

The work described in this document was performed by Transportation Technology Center, Inc.,
a wholly owned subsidiary of the Association of American Railroads.

A Review of Wheel Impact Measurement Variation

RB Wiley and Abeer Elsalebby

Summary

The repeatability of wheel impact load detector (WILD) measurements is often questioned because removal of high impact wheels constitutes a major expense to car owners. The WILD impact history of a typical out of round wheel can vary. The variations can appear to be quite large due to varying speeds and load conditions. Other factors related to the defect location on the wheel tread and with the tread position on the rail can also affect measurement consistency. Thus, the repeatability questions are understandable and to be expected.

A WILD data review and analysis for an example set of wheels was conducted. These wheels were selected for review by the owner because they had WILD alerts (i.e., at opportunistic repair or condemnable levels). The primary conclusion of these studies is that repeatability of optimal WILD measurements is more a function of normal variation in rail vehicle operating circumstances, with no relationship to the measurement system's capability itself. Specific to the example wheels, all were shown to be out of round, confirmed by reliable evidence from multiple readings from WILDs.

The analysis was intended to qualify and quantify the nature and causes of measurement variation. WILD data histories for 60 wheels from 56 cars were reviewed.

The analysis results for these 60 wheels showed that anomalous readings were not the cause of the reading variation. Of the 60 wheels, seven remained in the category of the original basic question, that of having a single high reading (i.e., qualifying as an opportunistic repair or as a condemnable level alert). These seven wheels showed clear evidence of being out of round, but had reduced opportunities to obtain optimum WILD measurements (e.g., low mileage or mostly measured while empty).

A separate review of industry-wide data for alerted wheels of the same levels also indicates consistently repeatable, although variable, vertical impact force measurements. An extension to this review shows a strong correlation between repeatability of high wheel impact readings and frequency of opportunities to optimally measure the same wheels at a WILD site. Conversely, reduced opportunities yielded low counts of repeatable readings. These results will be reported in a subsequent *Technology Digest* (TD).

This TD addresses results and conclusions for the Phase 1 wheel examples review. Next, Phases 2 and 3 will expand the data in an industry-wide study to establish the level of confidence in WILD measurements while considering a potentially larger set of variation causes. A follow-on TD will describe the results and conclusions from the industry-wide review.



BACKGROUND AND INTRODUCTION

To characterize wheel impact measurement repeatability, a review and analysis of WILD data was conducted. The intent was to characterize the example data provided by a car owner. The basic question was why some WILD impact measurements appear to be inconsistent and therefore, should wheel alert rules consider these variations in readings when identifying out of round wheel defects. For this review, the question was limited to evaluating WILD data variations from the examples of 60 wheels from 56 cars. A follow-on study will evaluate and compare these results with those for a broader range of industry-wide data.

The data characterization followed a stepwise analysis process that categorized the subject wheels according to similarities in WILD readings, travel frequency, speed, and load condition. The methodology provided understanding about the factors that limit the opportunities to optimally measure wheel out of round defects consistently with a WILD.

The study evaluated and if possible quantified the nature of WILD wheel impact measurement variations, relative to major contributing factors. It also identified important considerations required to qualify or mitigate the analysis results according to the causes of variation.

This TD addresses the Phase 1 WILD examples review. A follow-on TD will describe the results and conclusions from the industry-wide review (Phases 2 and 3).

ANALYSIS METHODOLOGY

Transportation Technology Center, Inc.’s standard methodology for analyzing detector data involves two initial analysis steps:

1. Exploratory Data Analysis (EDA): Analyze the data for the purpose of gaining maximum insight into that data. Produce frequency tables, descriptive statistics, categoric breakdowns, distributions and associated plots, charts, and graphs. Often, 80 percent of the required answers are found during EDA.
2. Determine data reliability: If a reading (e.g., WILD max_vert) is questioned, evaluate the subject data as meeting the hypothesis (H_0) of being good, reliable data versus the alternative (H_a) that the reading is anomalous or improperly assigned to the subject wheel or vehicle.

It was determined that a much larger date range query was required for complete analysis of the wheel histories. The sample data provided was for 60 days or less (beginning March 14, 2009).

Wheel data histories were identified with higher or questioned readings. To assist in the H_0 - H_a determination, more data was required to observe for trends toward the apparent single higher reading or simply to find previous readings indicating out of round or not. Therefore, the analysis for all example wheels was based on data recorded at WILDs from January 1, 2009, through May 9, 2010.

Scenarios for Classifying WILD History of the Subject Wheels

Following the EDA, the data reliability test was run to determine acceptance or rejection of H_0 (i.e., that the example reading in question is from good and reliable data). This test was primarily based on the dynamic vertical force measurement ($dyn_vert = max_vert - avg_vert$). Each wheel history was examined individually. A rule of thumb evaluation is that repeated readings of $dyn_vert \geq 20,000$ pounds indicate an out of round wheel. Figure 1 shows the evaluation.

The left plot in Figure 1 shows a normal (i.e., not out of round) wheel’s dynamic vertical force measurements over time. All of the values were below 20,000 pounds. The right plot shows an out of round wheel history from the same vehicle. Generally, from December 2005 through July 2007 almost all of the dynamic vertical values were above 20,000 pounds. Between July and August 2007, the wheelset was replaced, and subsequent dyn_vert readings were well below 20,000 pounds.

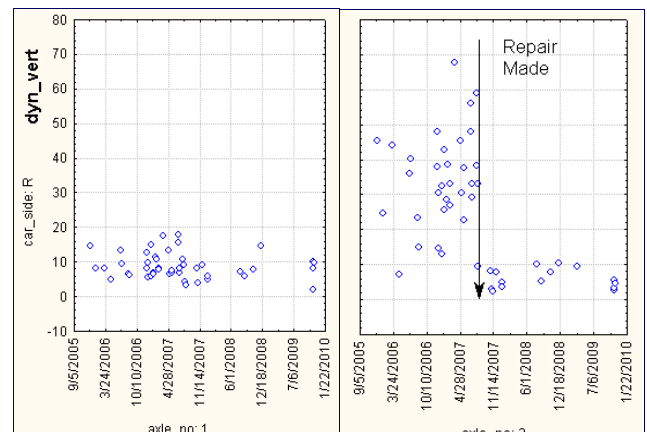


Figure 1. Not Out of Round Wheel (left); Out of Round Wheel with Repair (right)

It was, therefore, easy to distinguish between a good and an out of round wheel when extended data was available. Often the pattern in the data is not so clearly defined.

The examples given in Figure 1 are plotted without any adjustments for speed or load.

The test for data reliability assumed no automatic equipment identification tag swap problem where the complementary wheel on the car is the actual out of round wheel, but the detector assigns the location, and thus the high reading, to the wrong wheel (e.g., L1 readings are assigned to R4, R2 to L3, etc.). Other considerations were adequate speed is ≥ 20 mph and that the vehicle has no “bad data” flags, such as unknown end-leading or erroneous axle or truck counts. These factors were identified and filtered out during EDA.

Wheels passing the acceptance test H_0 (confirmed out of round) were evaluated further to classify the repeatability of the WILD readings. Table 1 shows the WILD reading variation hypothesis descriptions.

Table 1. WILD Reading Variation Hypothesis Descriptions

Symbol	Hypothesis (Scenario) Description
H _{1a}	Single higher reading
H _{1b}	Single higher reading; repeated readings below but relatively near (within 15- to 20-thousand pounds) the subject threshold (80- or 90-thousand pounds)
H ₂	Repeat higher readings found dated after the example data readings
H ₃	Repeat higher readings found dated before the example data readings
H ₄	Repeat higher readings found dated before and after the example data readings
H ₅	Single higher reading; warp or truck hunting found
H ₆	Mate wheel with higher level or more advanced out of round

A number of factors will reduce or limit the opportunities to obtain optimal WILD measurements. Each subject wheel was classified into the previously described hypothesis scenarios by observing the existence of these factors in the data history. Table 2 shows the factors that influence the repeatability or magnitudes of WILD readings.

Table 2. Factors Limiting Opportunities to Obtain Optimal WILD Readings

Symbol	Factor Description
F ₁	Number of visits to WILD sites
F ₂	Number (or percentage) of loaded visits to WILD sites
F ₃	Zero additional visits to WILD sites
F ₄	Reduced speed (< 50 mph)
F ₅	Truck warp or hunting causing inconsistent readings
F ₆	Wheelset replaced

Mitigation of Variation Factors

Some variation factors can be mitigated by standardizing (adjusting) readings, i.e., for speed and load.

The speed adjustment factor is based on the ratio of the train speed when the reading is taken versus a standard speed of 50 mph. The evaluation mode and formula used are: if the train speed is between 20 mph and 50 mph, then calculate a speed corrected (SC) dynamic vertical force (SC_dyn_vert) as follows: (50/spd)*dyn_vert (referred to as speed adjusted in Table 4). To calculate a speed corrected maximum vertical force (SC_max_vert), add the SC_dyn_vert to the measured nominal vertical force (avg_vert).

The load adjustment was made by finding the largest avg_vert value for each subject wheel and adding it to the respective SC_dyn_vert. This was done as a final step and includes both speed and load adjustments (LCSC) for simultaneous evaluation as the LCSC_max_vert (referred to as Load Adjusted in Table 4).

The 60 subject wheels were then reclassified and tallied according to both the SC_max_vert and the LCSC_max_vert.

The subject wheels that remained classified as H_{1a} (having a single higher reading) were evaluated for the factor F₅ (having hunting or warp characteristics). These two truck performance problems affect wheelset tracking.¹ The wheel out of round defect position on the tread, relative to the tread position on the rail, can affect the magnitude and consistency of the vertical force measurement. This will further limit the opportunity to obtain optimal WILD measurements.

None of the subject wheels exhibited the questioned variation due to a wheelset replacement (factor F₆); however, this is a common cause of suddenly reduced readings.

RESULTS

The 60 wheels were categorized according to their initial WILD reading threshold level measured in thousand pounds (i.e., opportunistic repair or condemnable), as Table 3 shows.

Table 3. Example Wheels with Questionable Higher WILD Readings, Grouped by Alert Level Impact Range

Opportunistic and Condemnable Event Examples			
	Count (N)	(80,000-90,000 pounds)	≥ 90,000 pounds
Wheels	60	41	19
Cars*	56	40	18

* Of the 56 cars represented, two have more than one wheel included in the sample set.

Following the EDA, the data reliability test confirmed H₀ (i.e., that all the example wheel readings in question are good and reliable data). In other words, all example wheels were out of round, confirmed by reliable evidence from multiple readings from WILDs. All of the wheels were then reclassified and tallied according to hypothesis scenarios H_{1a} to H₆, as Table 4 shows.

The SC_max_vert values were reclassified according to their new hypothesis scenario placement using the same definitions, as were the LCSC_max_vert values. Table 4 columns 3 and 4 show the adjusted values for the hypothesis scenarios.

Table 4. Cross Tabulation: Count (%) of Example Wheels, by Hypothesis Scenario Before/After Mitigation Adjustments

Hypothesis Scenario	WILD Readings		
	Unadjusted	Speed Adjusted	Load Adjusted
H _{1a}	3 (5.00%)	3 (5.00%)	0
H _{1b}	29 (48.33%)	13 (21.67%)	7 (11.67%)
H ₂	4 (6.67%)	6 (10.00%)	9 (15.00%)
H ₃	21 (35.00%)	22 (36.67%)	25 (41.67%)
H ₄	3 (5.00%)	16 (26.67%)	18 (30.00%)
H ₅	0	0	1 (3.67%)
H ₆	0	0	0

A key finding relating to Hypotheses 2 to 4 is that it is necessary to review a broad date range of WILD data before attempting to draw conclusions. As Figure 1 shows, a wheel with WILD readings that reflect out of round performance may do so over a long period of time (> 2 years). This time frame may reflect low mileage and few readings, high mileage and many readings, or a combination of high mileage but few readings. Having as much data as possible is necessary to assure having the most complete picture of a wheel's out of round health. Other information regarding mileage, ton-mileage, routes, repairs, storage, etc. may be important factors necessary for interpreting the WILD data.

The right-hand plot of Figure 1 shows that the dynamic vertical readings, indicating an out of round condition, existed for over 18 months. However, the beginning of this condition was not apparent prior to this particular dataset. For this particular wheel, a longer prior history was needed to see the beginning of the out of round condition and when the dynamic vertical readings initially exceeded 20,000 pounds.

CONCLUSIONS

All questionable WILD reading examples were confirmed as good, reliable data. Further, the example wheels were all confirmed to be out of round with valid alerts (i.e., repeated readings of dyn_vert ≥ 20,000 pounds).

Table 4 shows that WILDs do in fact produce wheel impact measurements that are not repeated at or near the same level of magnitude. From the example, these occur at a relatively small rate (5 percent), but are still considered significant. To fully understand the nature of these nonrepeatable readings, one must consider whether the initial observations do, in fact, represent the complete health picture for the sample (i.e., the example wheels).

It was found that the query time frame required substantial expansion (from 60 days to more than 15 months). The more complete set of data supports the determination of multiple possible scenarios that influence the repeatability of higher WILD readings for the example wheels. Five of six possible scenarios (H_{1b}-H₅) contributed to the limiting of the opportunities for obtaining optimal WILD measurements. These five scenarios are the primary factors that generally influence all WILD measurements. However, it is possible that given a larger sample size additional factors may be found, including wheels representing scenario H₆.

Two of the observed factors can be standardized: Speed and Load. Speed adjustments were made, and then an adjustment for wheel load followed. When speed and load variation were standardized, almost 87 percent of the example wheels did have repeatable higher readings. Only 10 percent were not repeated at or above the given threshold. However, these did have repeat readings below but near the alert threshold (i.e., >70,000 pounds for the 90,000-pound threshold and >65,000 pounds for the 80,000-pound threshold).

This sample did not have any wheels without repeatable high impact readings. However, it is expected that a much larger sample would. This sample was as provided, being of very limited size and not representative of the broad range of car types, WILD sites, and operating circumstances of the whole North American freight rail industry.

REFERENCES

1. Wu, H. and B. Madrill. June 2008. "Effect and Formation of Asymmetrically Worn Wheels," *Technology Digest* TD-08-021, Association of American Railroads, Transportation Technology Center, Inc., Pueblo, Colo.