Wayside Defect Detector Data Mining to Predict Potential WILD Train Stops

Leila Hajibabai[a], M. Rapik Saat[a], Yanfeng Ouyang[a], Christopher P.L. Barkan[a], Zengyi Yang[a],

Kim Bowling[b], Kamalesh Somani[b], Don Lauro[b], and Xiaopeng Li[a]

[a]*Department of Civil and Environmental Engineering, University of Illinois at Urban-Champaign*

[b]*CSX Transportation*

Abstract

Advanced wayside detector technologies can be used to monitor the condition of railcar components, notify railroads of probable failures to equipment and infrastructure in advance, and predict alert of imminent mechanical-caused service failures. Using some statistical data-mining techniques, historical railcar health records from multiple Wayside Defect Detector (WDD) systems can potentially provide the essentials to recognize the patterns and develop the reliable and innovative rules to predict the failures and reduce related risks on railroads. In this paper, data from Wheel Impact Load Detector (WILD) and Wheel Profile Detector (WPD) were analyzed through comparing historical measurements for failed and non-failed wheels on the same truck to predict train stops due to high impact wheels. An exploratory data analysis was performed to identify the most critical measurements from each detector by comparing the distributions of several measurements from failed wheels to the ones from non-failed wheels. A logistic regression approach was used to predict the probability of potential high impact wheel train stops. Initial results show a 90% efficiency to predict the failure within 30 days after the most recent WILD reading.

Keywords:

wayside defect detector, prediction, failure, non-failure, statistical analysis, logistic regression, WILD, WPD

1 Introduction

Train accidents, either track-or railcar-related causes, imposes huge costs, safety concerns, and affects train operations and efficiency. The likelihood of railcar mechanical failures may be reduced by (1) improving the mechanical components' resistance to bear higher stresses and loads, and/or (2) using wayside defect detectors (WDD) to monitor the performance of equipment and components to take proactive actions before mechanical failures occur.

Different types of WDD are used to monitor railcar health and/or detect immediate hazards that can result in train derailments. WDD can be broadly classified into two types (Lagnebäck 2007, Ouyang et al 2009):

Reactive Systems detect railcar component conditions that have a short latency between detectability and failure of the component. Failure to react promptly to these alarms may result in damage to railroad infrastructure and/or a derailment. Dragging Equipment Detectors (DED) and Hot Bearing Detectors (HBD) are widely used examples of reactive WDD technologies.

Predictive Systems are capable of measuring, recording, and trending the performance of vehicles and specific components. The information collected can be used to analyze the condition of equipment to predict possible failures and faults that may occur sometime in the future, thereby making it easier to plan maintenance activities. Wheel Impact Load Detectors (WILD) and Wheel Profile Detectors (WPD) are commonly used examples of these types of technologies on North American railroads. Some predictive technologies, such as WILD, may also detect certain imminent hazards and therefore can be classified as both predictive and reactive.

Statistical data analysis and data mining approaches can offer the basis to recognize the patterns and dependencies between various factors derived from the WDD data. These sets of analyses help generate innovative, reliable, and advanced rules to predict the occurrence of equipment failures, service disruptions, or possibly mechanical-caused derailments. Therefore, to reduce inspection costs and frequencies, and optimize railcar maintenance schedule, statistical and data mining techniques can be used to analyze historical railcar health record data from multiple WDD systems.

This paper describes an efficient data mining framework to analyze data from different WDDs and predict high impact wheel train stops in advance. Data from WILD and WPD detectors were analyzed through comparing historical measurements for failed and non-failed wheels in the same truck. Comparing the distributions of failed and corresponding non-failed detector readings for wheels helped identify critical measurements from each detector. Then, to predict the probability of potential high impact wheel train stops, logistic regression models were developed. The exposition of the paper is as follows. First, in Section 2, methodology including the exploratory analysis and regression approach will be discussed. Afterwards, data analysis results and performance measures will be introduced in Section 3, and conclusions will be discussed in Section 4.

2 Methodology

Detailed analysis of individual data sources enables further understanding of information availability and limitations. Since data obtained from multiple WDD is significantly large, the preprocessing and analyzing of the combined data from different sources is non-trivial and requires significant time and effort. Therefore, an efficient data mining framework is required to manage the large databases and prepare useful data for further railcar condition analysis. In this section, a data analysis approach to determine the failed and non-failed railcar wheels will first be introduced (Figure 1). Then the method to determine the *potential* and *true* failures and corresponding non-failures will be discussed. The exploratory approach to identify the most critical variables from the data attributes will be applied. Finally, those critical variables are considered in a regression model, which was developed to predict the failures based on the identified significant variables.

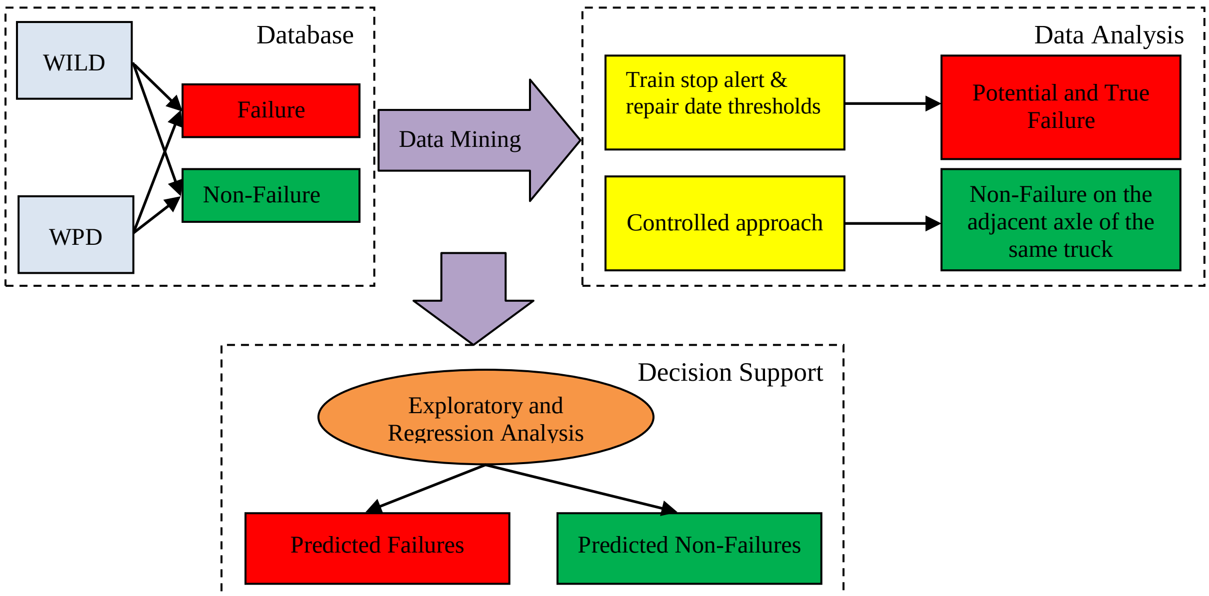


Figure 1. Data mining framework for condition monitoring to predict wheel failures

2.1 Data Analysis Approach

The WILD and WPD data were analyzed on railcars with high impact wheel train stops, and dates of these car repair billing records to identify railcar wheel failures, and non-failures. If an instance of train stop was recorded for a wheel it would be categorized as failure; otherwise, with no train stop nor repair record, it was categorized as a non-failed wheel.

Wheels in the failure category were further analyzed according to the time span between the most recent detector measurement and the train stop alert date (M-TS), and similarly based on the time span between each high impact train stop and repair dates (TS-R) (Figure 2).Certain thresholds for M-TS and TS-R were defined and tested to determine the railcar wheel *potential* vs. *true* failures. Data availability was an important issue to determine the appropriate thresholds. A maximum threshold of 30 days was selected for M-TS and TS-R to determine the *potential* and *true* failures for wheels. In other words, if reading date was within 30 days before the high impact train stop date, the wheels would be categorized as *potential* failures; and, the wheels would be classified as *true* failures if wheel repair date was within 30 days after the train stop.

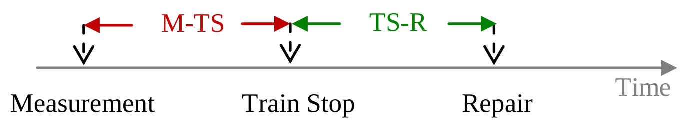


Figure 2. Time span approach to predict wheel failures

To reduce uncertainty and variance from different truck characteristics, a controlled approach has been developed by comparing failed and non-failed wheel historical measurements on the same truck. Therefore, the wheels with no records of high impact train stop nor wheel repair, that were located on the adjacent axle to the failed wheel would be considered in the non-failure category (Figure 3).

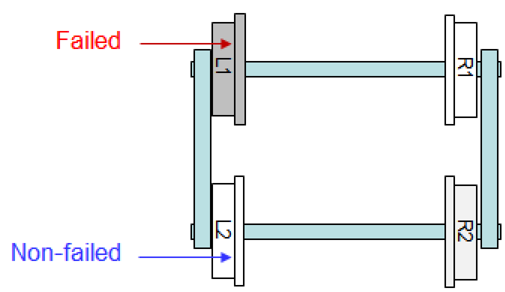


Figure 3. A controlled approach to compare failed and non-failed wheel measurements

2.2 Exploratory Analysis

After determining the failures and non-failures, a set of data attributes from each detector was selected for the exploratory analysis. From WILD measurement, vertical average weight, vertical peak force, lateral average force, lateral peak force, and dynamic ratio were selected. Flange height, flange thickness, rim thickness, tread hollow, wheel diameter, groove tread, vertical flange (flange angle), tread built-up, and wheel angle were considered for WPD records. Based on each detector database, the time distribution of failures of each variable was compared to the average of non-failures of the same variable. Then, percentage of the failures above (or below) the non-failures average was determined.

If a large percentage of failed-wheel measurements were higher (or lower) than the average measurements for non-failed-wheels, the variable would be considered as a critical variable. Figure 4 shows this exploratory analysis for the vertical average weight from WILD database. The variable is deemed critical since a large percentage of failed-wheel measurements were higher than the average measurements for the non-failures. Using this approach, vertical average weight, vertical peak force, lateral average force, lateral peak force, and dynamic ratio were all considered as critical variables for WILD data. Vertical flange, rim thickness, tread built-up, and wheel angle were the identified critical variables from WPD measurements.

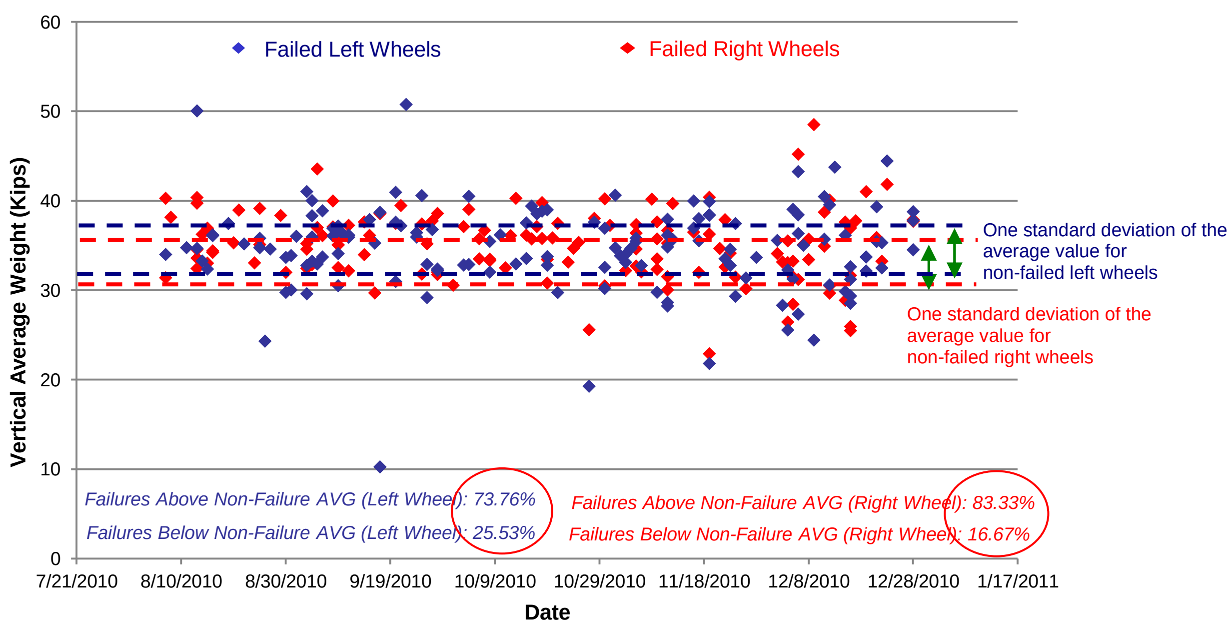


Figure 4. Exploratory analysis to determine the critical variables from each detector

2.3 Regression Analysis

A regression model is needed to predict the failures using the significant variables identified in Section 2.2. In order to predict the outcome of a categorical variable according to a set of predictor variables, a logistic regression was developed. The probability of a possible outcome was modeled as a function of exploratory variables. In this paper, the logistic regression setting includes detector settings $x_0, x_1, ..., x_n$ for a generic instance, regression coefficients as $b_0, b_1, \ldots, b_n$, and logit measure that is $z = b_0 + b_1 x_1 + b_2 x_2 + ... + b_n x_n$. Since there were two main categories of failures, two regression models could be considered as alternatives: regression model for *potential* failures and *true* failures. In this paper, we focused on the logistic regression model based on *potential* failures. 85% of the data was considered to develop the model, and 15% of the data was used to evaluate the effectiveness of the model to predict failures. Based on the *potential* failures, we developed three regression models using: (1) WPD critical variables, (2) WILD critical variables, and (3) combined WPD and WILD critical variables. Unit-less values for all critical variables were used in the regression models according to the (measurement – non-failure mean)/non-failure mean)×100 formulation.

Data structure for data analysis based on WILD *potential* failures was considered in two separate regression models: (1) ignoring dynamic ratio and (2) including dynamic ratio as a critical variable, as follows:

(1) Regression model for WILD data when dynamic ratio is ignored:

$$P(z) = \frac{1}{1+e^{-z}}$$
$$z = 0.4002 - 0.0211x_1 + 0.0261x_2 - 0.0031x_3 - 0.6290x_4 \qquad (1)$$

where

x_1 = vertical average force

x_2 = vertical peak force

x_3 = lateral peak force

x_4 = loaded/unloaded (0 or 1)

The regression model was found to be significant with p-value of 0.002.

(2) Regression model for WILD data when dynamic ratio is considered:

$$P(z) = \frac{1}{1 + e^{-z}}$$
$$z = 0.4103 + 0.0256x_1 - 0.6036x_2 \qquad (2)$$

where

x_1 = dynamic ratio

x_2 = loaded/unloaded (0 or 1)

The regression model was found to be significant with p-value of 0.013.

The following regression model for data analysis based on WPD *potential* failures was developed:

$$P(z) = \frac{1}{1 + e^{-z}}$$
$$z = -1.1029 - 0.0416x_1 + 0.0272x_2 \qquad (3)$$

where

x_1 = rim thickness

x_2 = flange angle

The regression model was found to be significant with p-value of 0.031.

For WILD and WPD data combined, the following regression model *without* dynamic ratio was developed.

$$P(z) = \frac{1}{1 + e^{-z}}$$
$$z = -2.4402 + 0.0300\, x_1 \qquad (4)$$

where

x_1 = vertical peak force

The regression model was found to be significant with p-value of 6.283e-10.

When dynamic ratio was considered with combined WILD and WPD data, the following regression model was developed:

$$P(z) = \frac{1}{1 + e^{-z}}$$
$$z = -2.3057 + 0.0317\,x_1 + 0.0107\,x_2 \tag{5}$$

where

x_1 = dynamic ratio

x_2 = lateral average force

The regression model was found to be significant with p-value of 1.305e-09.

2.4 Regression Model Test

As described above, the regression models were developed based on 85% of the data. Afterwards, using the remaining 15% unseen data, a test for each regression model was conducted. Below an example of the test is presented for regression model (1) for WILD data when dynamic ratio is ignored. Table 1 and Figure 5 illustrate the number of true failures, false positives, and false negatives based on predictions of this regression model. As it is shown, predicted true failures and false positives (non-failures that are predicted as failures) decrease by increasing the probability of failure, while false negatives (failures that are predicted as non-failures) increase. Therefore, it is useful to look at the summation of the two errors. The minimum total errors can be found at 50% failure probability (Table 1 and Figure 5).

Table 1. Determining the failure probability at the minimum total errors for regression model (1)

Failure Probability	>=20%	>=30%	>=40%	>=50%	>=60%	>=70%	>=80%	>=90%
True Failure	101	95	94	91	85	81	76	59
False Positive	33	31	15	5	3	3	2	2
False Negative	0	6	7	10	16	20	25	42

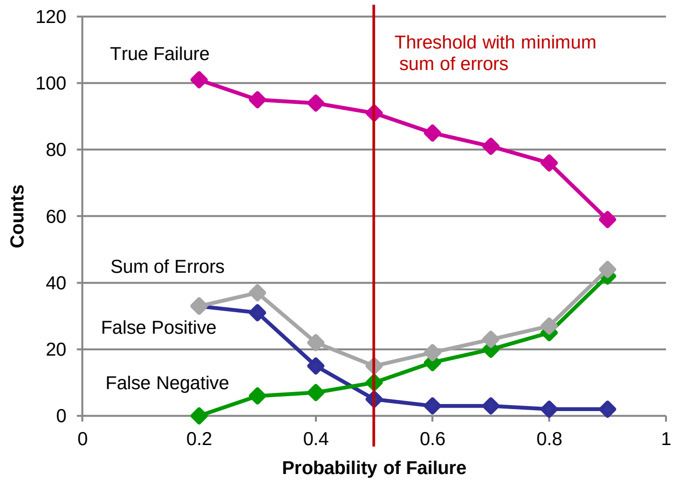


Figure 5. Threshold with minimum total errors for regression model (1)

The confusion matrix for data mining model evaluation, where failure probability was above 50% is illustrated in Table 2-a and 2-b. Probability of Type I error (non-failures that are predicted as failures) was 15% and probability of Type II error (failures that are predicted as non-failures) was 10%.

Table 2. Confusion matrix for data mining model evaluation for regression model (1)

(a): count		Prediction	
		Failure	Non-failure
Actual	Failure	91	10
	Non-failure	5	28

(b): percentage		Prediction	
		Failure	Non-failure
Actual	Failure	90%	10%
	Non-failure	15%	85%

3 Results

Table 3 summarizes the performance results of all five regression models using the unseen data.

Table 3. Summary of regression model performance using the 15% unseen data

Regression Model	True Failure	Type I Error	Type I Error
WILD without dynamic ratio	90%	15%	10%
WILD with dynamic ratio	89%	15%	11%
WPD	40%	0%	60%
WILD and WPD without dynamic ratio	20%	0%	80%
WILD and WPD with dynamic ratio	60%	0%	40%

As Table 3 illustrates, WILD data leads to an efficient prediction of failures. We found promising results with about 90% efficiency to predict high impact wheel train stops in the next 30 days based on the most recent WILD measurements. Failure prediction for WPD model is not as efficient as WILD model, since the number of records in WPD database were much fewer than WILD readings. Furthermore, due to the lack of symmetry between the total number of WILD and WPD records, development of a reliable combined model was prohibited at this time.

4 Discussion and Conclusions

In this paper, data from WILD and WPD databases were analyzed through comparing historical measurements for failed and non-failed wheels on the same truck to predict train stops due to high wheel impacts. Through developing a controlled approach, we performed an exploratory data analysis to identify the most critical measurements from each detector by comparing the distributions of several measurements from failed wheels to the ones from non-failed wheels on the same truck, which reduces variance to identify critical variables and develop efficient prediction models. Then, a set of logistic regression models was developed to predict the probability of occurrence of potential high wheel impact train stops. Results show a 90%

efficiency to predict high wheel impact train stops within 30 days after the most recent WILD measurements.

The lack of symmetry between the total number of WILD and WPD records prohibited development of a reliable combined model. Moreover, difficulty in analyzing combined data from different sources/detectors reflects the need to focus on data from different detectors that are co-located. Therefore, future research directions may focus on using WILD and WPD data from locations where they co-exist, validating the WILD trending model identified in this phase, and re-developing combined WILD and WPD regression models.

References

Lagnebäck, R. Evaluation of wayside condition monitoring technologies for condition-based maintenance of railway vehicles. Licentiate Thesis, Lulea University of Technology, Department of Civil, Mining and Environmental Engineering Division of Operation and Maintenance Engineering, Lulea, Sweden, 2007.

Ouyang, Y., X. Li, C. P. L. Barkan, A. Kawprasert, and Y.-C. Lai. Optimal Locations of Railroad Wayside Defect Detection Installations. Computer-Aided Civil and Infrastructure Engineering 24 1–11, 2009.

Wayside Defect Detector Data Mining to Predict Potential WILD Train Stops

Leila Hajibabai[a], M. Rapik Saat[a], Yanfeng Ouyang[a], Christopher P.L. Barkan[a], Zengyi Yang[a], Kim Bowling[b], Kamalesh Somani[b], Don Lauro[b], and Xiaopeng Li[a]

[a] University of Illinois at Urbana-Champaign

[b] CSX Transportation

Outline

- Problem statement
- Data mining framework for condition monitoring
- Exploratory data analysis
- Regression approach
- Data analysis results and performance measures
- Conclusion

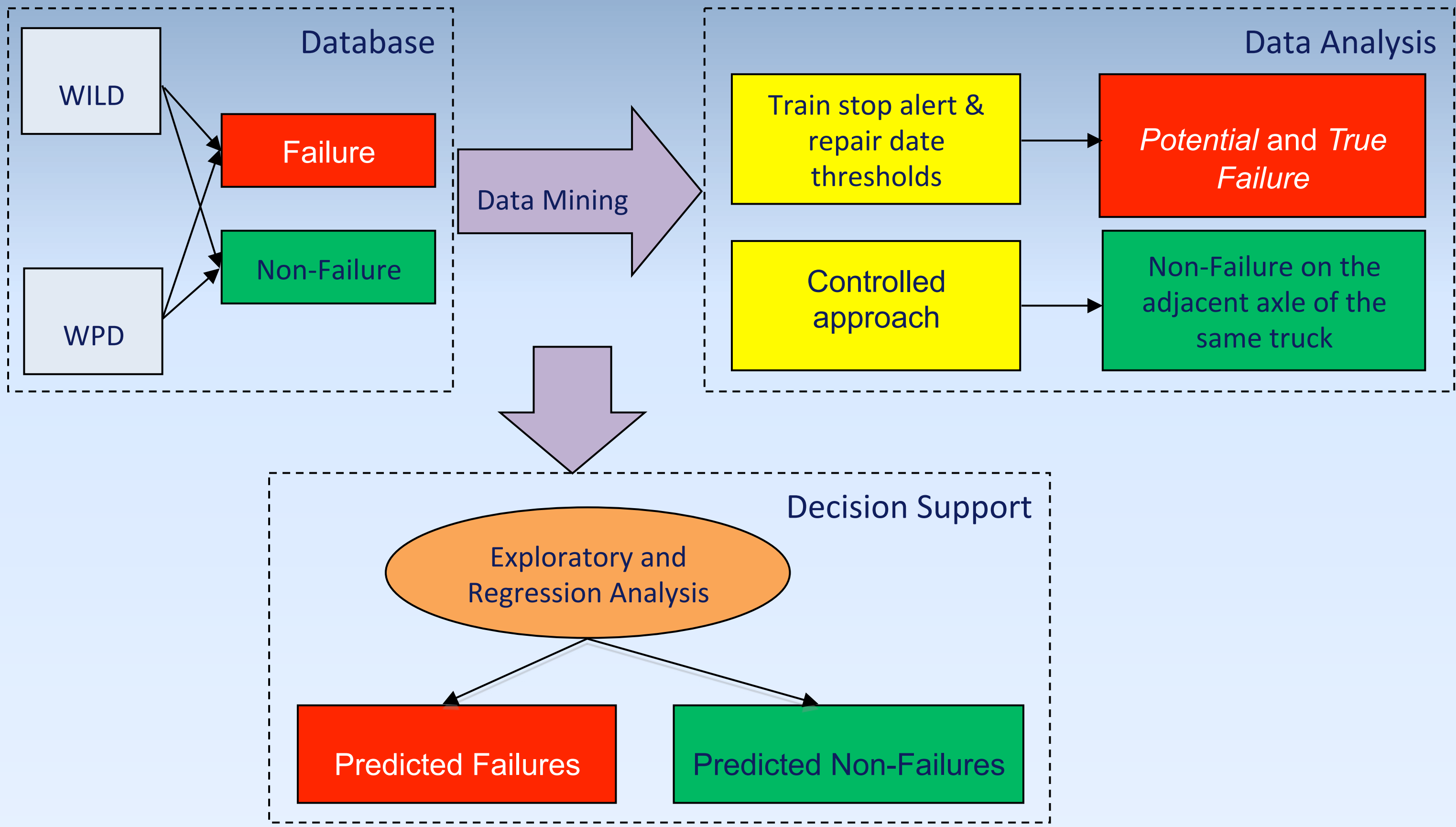
Problem Statement

- Advanced wayside detector technologies can be used to:
  - Monitor the condition of railcar components
  - Notify railroads of probable failures to equipment and infrastructure in advance
  - Predict alert of imminent mechanical-caused service failures
- More effective use of wayside mechanical inspection data can lead to improved safety, service reliability and economics through more effective preventive maintenance practices
- Objectives:
  - Recognize patterns that ultimately can be used to generate new, reliable, advanced rules to predict the occurrence of service disruptions
  - Develop the reliable and innovative rules to predict the failures (due to high impact wheels) and reduce related risks on railroads
    - Using statistical data-mining techniques on historical railcar health records from multiple Wayside Defect Detector (WDD) systems

Problem Statement

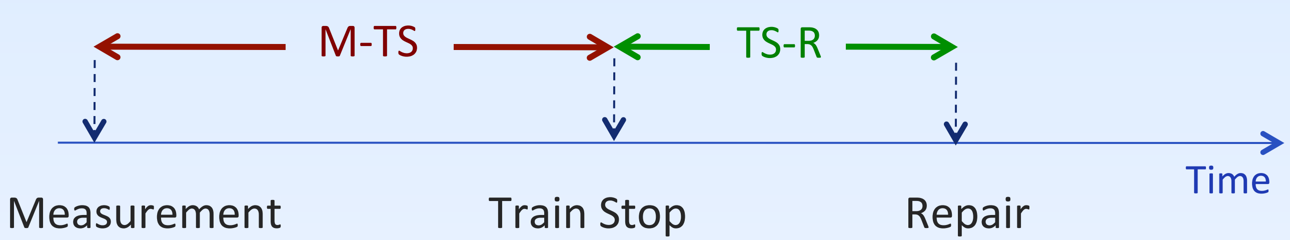
- Data from Wheel Impact Load Detector (WILD) and Wheel Profile Detector (WPD) are analyzed:
  - Comparing historical measurements for failed and non-failed wheels on the same truck
  - Exploratory data analysis to identify the most critical measurements from each detector by comparing the distributions of several measurements from failed wheels to the ones from non-failed wheels
  - Logistic regression approach to predict the probability of potential high impact wheel train stops
- Value proposition of individual and multiple WDDs used in combinations are evaluated:
  - WILD only
  - WPD only
  - WILD and WPD

Data Mining Framework for Condition Monitoring to Predict Wheel Failures



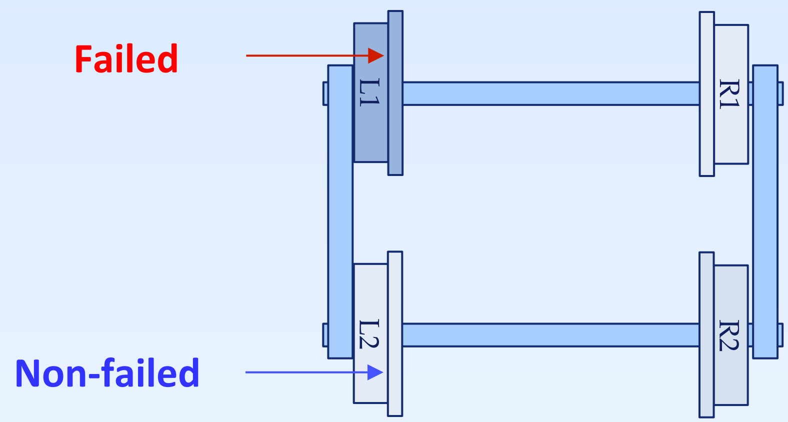
Data Analysis Approach: Time Span Method to Predict Wheel Failures

- Wheels in the failure category are further analyzed according to the time span between:
  - The most recent detector measurement and the train stop alert date (M-TS)
  - Each high impact train stop and repair dates (TS-R)
- Certain thresholds for M-TS and TS-R are defined and tested to determine the railcar wheel *Potential* vs. *True Failures*
- Data availability is an important issue to determine the appropriate thresholds



Data Analysis Approach: Time Span Method to Predict Wheel Failures

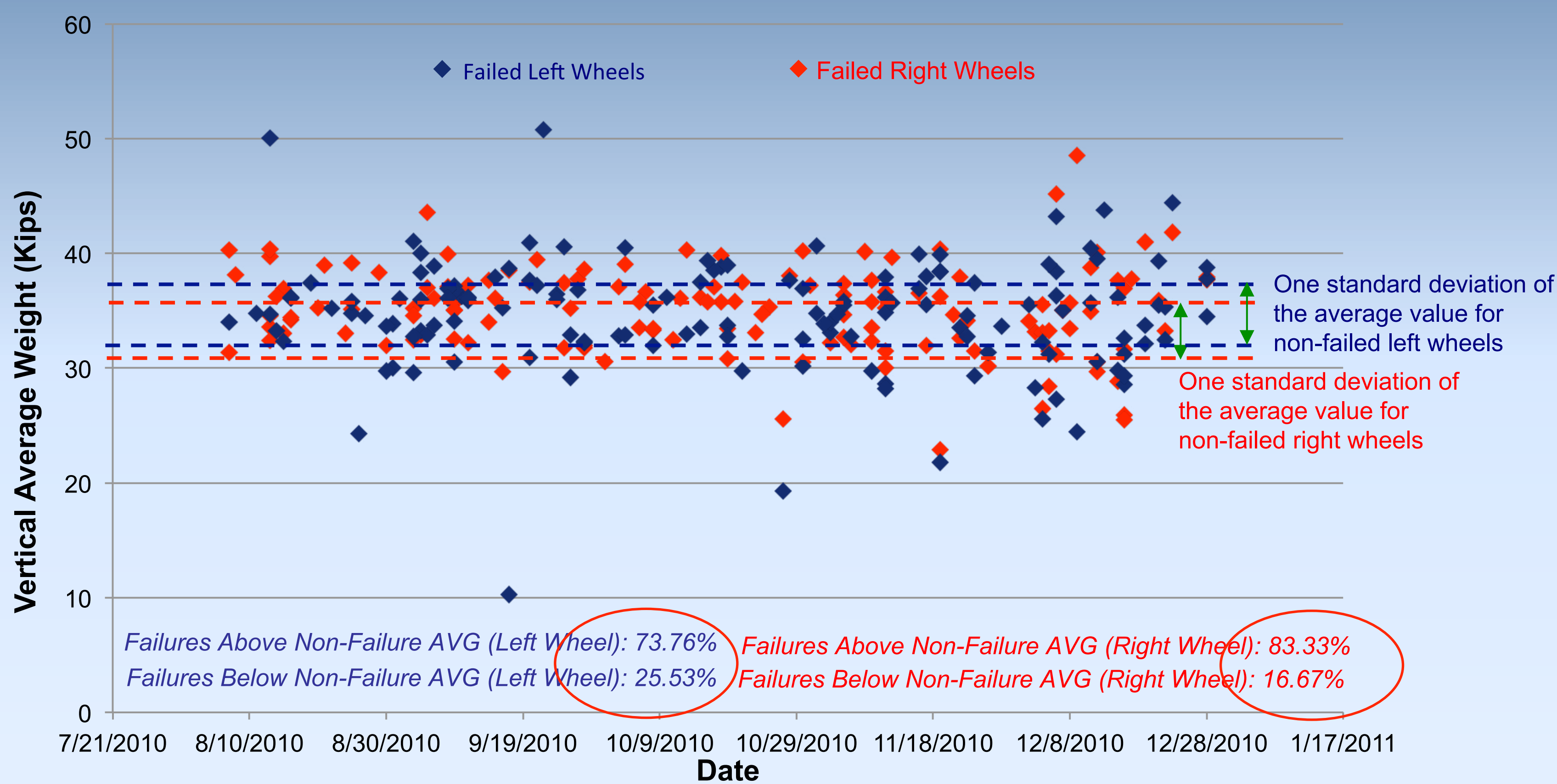
- Compared historical measurements for failed and non-failed wheels on the same truck to reduce uncertainty and variance
- Failed wheel (train stop due to high wheel impact)
  - *Potential Failure*: Reading date is within 30 days range before train stop date
  - *True Failure*: Wheel repair date is within 30 days range after train stop date
- Non-failed wheel
  - On the adjacent axle to the failed wheel
  - No train stop nor wheel repair



Exploratory Data Analysis to Identify Critical Variables

- After determining the failures and non-failures, a set of data attributes from each detector is selected for the exploratory analysis
  - WILD measurements: vertical average weight, vertical peak force, lateral average force, lateral peak force, and dynamic ratio
  - WPD measurements: flange height, flange thickness, rim thickness, tread hollow, wheel diameter, groove tread, vertical flange (flange angle), tread built-up, and wheel angle
- Based on each detector database, the time distribution of failures of each variable was compared to the average of non-failures of the same variable
- The percentage of the failures above (or below) the non-failures average is determined
  - If a large percentage of failed-wheel measurements were higher (or lower) than the average measurements for non-failed-wheels, the variable would be considered as a critical variable

Sample Exploratory Data Analysis to Identify Critical Variables



Based on this result Vertical Average Weight is deemed critical because large percentage of failed-wheel measurements are higher than the average measurements for non-failed-wheel

Summary of Critical Variables Identified

WILD:

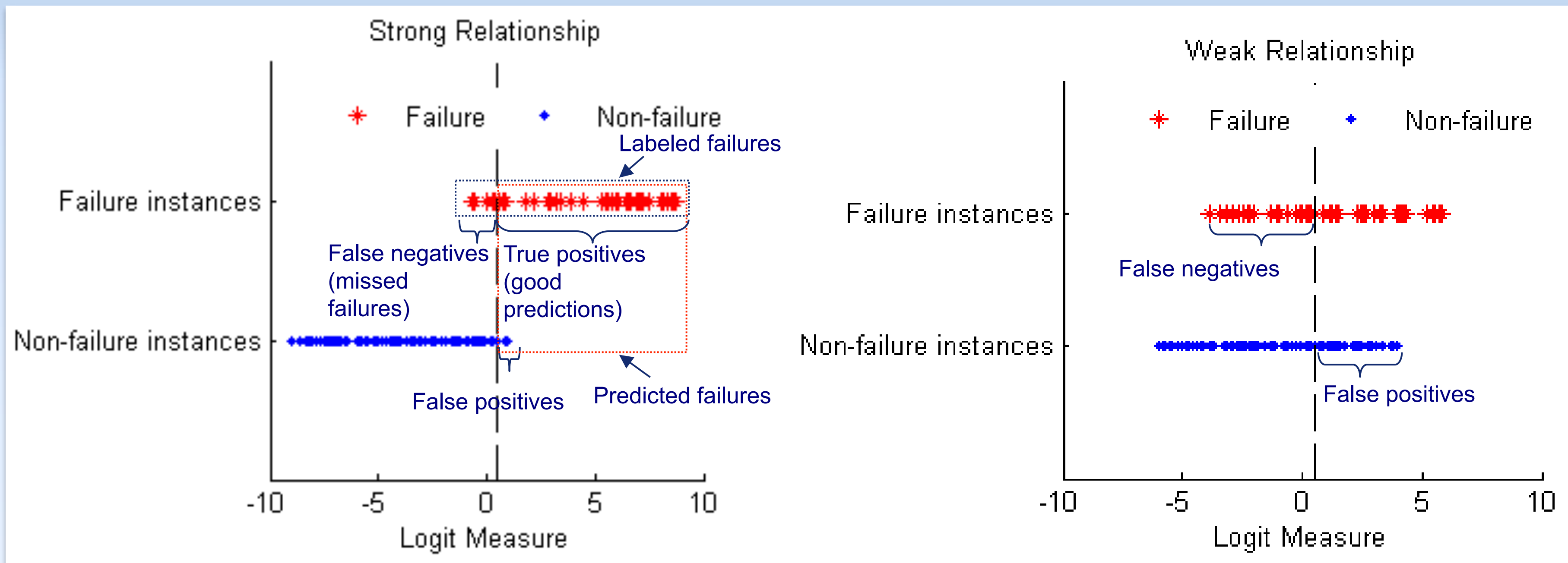
- Vertical Average Weight
- Vertical Peak Force
- Lateral Average Force
- Lateral Peak Force
- Dynamic Ratio

WPD:

- Vertical Flange (Flange Angle)
- Rim Thickness
- Tread Built-up
- Wheel Angle

Logistic Regression Approach

- Logistic regression setting
  - Detector readings $x_1, x_2, \ldots, x_n$ for a generic instance
  - Regression coefficients: $b_0, b_1, \ldots, b_n$
  - Logit measure $z = b_0 + b_{1x1} x_1 + b_{2x2} x_2 + \ldots + b_{nxn} x_n$



Regression Model Alternatives

- *Potential Failure* (85% of the data for model construction and 15% for test)
  - WPD critical variables (failure and non-failure records)
  - WILD critical variables (failure and non-failure records)
  - Combined WPD and WILD critical variables
- Sparse data to perform *True Failure* (verified repair) analysis
- Used unit-less values for all critical variables in regression models:

(measurement – non-failure mean)/non-failure mean × 100

Regression Model: Data Structure for Data Analysis based on WILD Potential Failure

(Absolute Value for Lateral Forces)

- Logistic Regression (if Dynamic Ratio is ignored):

$$P(z) = \frac{1}{1+e^{-z}}$$

$$z = 0.4002 - 0.0211x_1 + 0.0261x_2 - 0.0031x_3 - 0.6290x_4$$

- The critical variables in the model (0.1 significance):
  - Vertical Average Force(0.001), Vertical Peak Force(0.001), and Lateral Peak Force(0.1), respectively
  - Loading condition(0.01) has also been considered in this model
  - Significance Test of Logistic Regression:

    H_0: coefficients=0; H_1: O.W.

    The test is rejected, therefore the regression model between the dependent and independent variables is significant with p-value=0.002

Test of Logistic Regression using the 15% Unseen WILD Data: without Dynamic Ratio

(Absolute Value for Lateral Forces)

Failure Probability	>=0.2	>=0.3	>=0.4	>=0.5	>=0.6	>=0.7	>=0.8	>=0.9
True Failure	101	95	94	91	85	81	76	59
False Positive	33	31	15	5	3	3	2	2
False Negative	0	6	7	10	16	20	25	42

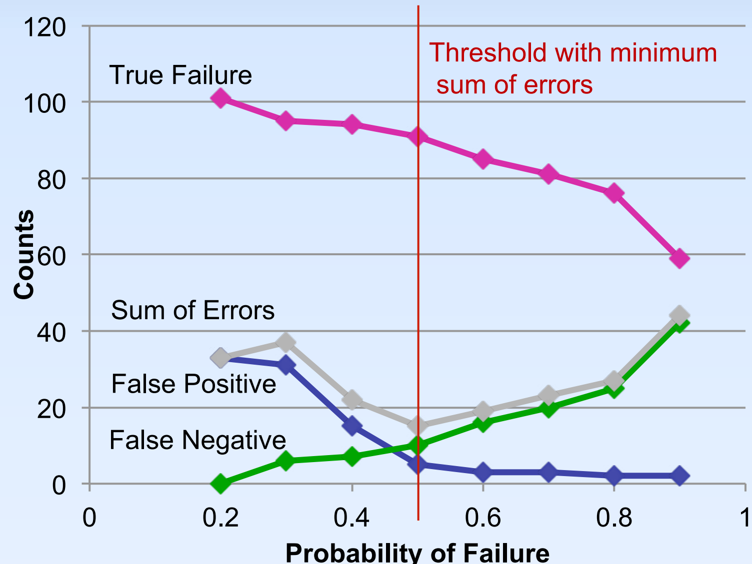
Confusion Matrix for Data Mining Model Evaluation (P>=50%):

Count:

		Prediction	
		Failure	Non-Failure
Actual	Failure	91	10
	Non-Failure	5	28

Percentage:

		Prediction	
		Failure	Non-Failure
Actual	Failure	90%	10%
	Non-Failure	15%	85%



Probability of Type II Error

Probability of Type I Error

Summary of Regression Model Performance Using the 15% Unseen Data

Promising results

Model	True Failure	Type I Error	Type II Error
WILD-w/out Dyn. Ratio	90%	15%	10%
WILD-w/ Dyn. Ratio	89%	15%	11%
WPD	40%	0%	60%
WPD & WILD-w/out Dyn. Ratio	20%	0%	80%
WPD & WILD-w/ Dyn. Ratio	60%	0%	40%

Conclusion

- We developed a methodical, controlled approach to analyze data by comparing failed and non-failed cases on the same truck
  - Reduce variance to make it easier to identify critical variables and develop good prediction models
- We found promising results with about 90% efficiency to predict high wheel impact train stops in the next 30 days based on the most recent WILD measurements
- Detailed analysis of individual data sources enable further understanding of information availability and limitations
  - Increase value proposition by identifying potential new rules to avoid service failures and derailments

Conclusion

- The lack of symmetry between the total number of WILD and WPD records prohibit development of a reliable combined model
- Difficulty in analyzing combined data from different sources/detectors reflects the need to focus on data from different detectors located at the same location
  - Could be helpful to guide future plan to install which detectors at super sites

QUESTIONS?

Appendix

Regression Model: Data Structure for Data Analysis based on WILD Potential Failure

(Absolute Value for Lateral Forces)

- Logistic Regression (if Dynamic Ratio is considered):

$$P(z) = \frac{1}{1+e^{-z}}$$

$$z = 0.4103 + 0.0256x_1 - 0.6036x_2$$

- The critical variables in the model (0.01 significance):
  - Dynamic Ratio(0.001)
  - Loading condition(0.01) has also been considered in this model
  - Significance Test of Logistic Regression:

    H_0: coefficients=0; H_1: O.W.

    The test is rejected, therefore the regression model between the dependent and independent variables is significant with p-value=0.013

Test of Logistic Regression using the 15% Unseen WILD Data: with Dynamic Ratio
(Absolute Value for Lateral Forces)

Failure Probability	>=0.2	>=0.3	>=0.4	>=0.5	>=0.6	>=0.7	>=0.8	>=0.9
True Failure	101	96	94	90	84	81	74	63
False Positive	33	33	13	5	3	3	3	2
False Negative	0	5	7	11	17	20	27	38

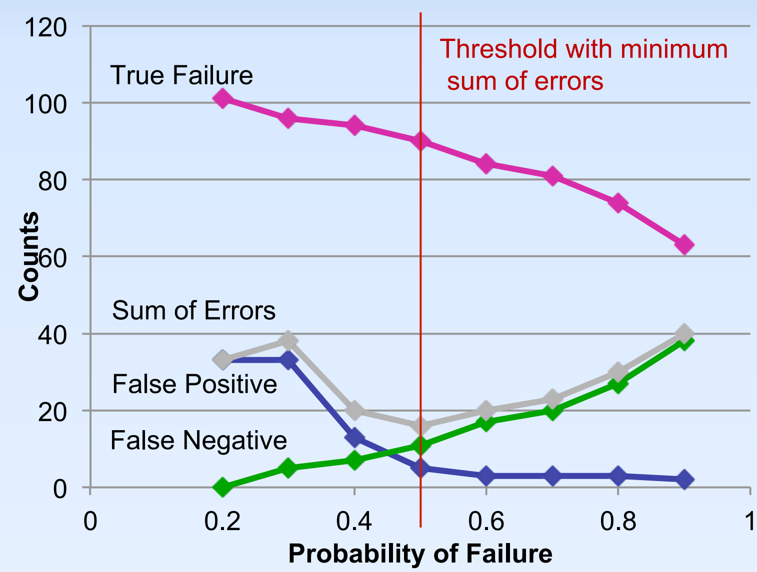
Confusion Matrix for Data Mining Model Evaluation (P>=50%):

Count:

		Prediction	
		Failure	Non-Failure
Actual	Failure	90	11
	Non-Failure	5	28

Percentage:

		Prediction	
		Failure	Non-Failure
Actual	Failure	89%	11%
	Non-Failure	15%	85%



Probability of Type II Error

Probability of Type I Error

Regression Model: Data Structure for Data Analysis based on WPD Potential Failure

- Logistic Regression:

$$P(z) = \frac{1}{1+e^{-z}}$$

$$z = -1.1029 - 0.0416x_1 + 0.0272x_2$$

- The critical variables in the model (0.1 significance):
  - Rim Thickness(0.05) and Flange Angle(0.1), respectively
  - Significance Test of Logistic Regression:

    H_0: coefficients=0; H_1: O.W.

    The test is rejected, therefore the regression model between the dependent and independent variables is significant with p-value=0.031

Test of Logistic Regression using the 15% Unseen WPD Data

Failure Probability	>=0.2	>=0.3	>=0.4	>=0.5	>=0.6	>=0.7	>=0.8	>=0.9
True Failure	3	3	3	2	1	0	0	0
False Positive	8	5	2	0	0	0	0	0
False Negative	2	2	2	3	4	5	5	5

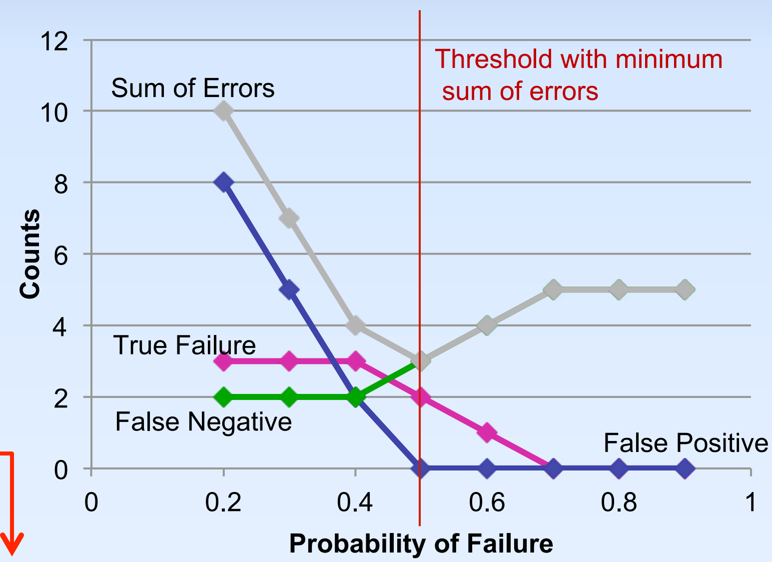
Confusion Matrix for Data Mining Model Evaluation (P>=50%):

Count:

		Prediction	
Actual		Failure	Non-Failure
	Failure	2	3
	Non-Failure	0	10

Percentage:

		Prediction	
Actual		Failure	Non-Failure
	Failure	40%	60%
	Non-Failure	0%	100%



Probability of Type II Error

Probability of Type I Error

Regression Model: Data Structure for Data Analysis based on WPD and WILD Potential Failure

(Absolute Value for Lateral Forces)

- Logistic Regression (if Dynamic Ratio is ignored):

$$P(z) = \frac{1}{1+e^{-z}}$$

$$z = -2.4402 + 0.0300 x_1$$

- The critical variables in the model (0.001 significance):
  - Vertical Peak Force (0.001)
  - Significance Test of Logistic Regression:

    H_0: coefficients=0; H_1: O.W.

    The test is rejected, therefore the regression model between the dependent and independent variables is significant with p-value=6.283e-10

Test of Logistic Regression using the 15% Unseen WPD & WILD Data: without Dynamic Ratio

(Absolute Value for Lateral Forces)

Failure Probability	>=0.2	>=0.3	>=0.4	>=0.5	>=0.6	>=0.7	>=0.8	>=0.9
True Failure	1	1	1	1	1	1	1	1
False Positive	0	0	0	0	0	0	0	0
False Negative	4	4	4	4	4	4	4	4

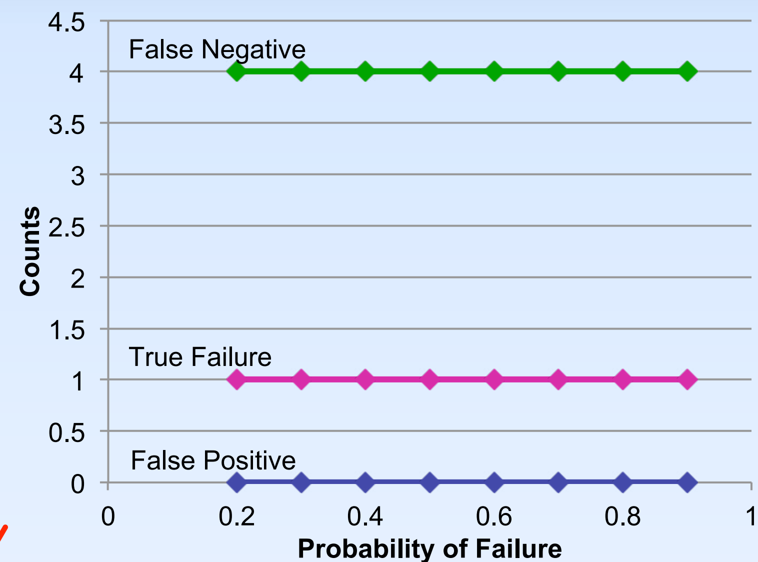
Confusion Matrix for Data Mining Model Evaluation:

Count:

		Prediction	
		Failure	Non-Failure
Actual	Failure	1	4
	Non-Failure	0	10

Percentage:

		Prediction	
		Failure	Non-Failure
Actual	Failure	20%	80%
	Non-Failure	0%	100%



Probability of Type II Error

Probability of Type I Error

Regression Model: Data Structure for Data Analysis based on WPD and WILD Potential Failure

(Absolute Value for Lateral Forces)

- Logistic Regression (if Dynamic Ratio is considered):

$$P(z) = \frac{1}{1+e^{-z}}$$

$$z = -2.3057 + 0.0317x_1 + 0.0107x_2$$

- The critical variables in the model (0.1 significance):
  - Dynamic Ratio (0.001), and Lateral Average Force(0.1), respectively
  - Significance Test of Logistic Regression:

    H_0: coefficients=0; H_1: O.W.

    The test is rejected, therefore the regression model between the dependent and independent variables is significant with p-value=1.305e-09

Test of Logistic Regression using the 15% Unseen WPD & WILD Data: with Dynamic Ratio

(Absolute Value for Lateral Forces)

Failure Probability	>=0.2	>=0.3	>=0.4	>=0.5	>=0.6	>=0.7	>=0.8	>=0.9
True Failure	3	3	1	1	1	1	1	1
False Positive	0	0	0	0	0	0	0	0
False Negative	2	2	4	4	4	4	4	4

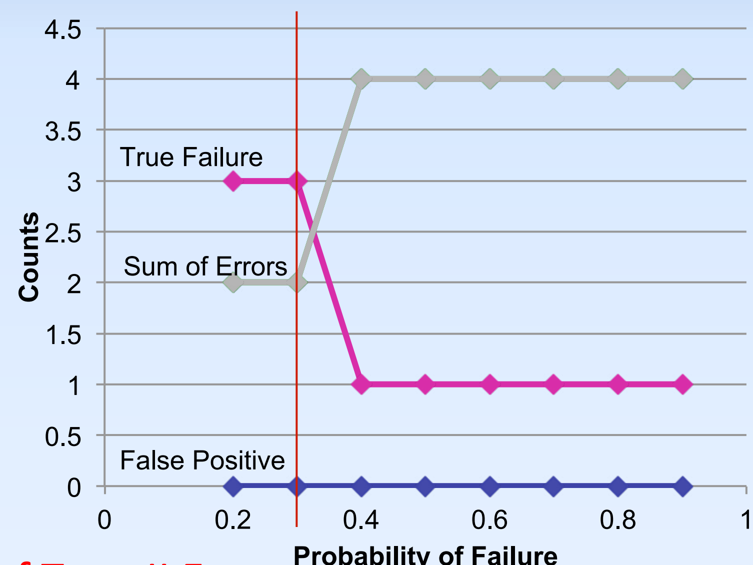
Confusion Matrix for Data Mining Model Evaluation (P>=30%):

Count:

		Prediction	
		Failure	Non-Failure
Actual	Failure	3	2
	Non-Failure	0	10

Percentage:

		Prediction	
		Failure	Non-Failure
Actual	Failure	60%	40%
	Non-Failure	0%	100%



Probability of Type II Error

Probability of Type I Error