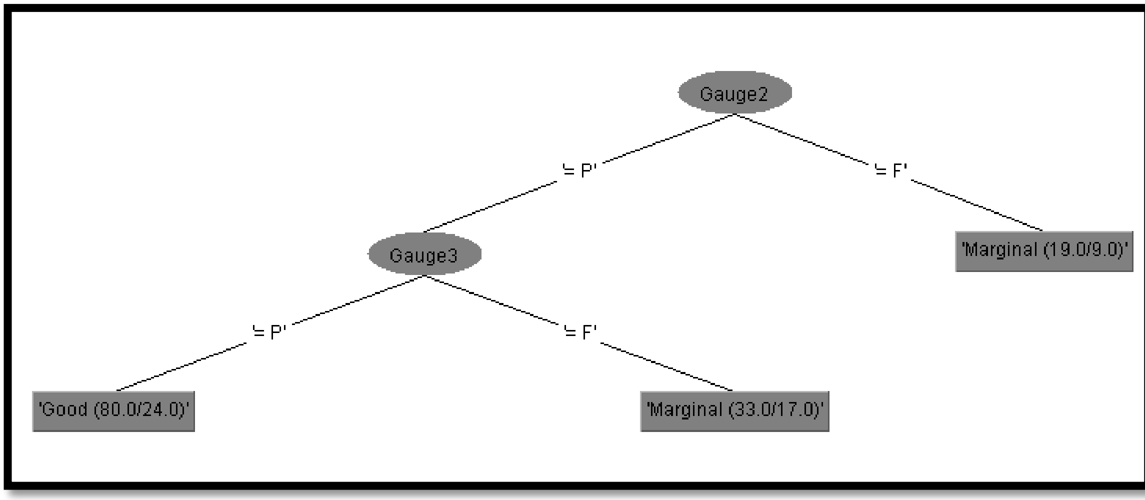


Figure 12: Decision Tree diagram for Case 3: All non-CN railroads.

The results of the Decision Tree analysis were:

- AAR 1B gauge, Gauge 2, is the most relevant based on analysis of Case 3, followed by the severely worn gauge, Gauge 3.
- Chipped point gauge, Gauge 1and gage face wear, Gauge 4, were not considered in the tree as they had a low relationship between the gauge results and the assessment based on the analysis.

The accuracy is 62.1%, with 82 instances out of total 132 found to be correctly classified. This compares with the direct correlation agreement of 67%. The Kappa statistic is 0.567, which is the highest among all analysis preformed in this section.

Summary

Summarizing the results of the Decision Tree analysis shows that for:

- Case 1A: Severely worn gauge, Gauge 3, is the most relevant based on analysis of Case 1A, followed by the gage-face angle gauge, Gauge 4, and the chipped point gauge, Gauge 1.
- Case 1B: Gage-face angle gauge, Gauge 4, is the most relevant based on analysis of Case 1B, followed by the severely worn gauge, Gauge 3.
- Case 2: Chipped point gauge, Gauge 1, is the most relevant and there is no relationship between the outcomes “assessments” the other three gauges.
- Case 3: AAR 1B gauge, Gauge 2, is the most relevant based on analysis of Case 3, followed by the severely worn gauge, Gauge 3.

These results are summarized in Table 5.

Gauges	Case 1A: Full Data Set.	Case 1B: Full Data Set-Alternate CN Logic.	Case 2: CN Data Only.	Case 3: Non-CN Data.
Chipped point gauge	*Not relevant*	*Not relevant*	*Most sufficient*	*Not relevant*
AAR 1B wheel contact gauge	*Least sufficient*	*Not relevant*	*Not relevant*	*Most sufficient*
Severely worn wheel profile gauge	*Most sufficient*	*Less sufficient*	*Not relevant*	*Less sufficient*
Gage-face wear angle gauge	*Less sufficient*	*Most sufficient*	*Not relevant*	*Not relevant*

Table 5: Summary of Decision Tree analyses

As can be seen in Table 5, each of the four gauges is "Most Sufficient" in at least one case. This suggests that the "best" gauge varies with condition, type of problem, railroad standards, and/or inspector experience. It also indicates that at this stage all four gauges are needed in order to properly inspect a switch point.

Review of Gauges and Modifications

The results of the analyses presented here were reviewed by the Expert Review Panel at a meeting on April 20, 2012, at the University of Delaware. (For a more detailed discussion of these results see reference 7.) Specific focus was on review of the prototype gauges, recommendations for changes, and follow-up field investigations. Based on the results of the Task 2 field assessments, analyses, and railroad comments it was determined that all four gauges were of value and should be retained. However, these analyses, together with a review of current railroad standards and practices, resulted in recommendations for modification of the gauges and gauge use instructions which were then implemented in Task 3.

Gage 1: Chipped point gauge

This gauge addressed the potential for wheel climb at chipped or damaged switch points and was found to have significant variation in agreement between gauges and inspector between different railroads. For example, CN showed double the percentage of failures of this gauge compared with other railroads. This may be due to variations in internal railway standards or a different focus of switch point inspections. The key dimension of the gauge was the vertical dimension (height) of 0.70" (as illustrated in Figure 1). The gauge indicates the switch "fails" if the bottom edge of the gauge is vertically above the top of a broken or worn point. This 0.70" height was calculated during the IDEA S-23 activity based on the potential for climb of a worn will on a chipped switch point. However, comparison with different railroad and transit standards showed that the actual range of condemnable chipped point height was between 0.75" (FTA[8]) and 0.875" (Amtrak, BNSF), both of which allowed for deeper chipping or damage. Therefore, it was decided to increase the height of the gauge (allowable depth of the chip or damage) to 0.75", which is still at the conservative end of the range of current practice.

In addition, it was noted that the original instructions stated that the first 12" of the switch point should be inspected by this gauge. However, reports from the field indicated the presence of chipped points further back than 12", so the instructions were modified to say, "Apply this gauge where the point is worn or broken. This is typically within 24 inches of the tip, but could be anywhere along the point."

Gauge 2: AAR 1B wheel contact gauge

This gauge uses the AAR 1B wheel profile with a 60^{o} contact mark to monitor for inappropriate wheel/rail contact on switch points; specifically, to ensure there is no contact made between the bottom of the wheel

[8]Federal Transit Administration, "Recommended Practice for Rail Transit Track Inspection" (2009).

flange below the 60^o mark and the switch rail, where there may be a potential for the wheel climb. Reports from the field indicated this gauge was useful, but it was hard to see the 60^o mark. As a result, this mark was machined onto both sides of the gauge and made bolder.

The instructions for the AAR 1B wheel gauge were revised with the same language as for the chipped point gauge: "Apply this gauge where the point is worn or broken. This is typically within 24 inches of the tip, but could be anywhere along the point."

Gauge 3: Severely worn wheel profile gauge

This gauge gives an indication of the potential for wheel climb derailment for severely worn (but non-condemnable) wheel profiles. This gauge can be used to identify gapping or broken points or a top-worn stock rail that, when combined with a severely worn wheel flange, can result in a wheel climb derailment. No changes were recommended for this gauge. However, the instructions were modified to say, "Apply this gauge within the first inch (approximately) of the tip of the point, or where wheel flange contact is first evident."

Gauge 4: Gage face wear angle gauge

This gauge checks on the gage face wear angle at the switch point and the potential for wheel climb, particularly for high L/V conditions. The gage was designed with a 32º gage face angle (58º if the complementary angle is used), so that if the switch point's gage face angle is greater than 32º, the switch condition is safe. However, after review of Amtrak and LIRR standards, this gauge was changed to a 30º gage face angle (60º if the complementary angle is used). The instructions were modified to state, "Apply this gauge where gage face wear is evident anywhere along the point."

IDEA S-28: Stage 2

The high level of interest by all participating railroads, together with the recommendations for gauge and instruction modification noted previously, indicated the need for a second round of field inspections, using the modified gauges, as called for in Task 3. Norfolk Southern again agreed to fabricate replacement measurement heads for the seven gauges currently in service, and to manufacture an additional two gauges for field inspection use.

The participating railways/consultants for Stage 2 field evaluation (Task 3) were as follows:

- BNSF
- CP
- Long Island Rail Road
- Norfolk Southern
- TTCI
- Gary Wolf

The gauges were fabricated and distributed to the participants in early September 2016.

Task 3. Field Deployment of Modified Gauges and Analysis of Results

Task 2 of the project examined the results of the initial field validation activity (which were collected under Task 1) with a focus on the effectiveness of the gauges and if any of the gauges needed to be modified or improved. Task 3 included a second field validation, using the modified gauges and a revised set of instructions, as well as the analysis of the results from that field work. .

Field Evaluation (*September 2016–October 2016*)

The follow-up field evaluation of the modified gauges was conducted from September 2016 to October 2016. Measurements were taken by four railways, TTCI, and an independent consultant who reviewed 84 switch points of different design configurations and conditions using the four study gauges (see Appendix A). The following table shows the number of switches inspected by railroad in the field validation activity.

	Switches inspected
NS	10
LIRR	19
CP	7
Wolf[9]	15
TTCI	4
BNSF	19
Total	74

Table 6: Summary of Stage 2 Inspections by Railroads

In this round, all participants used the four-level assessment; therefore, there was no need to try to match logic as in the case of the first round of inspection data.

Data Analysis

The 74 sets of measurements were again analyzed by comparing the evaluation of the switch point condition with the results of the measurement gauges. All of switch point inspections used the four categories of switch point condition assessment[10]. As in the earlier analysis, the comparison was between the gauge measurements (Pass/Fail) and the inspector assessments. However, in this set of results, the field inspector results were reviewed[11] by a member of the Review Team to ensure that the inspection gauges were properly applied and the switch condition assessment was consistent.

Based on these data the following sets of analyses were performed:

- Full data set (74 inspections)
- Analysis by individual railroad.

As in the earlier analysis, for the sake of easy identification, the inspection results were color coded as follows:

Red—indicating failed or near failed condition. For a gauge measurement, this represents a failure reading of the gauge.

Yellow—indicating degraded or poor condition. If the gauge is on the border, it is set to yellow.

"No color"—indicating good or marginal condition, but not near failing. If the gauge does not indicate failure, the gauge reading is set to "no color."

These analysis results are summarized here.

Case 1: Analysis of full set of inspections

The total number of inspections in this case is the full "valid" dataset of 74 switch inspections, which included the 135 CN inspections with three condition categories and the 137 other inspections with four

[9]Gary Wolf is an independent railroad consultant who agreed to test the gauges on several local railroads.
[10]"Good," "Marginal," "Poor," and "Failed."
[11]Photos were provided by the field inspector for all the evaluated gauges and these were reviewed by Brad Kerchof of NS, the chair of the Expert Review Panel.

condition categories. To allow for a consistent analysis, the following definitions were used based on the three color-coded categories:

- "Good" and " Marginal"
- "Poor" is considered yellow condition
- "Failed" is considered red condition

Based on these color definitions, the following combination of events was tabulated:

- Inspector failed point (red)/one of more gauges failed the point [red positive].
- Inspector failed point (red)/point passed all gauges [red negative].
- Inspector rated point good or marginal (no color)/one of more gauges failed the point [red negative].
- Inspector rated point poor (yellow)/one of more gauges failed the point [yellow positive].
- Inspector rated point poor (yellow)/point passed all gauges [yellow negative].
- Inspector rated point good or marginal (no color)/point passed all gauges. No color [good condition agreement or good-good].

The results of this comparison are presented in Table 7.

	R+	R-	Y+	Y-	Total R/Y
Count	3	12	11	1	27

	Count	% of total
Total	74	100
Subtotal with one or more indicator	27	36
Positives	14	19
Negatives	13	18
Good condition agreement	47	64
Total agreements	61	82
Disagreements	13	18%

Table 7: Summary of Analysis for Case 1 (all data)

Thus, for Case 1 there was good agreement between the gauges and the inspector (to include yellow conditions) for 61 of the 74 switches or 82% of the switches inspected.

Table 8 presents the distribution of failures identified by gauge.

	Failed Gauges count				
	G1	G2	G3	G4	total
Red / yellow	6	10	11	4	31
no color	2	2	9	1	14
total	8	12	20	5	45
	18%	27%	44%	11%	100%

Table 8 Failed Gauge Summary for Case 1

Thus, based on Table 8, gauge G3 failed the most switches (44%), followed by G2 and G1.

Case 2: Analysis of individual railroad inspections

In this section, the analysis results for the six railroads/consultants summarized in Case 1 are presented individually These six railroads/consultants used the four-condition assessment for the switch point inspections consisting of: Good, Marginal, Poor, and Failed. Thus, the analysis logic used here is the same as that presented in Case 1.

Case 2A: BNSF

This set of BNSF inspections consisted of 19 inspected switches. Table 9 presents a summary of the Case 2A results.

	R+	R-	Y+	Y-	Total R/Y
Count	2		8	1	15
	Count	% of total			
Total	19	100			
Subtotal with one or more indicator	15	79			
Positives	10	53			
Negatives	5	26			
Good condition agreement	4	21			
Total agreements	14	74			
Disagreements	5	26			

Table 9: Case 2A Summary (BNSF)

As can be seen in Table 9, for Case 2A (BNSF Only) there was good agreement between the gauges and the inspector (to include yellow conditions) on 14 of the 19 switches inspected, or 74%.

Table 10 presents a more detailed analysis of these inspections, looking at the data by gauge.

	Failed Gauges count				
	G1	G2	G3	G4	total
Red / yellow	3	7	8	2	20
no color	1	2	3	0	6
total	4	9	11	2	26
	15%	35%	42%	8%	100%

Table 10: Failed Gauge Summary for Case 2A BNSF

Thus, based on Table 10, Gauge G3 failed the most switches (42%) followed by G2.

Case 2B: Long Island Rail Road

This set of LIRR inspections consisted of 19 inspected switches. Table 11 presents a summary of the Case 2B.

	R+	R-	Y+	Y-	Total R/Y
Count	0	1	0	0	1
% of Total	0%	100%	0%	0%	100%
	Count	% of total			
Total	19	100			
Subtotal with one or more indicator	1	5			

Positives	0	0
Negatives	1	5
Good condition agreement	18	95
Total agreements	18	95
Disagreements	1	5

Table 11: Case 2B Summary (LIRR)

As can be seen in Table 11, for Case 4B (LIRR Only) there was excellent agreement between the gauges and the inspector (to include yellow conditions) on 18 of the 19 switches inspected, or 95%.

Table 12 presents a more detailed analysis of these inspections, looking at the data by gauge.

	Failed Gauges count				
	G1	G2	G3	G4	total
Red / yellow	0	0	0	0	0
no color	0	0	1	0	1
total	0	0	1	0	1
	0%	0%	100%	0%	100%

Table 12: Failed Gauge Summary for Case 2B LIRR

Thus, based on Table 12, gauge G3 failed the most switches (100%).

Case 2C: Norfolk Southern

This set of NS inspections consisted of 10 inspected switches. Table 13 presents a summary of the Case 2C results.

	R+	R-	Y+	Y-	Total R/Y
Count	0	0	1	0	1

	Count	% of total
Total	10	100
Subtotal with one or more indicator	1	10
Positives	1	10
Negatives	0	0
Good condition agreement	9	90
Total agreements	10	100
Disagreements	0	0%

Table 13: Case 2C Summary (NS)

As can be seen in Table 13, for Case 2C (NS Only) there was excellent agreement between the gauges and the inspector (to include yellow conditions) on 10 of the 10 switches inspected, or 100%.

Table 14 presents a more detailed analysis of these inspections, looking at the data by Gauge.

	Failed Gauges count				
	G1	G2	G3	G4	total
Red / yellow	1	1	0	1	3
no color	0	0	0	0	0
total	1	1	0	1	3
	33%	33%	0%	33%	100%

Table 14: Failed Gauge Summary for Case 2C NS

Thus, based on Table 14, gauges G1, G2, and G4 each failed the same "failed" switch.

Case 2D: Gary Wolf (consultant)

This set of inspections by consultant Gary Wolf consisted of 35 inspected switches. Table 15 presents a summary of the Case 2D results.

	R+	R-	Y+	Y-	Total R/Y
Count	0	1	1	0	2

	Count	% of total
Total	15	100
Subtotal with one or more indicator	2	13
Positives	1	7
Negatives	1	7
Good condition agreement	13	87
Total agreements	14	93
Disagreements	1	7

Table 15: Case 2D Summary (Gary Wolf)

As can be seen in Table 15, for Case 2D (Gary Wolf Only) there was excellent agreement between the gauges and the inspector (to include yellow conditions) on 14 of the 15 switches inspected, or 93%.

Table 16 presents a more detailed analysis of these inspections, looking at the data by Gauge.

	Failed Gauges count				
	G1	G2	G3	G4	total
Red / yellow	0	0	1	0	1
no color	0	0	1	0	1
total	0	0	2	0	2
	0%	0%	100%	0%	100%

Table 16: Failed Gauge Summary for Case 2D Gary Wolf

Thus, based on Table 16, gauge G3 failed the most switches (100%).

Case 2E: CP Rail

This set of CP rail inspections consisted of seven inspected switches. Table 17 presents a summary of the Case 2E results.

	R+	R-	Y+	Y-	Total R/Y
Count	0	4	1	0	5

	Count	% of total
Total	7	100
Subtotal with one or more indicator	5	71
Positives	1	14
Negatives	4	57
Good condition agreement	2	29
Total agreements	3	43
Disagreements	4	57

Table 17: Case 2E Summary (CP Rail)

As can be seen in Table 17, for Case 2E (CP Rail Only) there was moderate agreement between the gauges and the inspector (to include Yellow conditions) on 3 out of the 7 switches inspected, or 43%.

Table 18 presents a more detailed analysis of these inspections, looking at the data by gauge.

	Failed Gauges count				
	G1	G2	G3	G4	Total
Red /yellow	1	1	1	0	3
no color	1	0	2	1	4
Total	2	1	3	1	7
	29%	14%	43%	14%	100%

Table 18: Failed Gauge Summary for Case 2E CP Rail

Thus, based on Table 18, gauge G3 failed the most switches (43%) followed by G1.

Case 2F: TTCI

This set of TTCI inspections consisted of for inspected switches. Table 19 presents a summary of the Case 2F results.

	R+	R-	Y+	Y-	Total R/Y
Count	0	1	1	0	2
% of Total	0%	50%	50%	0%	100%

	Count	% of total
Total	4	100
Subtotal with one or more indicator	2	50
Positives	1	25
Negatives	1	25
Good condition agreement	2	50
Total agreements	3	75
Disagreements	1	25

Table 19: Case 2F Summary (TTCI)

As can be seen in Table 19, for Case 2F (TTCI Only) there was agreement between the gauges and the inspector (to include yellow conditions) on three of the four switches inspected, or 75%.

Table 20 presents a more detailed analysis of these inspections, looking at the data by gauge.

	Failed Gauges count				
	G1	G2	G3	G4	total
Red / yellow	0	0	1	0	1
no color	1	1	1	1	4
total	1	1	2	1	5
	20%	20%	40%	20%	100%

Table 20: Failed Gauge Summary for Case 4C NS

Thus, based on Table 20, gauge G3 failed the most switches (40%) followed by G1, G2, and G4, respectively.

Summary of Comparisons

Table 21 summarizes the results of the comparisons presented in this section.

Case/Railroad	Number of inspections	% Agreement	% Disagreement
All inspections	74	82	18
BNSF	19	74	26
LIRR	19	95	5
NS	10	100	0
CP	7	43	57
Gary Wolf	15	93	7
TTCI	4	75	25

Table 21: Summary of Comparison Between Inspectors and Gauges

As can be seen in this table, there was generally good agreement with the gauges, ranging from 74% to 100%,[12] and an overall agreement of 82%.

Of the four gauges, gauge G3 recorder the most "fails," representing 44% of all the recorded gauge failures, followed by G2 and G1, respectively.

Probabilistic Assessment of Agreement between Gauges and Railroad Inspectors

As an alternate assessment approach, the researchers performed a probabilistic simulation analysis, where the agreement percentages for the individual gauges and for all gauges combined were used with the following analysis logic based on Monte Carlo Simulation.

1. The agreement percentages for all inspection cases (Table 21) were taken to be the initial probability of agreement. The probabilities distribution is discrete with only two outcomes "Agreement" and "Disagreement." The graph in Figure 13A illustrates the Probability Mass Function (PMF) of the agreement probability for the all gauges case.

[12]Not including CP, which appears as an outlier.

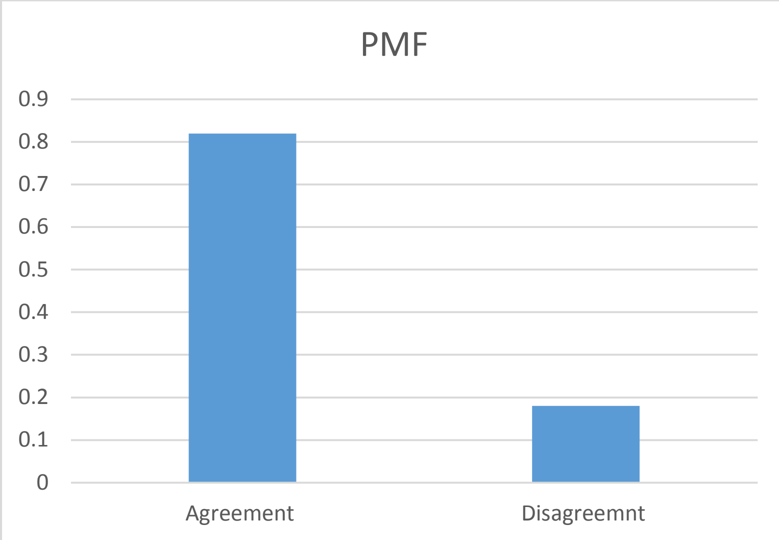


Figure 13A: Probability Mass Function (PMF).

2. A random number generator was used to generate a random number from 0 to 1.
3. 10,000 points were generated and then compared with the probabilities of agreement for each gauge and for all gauges combined. This process was completed by developing a Cumulative Distribution Function (CDF) for each PMF produced. The graph (Figure 13B) illustrates the CDF of the agreement probability for the all gauges case.

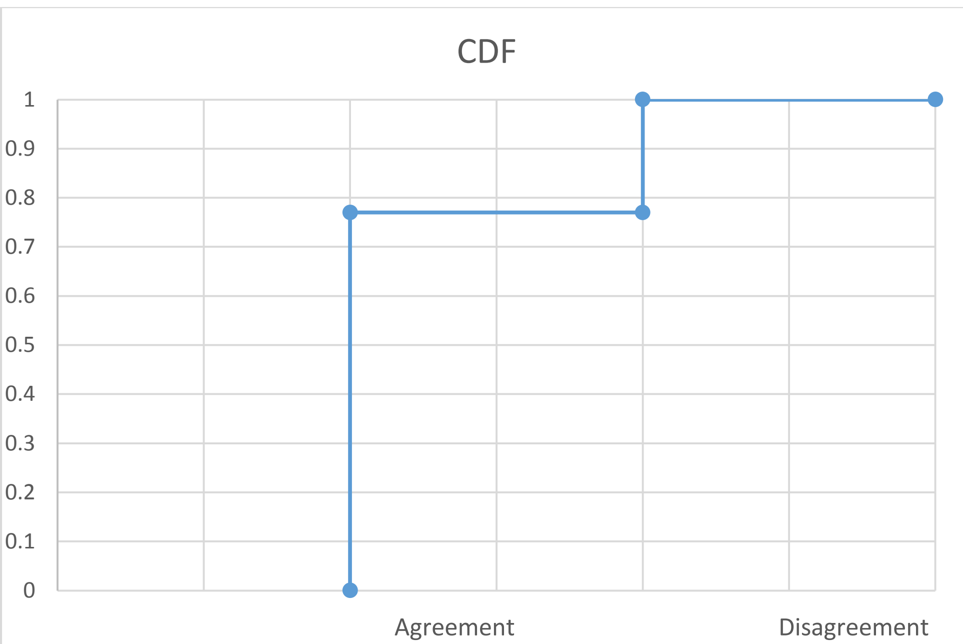


Figure 13B: Cumulative Distribution Function.

4. Based on the CDF each generated point is determined to be an "Agreement" or "Disagreement." For example, if the software generates a number equal to 0.5, then the imaginary situation would be an "Agreement."
5. This probability of agreement of the 10,000 points was calculated 1,000 times. Thus, the researchers ran this simulation 1,000 times and calculated each of the probabilities.

6. The minimum and maximum probabilities that might occur were recorded.
7. The average was calculated and taken as the “agreement.”
8. The process was then repeated for each individual gauge, using agreement percentages based on the data obtained from the field validation.

The results are presented in Table 22.

Gauge	1	2	3	4	All
Max	86%	91.8%	83.2%	85.1%	83.2%
Min	83.9%	90.1%	80.8%	82.9%	80.8%
Chance of Success (Agreement)	85%	91%	82%	84%	82%

Table 22: Probabilistic Analysis Summary

It can be seen that each of the gauges had agreements of no less than 82%, but with some gauges (e.g., gauge 2) having agreement as high as 92%. Furthermore, the overall agreement is 82%, which is consistent with the correlation results presented in Table 21 and supports the validity of the gauges in evaluating switch point condition.

Railroad Comments

In a manner similar to that presented after the first set of field evaluations, additional comments were solicited from the railroad users, particularly those who were new to this round of inspections. The following comments were obtained from the railroad inspectors.

CP Rail
“…. the use of these gauges has the potential to help reduce switch point derailments, and also help maintenance forces prioritize switch point replacement. The gauges provide a consistent means of evaluating switch points, and reduce the subjective opinion as to what constitutes a switch point in need of replacement. It has been my experience that G1 and G4 measurements are the most frequent causes of derailments. Often in these situations the point is worn, not chipped, or the angle of the point is flattened to where the flange will climb the side of the point. The G2 and G3 gauges are also beneficial.”

NS
“With the everyday methods of inspecting a turnout combined with the concepts of the switch point inspection gauge (SPIG), a more detailed inspection is performed. The SPIG was a great tool for me in learning what to look for when inspecting a point. It increased my awareness on how the wheel actually interacted with the point. Also, it helped me realize different situations that benefited the point such as point protectors, adequate tension and ensuring there were no gaps between the points and stock rail. This gauge caused me to ask the proper questions during inspections, and helped my judgment with determining if a switch point is in correspondence.”

TTCI
“We have used the gauges on a variety of switches, including FAST HTL and the Pueblo Chemical Depot. …. We are impressed that the gauges seem “about right” as compared to our visual ratings.”

These comments, combined with the comments provided by the different railroads during the initial field evaluation (see Task 2, preliminary railroad comments, in this report and in Reference 7), show that the railroad inspectors and supervisory personnel view these gauges as useful and valuable in helping evaluate switch point condition.

Conclusions and Recommendations

This project was an implementation and field assessment of a set of four gauges, originally developed under the earlier IDEA Project S-23 activity, with the goal of providing railroads and transit systems with a set of practical and directly useable gauges for track inspectors to evaluate the condition of switch points. Field assessment of these gauges was performed in two steps; an initial assessment activity, after which modification were made to the gauges to reflect the field experience, and a final assessment activity.

During both field assessments, approximately 350 switches were inspected and evaluated by six railroads and one consultant. The evaluations included direct correlation between inspector and gauges, statistical analyses, and Decision Tree analyses all looking at the effectiveness of the gauges in identifying switch points that are in a condition that can contribute to or result in a derailment. The gauges themselves were designed as simple “go/no go” gauges to advise the inspector as to whether the switch point “passes” or “fails.” Each gauge addresses a different failure mechanism and potential derailment condition.

The results of the two field inspections showed extremely good agreement, with overall ratings in the range of 80%; in line with the statistical analysis. In several cases agreement was over 90% (and in one case 100%). Even the two cases where the railroad inspectors had moderate agreement with the gauges (of the order of 50%), the direct feedback indicated that the inspectors viewed the gauges as being “beneficial” in identifying poor switch points and reducing derailments.

Other observations showed particularly good agreement with experienced inspectors, while other railroad inspectors may not be familiar with certain derailments categories for which specific gauges are intended.

Feedback from the railroads and inspectors was very favorable, noting that the gauges addressed several conditions which are often overlooked, such as high gage face wear (Gauge 4) and in general made for a more detailed inspection of the switch point. As one inspector noted in his report, “The Switch Point Inspection Gauges were a great tool for me in learning what to look for when inspecting a point. It increased my awareness on how the wheel actually interacted with the point …. This gauge caused to me to ask the proper questions during inspections, and helped my judgment with determining if a switch point is in correspondence.” This is in line with some of the Review Panel’s views that the gauges will help train inexperienced inspectors and refine their judgment.

Based on the results of this program, the committee has recommended that these gauges be made available to the railroad industry as well as to such railroad standards organizations as AREMA for industry implementation. The review panel has already received a request from AREMA committee 5 to review the gauges, and the specific information on the gauges will be forwarded to the relevant AREMA committees for review and potential incorporation into AREMA’s *Manual of Recommended Practices*.

In addition, the University of Delaware team and the Review Panel will begin to disseminate information about these gauges through railway media, conferences, and organizations. This is to include preparation of articles in *Railway Track & Structures*, presentation at AREMA and TRB annual conferences, and direct discussion with railroads and railway suppliers. It should be noted that at least one supplier has expressed interest in manufacturing and providing these gauges.

REFERENCES

1. Federal Railroad Administration, Track Safety Database [Online]. Available: www.fra.dot.gov.

2. Zarembski, A.M., *Safety IDEA Project 23 Final Report*: *Reducing Wheel Climb at Switch Points to Reduce Derailments*, Transportation Research Board of the National Academies, Washington, D.C., April 2014.

3. Network Rail, Inspection and Repair to Reduce the Risk of Derailments at Switches, Oct. 2007, NR/L2/TRK0053.

4. SBB (Swiss Federal Railways), Anleitung zum Gebrauch der Kontrolllehren fur die Prufung der Entgleisungssicherheit in Weichen, (Use of Inspection Gauges for Testing Switch Points for Derailment Safety), May 2009.

5. Zarembski, A.M., “Development of Rail Gage Face Angle Standards to Prevent Wheel Climb Derailments,” American Railway Engineering Association Annual Technical Conference, Chicago, Ill., Mar. 1996.

6. Hall, M., F. Eibe, G. Holmes, B. Pfahringer, P. Reutemann, and I.H. Witten, The WEKA Data Mining Software: An Update, *SIGKDD Explorations*, Vol. 11, Issue 1, 2009.

7. Zarembski, A.M., *Safety IDEA Project 28 Stage 1 Report*: *Field Validation of Inspection Gauges for Wheel Climb Safety at Switch Points*, Transportation Research Board of the National Academies, Washington, D.C., April 2016.

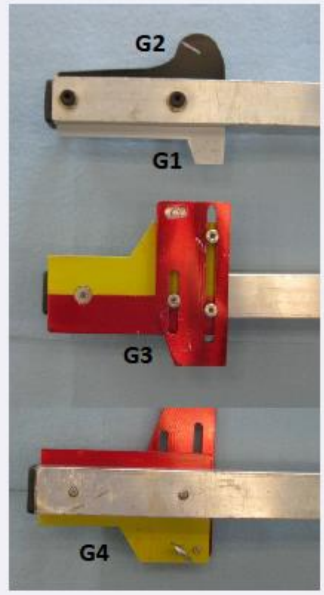
APPENDIX A

Gauge Instruction Sheets for Final Versions of Gauges

Instructions for Using Switch Point Inspection Gauges

Revised May 8, 2016

- G1 -Chipped point
- G2 -AAR 1B wheel contact
- G3 -Severely worn wheel profile
- G4 -Gage-face wear angle



1

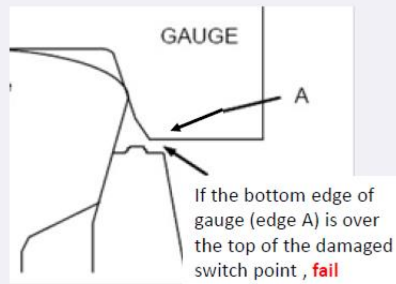
When evaluating a switch point that has been ground back

Apply gauges starting at the ground-back tip (yellow arrow)

G1 - Chipped point

Apply this gauge where the point is worn or broken. This is typically within 24 inches of the tip, but could be anywhere along the point.

- If the bottom edge slides <u>over</u> the top of a broken or worn point, **Fail**
- Otherwise, **Pass**



3

G2 - AAR 1B wheel contact

Apply this gauge where the point is worn or broken. This is typically within 24 inches of the tip, but could be anywhere along the point.

If contact with the switch point is…

- above the 60˚ mark, **Pass**
- below the 60˚ mark, **Fail**

4

G3 - Severely worn wheel profile

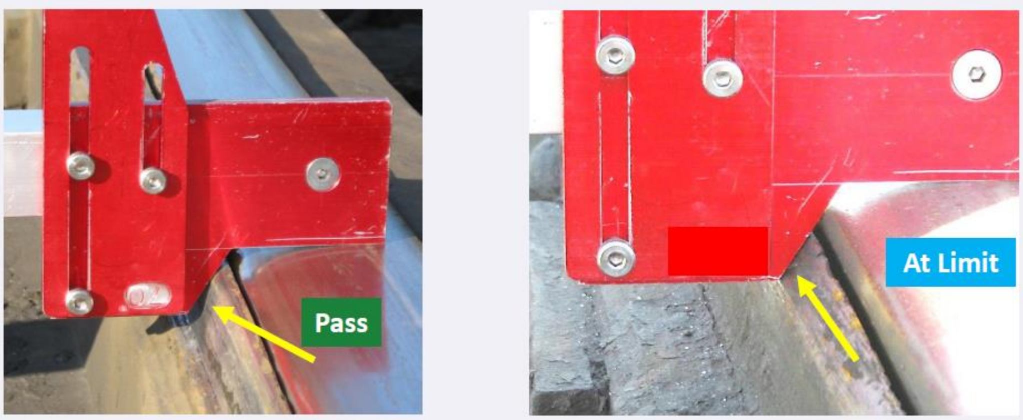
Apply this gauge within the first inch of the tip of the point, or where wheel flange contact is first evident. Start with the 70° slider on top of the stock rail, then slowly shift the gauge until the slider slides off the stock rail.

- If the slider slides down the gage face of the switch point, **Pass**
- If the corner of the slider lands on top of the switch point, **Fail**

This gauge can be used to identify gapping points, broken points and points that are exposed due to a worn stock rail.

5

G4 - Gage-face wear angle



Apply this gauge anywhere along the length of the point where it is worn.

If the switch point gage-face angle is...

- greater than 60˚, **Pass**
- less than 60˚, **Fail**

This gauge protects against a low gage face angle, which is more likely to contribute to wheel climb.

6