A Quantitative Analysis of Factors Affecting Broken Rails

Darwin H. Schafer II & Christopher P.L. Barkan
May 9th, 2008
The William W. Hay Railroad Engineering Seminar Series
University of Illinois at Urbana-Champaign
Introduction to Broken Rails

• Broken rails are caused by inherent flaws (internal defects) in the rail and/or by fatigue on the surface of the rail.
  
  – A **broken rail derailment** is the occurrence of a train accident due to the rail failure
  
  – A **service failure** is a detected broken rail event without the occurrence of a train derailment.
    
    • Usually detected by the signal system
Probability of Broken Rail Events

Broken rails result from the growth of internal or surface defects:

- Internal rail defects typically start as microscopic imperfections during the manufacturing of the rail.
  - Defects grow slowly due to cyclic loading and unloading of the rail from traffic and thermal stresses
  - Under the “right” conditions they may grow rapidly leading to a service failure or broken rail derailment

- The growth of defects is linked to a number of variables:
  - Wheel loads, types of rail, traffic levels, maintenance activities, changes in track modulus, etc.
Consequences of Broken Rail Events

Detection of broken rails is vital:

- Broken rails are the leading causes of major train derailments in U.S. and accident-caused hazardous materials releases

- From 2003 through 2006 broken rails accounted for:
  - 335 mainline derailments for Class I railroads in the U.S.
  - Over $176 million of equipment & track damage
    - Approximately $525,000 per incident
  - 14 hazardous material release accidents

- Service failures also cause delay and track repair cost
  - This is especially true on high density lines
Importance of Broken Rail Prevention

Railroad Mainline Accidents by Cause (1996-2005)

- Joint Bar Defects
- Buckled Track
- Wide Gage
- Track Geometry (excl. Wide Gage)
- Grade Crossing Collisions
- Broken Rails or Welds
Presentation Outline

• Economic Impact
  – Cost associated with broken rail derailments and service failures
  – Cost associated with train delay
  – Cost of typical preventive measures

• Prediction of broken rails
  – Description of broken rail data
  – Development of new statistical and neural network prediction models
  – Presentation of practical model for broken rail prediction

• Conclusions & future work
Presentation Outline

• Economic Impact
  – Cost associated with broken rail derailments and service failures
  – Cost associated with train delay
  – Cost of typical preventive measures

• Prediction of broken rails
  – Description of broken rail data
  – Development of new statistical and neural network prediction models
  – Presentation of practical model for broken rail prediction

• Conclusions & future work
Economic Impact of Broken Rails

Costs of rail failures include (Cannon et al. 2003):

- **Cost of broken rail derailments**
  - Track & equipment damage
  - Clean-up costs, lading damage, loss of business, etc.

- **Remedial treatments**
  - Replacement or repair of rails and welds

- **Train delay**
  - Based on time out-of-service and line density

- **Track inspection**
  - Ultrasonic & geometric testing

- **Preventive methods**
  - Rail grinding, rail replacement, track surfacing
Cost of Broken Rail Derailments

Reportable accidents to the FRA:

• Accident cost must be greater than $7,700 (for 2006)
• Equipment damage reported to the FRA
  – Repair or replacement of on-track rail equipment.
    • Cars, locomotives, and maintenance equipment
  – Includes cost labor and materials needed.
• Track damage reported to the FRA
  – Repair or replacement of any track, signals, or track structures.
  – Takes into account material and labor costs associated with clearing the right-of-way and repairing the roadbed.
  – Includes contracted third-party labor fees.
# Cost of Broken Rail Derailments

FRA reportable broken rail accidents:

- **Class I US freight railroad mainline accidents**
  - 335 broken rail derailments from 2003 - 2006
  - Total cost of $176 million

- Average cost of broken rail derailment is $525,400

<table>
<thead>
<tr>
<th>FRA Code</th>
<th>Cause Description</th>
<th>Frequency</th>
<th>Total Cost</th>
<th>Cost Per Incident</th>
</tr>
</thead>
<tbody>
<tr>
<td>T201</td>
<td>Bolt hole crack or break</td>
<td>14</td>
<td>$12,854,596</td>
<td>$918,185</td>
</tr>
<tr>
<td>T202</td>
<td>Broken Base</td>
<td>26</td>
<td>5,804,332</td>
<td>223,244</td>
</tr>
<tr>
<td>T203</td>
<td>Broken Weld (plant)</td>
<td>3</td>
<td>1,026,794</td>
<td>342,265</td>
</tr>
<tr>
<td>T204</td>
<td>Broken Weld (field)</td>
<td>25</td>
<td>17,338,957</td>
<td>693,558</td>
</tr>
<tr>
<td>T207</td>
<td>Detail fracture from shelling or head check</td>
<td>82</td>
<td>52,792,131</td>
<td>643,806</td>
</tr>
<tr>
<td>T210</td>
<td>Head and web separation (outside joint bar)</td>
<td>23</td>
<td>5,380,144</td>
<td>233,919</td>
</tr>
<tr>
<td>T211</td>
<td>Head and web separation (within joint bar)</td>
<td>7</td>
<td>2,042,042</td>
<td>291,720</td>
</tr>
<tr>
<td>T212</td>
<td>Horizontal split head</td>
<td>5</td>
<td>1,967,657</td>
<td>393,531</td>
</tr>
<tr>
<td>T213</td>
<td>Joint bar broken (compromise)</td>
<td>6</td>
<td>3,111,204</td>
<td>518,534</td>
</tr>
<tr>
<td>T214</td>
<td>Joint bar broken (insulated)</td>
<td>9</td>
<td>8,152,304</td>
<td>905,812</td>
</tr>
<tr>
<td>T215</td>
<td>Joint bar broken (noninsulated)</td>
<td>15</td>
<td>14,225,856</td>
<td>948,390</td>
</tr>
<tr>
<td>T219</td>
<td>Rail defect with joint bar repair</td>
<td>1</td>
<td>664,622</td>
<td>664,622</td>
</tr>
<tr>
<td>T220</td>
<td>Transverse/compound fissure</td>
<td>90</td>
<td>42,315,036</td>
<td>470,167</td>
</tr>
<tr>
<td>T221</td>
<td>Vertical split head</td>
<td>29</td>
<td>8,333,190</td>
<td>287,351</td>
</tr>
</tbody>
</table>

| Total    | $176,008,865 | $525,400 |
Cost of Broken Rail Derailments

Costs not reported to the FRA:

• Train delay
  – Depends on severity of accident, access to site, if hazmat is involved or if near a metropolitan area
  – One railroad estimates that a “moderate to large” broken rail derailment has an average of 24 hours of track outage.
  • Typically broken rail derailments pile cars up in the location at the point of derailment and only damage track in that immediate area.
  – Cost can be approximated using developed train delay calculator

• Lading damage
  – Not reported to FRA, future research needed.

• Loss of business or customer support
  – Very difficult to quantify, future research needed.
Cost of Broken Rail Service Failures

Costs associated with broken rail service failures:

• Material and labor costs:
  – Railroad estimates range from $750 to $1,500 for material, labor, and mobilization costs to repair a service failure.
    • The cost of mobilization varies significantly based on availability of material and access to the site.

• Train delay costs:
  – Length of delay depends on accessibility to the site, repair crew availability, and if proper materials are readily available.
  – One railroad estimates that the average track-outage for a broken rail service failure is 3-4 hours from the time of notification.
  – Delay cost can be estimated using train delay calculator.
Cost of Train Delay

Breakdown of estimated train delay cost:

- Cost based on four categories (calculated from industry operating averages):
  1) Car cost: **$20.67 per train-hour**
     - Cost of delaying railroad owned cars that cannot be used.
     - Depends on % or railroad owned cars in train consist.
  2) Locomotive Cost: **$64.54 per train hour**
     - Cost from loss of the use of locomotives delayed.
  3) Fuel Cost: **$18.24 per train-hour**
     - Fuel used while train is waiting in idle.
     - Depends on locomotive type.
  4) Crew Cost: **$110.08 per train-hour**
     - Train crew labor cost during a delay (including fringe benefits).
Cost of Train Delay

Total train delay cost:

- Calculated total cost per train-hour of delay is $213.52.
  - Value calculated based on industry averages for Class I railroads.
  - Railroads in the past have used train delay cost values ranging from $200 to $300 per train-hour.
    - Varies based on type of train, geographic location, etc.
- Additional train delay cost sources that were not considered:
  - Crew replacement due to federal hours-of-service regulations.
  - Fuel consumed for acceleration after train stoppage.
  - Delay cost of privately owned cars.
  - Cost of delay for multiple trains.
Train Delay Cost Calculator

Train delay due to multiple trains:

• Will vary based on the density of the line, the number of mainline tracks, and the length of the delay.

• Assume that trains will arrive in constant time intervals
  – This assumption is less valid for low density lines.
  – Interval between train arrivals can be calculated based on the number of Annual Million Gross Tons (MGTs)
    • Average train for Class I freight RR’s is 6,312 gross tons.

\[
n = \frac{\text{Annual MGTs}}{\text{tons per train (millions)}} = \frac{ANMGT}{0.006312}
\]

\[
t = \frac{\text{hours per year}}{\text{trains per year}} = \frac{8,766}{n} = \frac{55.33}{ANMGT}
\]
Train Delay Cost Calculator

Train delay cost formula:

\[ C = T \times x + \sum_{n=1}^{m} (T - nt) \times x \]

where,

- \( C \) = total train delay cost for multiple trains
- \( T \) = total delay time for service interruption
- \( x \) = cost of delay per train-hour ($213.52)
- \( m \) = number of following trains delayed = \( T / t \) (rounded to the nearest integer)
- \( t \) = hours per train arrival = 55.33 / ANMGT

• Train delay scenarios:
  1) Single track service failure or derailment (100% delay)
  2) Multiple track broken rail derailment (100% delay)
  3) Multiple track service failure (50% or less of delay)
     - Adjustment factor needed based on the density of the line.
Cost of Preventive Measures

Typical practices for preventing broken rails:

• Rail inspection
  – Ultrasonic inspection to search for internal defects in rail.
  – Inspection to detect geometric defects of the track.

• Rail grinding
  – Grind rail and remove surface defects – extends life of rail.

• Rail surfacing
  – Maintains stable and properly aligned track.

• Rail renewal and replacement
  – Occurs due to both wear on the rail and crack growth.
  – Replaced at local level or by capital rail replacement projects.
Cost of Preventive Measures

<table>
<thead>
<tr>
<th>Broken Rail Prevention Technique</th>
<th>Annual Cost per Track Mile ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultrasonic and Geometric Track Inspection</td>
<td>900</td>
</tr>
<tr>
<td>Rail Grinding</td>
<td>1,900</td>
</tr>
<tr>
<td>Rail Surfacing</td>
<td>700</td>
</tr>
<tr>
<td>Rail Renewal and Replacement</td>
<td>2,500</td>
</tr>
<tr>
<td><strong>TOTAL INDIRECT COST</strong></td>
<td><strong>6,000</strong></td>
</tr>
</tbody>
</table>

- Cost estimated to be $6,000 per track mile maintained.  
  - Based on industry reported estimates and averages.

- The cost of preventive measure are considered to be indirect costs of broken rails:
  - These practices have other benefits unrelated to broken rail prevention.
    - Rail grinding and surfacing extends overall rail life
    - Rail replacement is completed for many reasons
Summary of Economic Impact

- Based on this analysis estimated costs for broken rails:
  - Broken rail derailment: $550,000 and higher per incident
  - Service failure: $2,500 and higher per incident
  - Indirect cost of preventive measures: $6,000 per track mile

- Estimated annual costs for all Class I railroads:

<table>
<thead>
<tr>
<th>Broken Rail Associated Cost</th>
<th>Cost ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derailment Damage</td>
<td>44,002,216</td>
</tr>
<tr>
<td>Derailment Delay Cost</td>
<td>3,006,793</td>
</tr>
<tr>
<td>Service Failure Repair Cost</td>
<td>19,141,747</td>
</tr>
<tr>
<td>Service Failure Delay Cost</td>
<td>16,768,170</td>
</tr>
<tr>
<td><strong>TOTAL DIRECT COST</strong></td>
<td><strong>82,918,926</strong></td>
</tr>
<tr>
<td>Track Inspection</td>
<td>128,185,200</td>
</tr>
<tr>
<td>Rail Grinding</td>
<td>270,613,200</td>
</tr>
<tr>
<td>Rail Surfacing</td>
<td>99,699,600</td>
</tr>
<tr>
<td>Rail Renewal and Replacement</td>
<td>356,070,000</td>
</tr>
<tr>
<td><strong>TOTAL INDIRECT COST</strong></td>
<td><strong>854,568,000</strong></td>
</tr>
</tbody>
</table>
Presentation Outline

• Economic Impact
  – Cost associated with broken rail derailments and service failures
  – Cost associated with train delay
  – Cost of typical preventive measures

• Prediction of broken rails
  – Description of broken rail data
  – Development of new statistical and neural network prediction models
  – Presentation of practical model for broken rail prediction

• Conclusions & future work
Objective: Develop a model to identify locations in the rail network that have a high likelihood of experiencing a broken rail event.

Classification Problem:

- Failure vs. non-failure – if a broken rail event occurred
- Input variables:
  - Rail characteristics (Rail age, type, weight)
  - Track characteristics (Curvature, superelevation, gradient)
  - Traffic levels (Annual MGTs, speed, car loads, wheel passes)
  - Infrastructure (Bridges, turnouts, culverts, grade crossings)
  - Rail testing results (Geometric and ultrasonic inspections)
Data Available for Study

Service failure data used in analysis:

• 25,370 unique track sections from the BNSF Railway
  – Each track section has length 0.01 miles (~53 feet).
  – 12,685 sections experienced a rail failure in four year period.
    • 2003 through 2006
  – 12,685 randomly selected sections without a broken rail event.

• Rail network data provided for each track segment
  – 28 input variables for the model:
    • Rail characteristics, track characteristics, traffic levels, infrastructure features, and rail defect testing results
Previously Developed Statistical Model

Previous statistical model developed by Dick (2001):

- Model was formulated using broken rail data for two-year time period on the BNSF network.
  - May 1998 to May 2000 service failure data
- Logistic regression used to predict broken rail locations.
  - Dick’s (2001) model stated a 87.4% level of accuracy
- Model was tested for our current service failure data
  - Found to be only 54.8% accurate for current data
  - Further examination of the data used in Dick’s (2001) study revealed that the data was missing service failure locations
Updated Statistical Classification Model

Logistic regression (Logit) model development:

• Logistic regression technique used again to develop classification model with updated service failure data.
  
  – Logit is a discrete-choice, statistical model that constructs an equation based on the input variables to predict failure.
  
    • Only variables significant to the model are incorporated.
    
    • Output of the model is a value between 0 and 1 indicating the probability of failure.
  
  – Different variable selection techniques:
  
    • Forward selection
    
    • Backward selection
    
    • Step-wise selection
  
  – Tested model against “unseen” data (validation sample)
Logit Modeling Techniques

Different modeling techniques were used:

• Transformation of input parameters:
  – Changing applicable inputs from continuous to discrete variables
    • Rail weight transformed into 13 distinct rail weights
  – Binary inputs already act as discrete variables
    • Rail type (bolted or welded), occurrence of a defect, bridge present

• Multiple term interaction:
  – Model allows for variables to have interdependency
    • For example, rail age and degree of curve may have a combined effect on service failure probability
    • Computational power and resources only allow for up to two-term interaction models using logistic regression
Statistical Model Results

<table>
<thead>
<tr>
<th>Logistic Regression Model</th>
<th>Number of Parameters</th>
<th>Accuracy of Initial</th>
<th>Accuracy for Testing Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Logit</td>
<td>23</td>
<td>66.3%</td>
<td>66.3%</td>
</tr>
<tr>
<td>Simple Logit w/ Transformation</td>
<td>34</td>
<td>68.5%</td>
<td>67.7%</td>
</tr>
<tr>
<td>Two Term Interaction Logit</td>
<td>145</td>
<td>71.1%</td>
<td>57.1%</td>
</tr>
<tr>
<td>Two Term Interaction Logit w/ Transformation</td>
<td>336</td>
<td>72.3%</td>
<td>67.3%</td>
</tr>
<tr>
<td>Practical Logit Model</td>
<td>8</td>
<td>64.7%</td>
<td>64.1%</td>
</tr>
</tbody>
</table>

• Results:
  – Logit models doing a fairly good job but required a lot of parameters
  – Two term interaction models also appeared to be doing well, but were “over-fitting” the data
  • Not accurate for “unseen” data
“Practical” Model

- Practical model was both simple and robust for data set
  - ~65% of locations can be identified correctly by only 8 inputs
  - This model is understandable and easy to implement

Accuracy of Classification

Practical Model

Number of Parameters

0%  10%  20%  30%  40%  50%  60%  70%

1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23
Artificial Neural Network Model

Artificial Neural Networks (ANN) for classification:

- Alternative to traditional logistic regression analysis:
- Advantages of Neural Networks:
  - Can detect complex non-linear relationships between input and output variables.
  - Can detect all possible interactions between predictor variables.
- Disadvantages of Neural Networks:
  - Greater computation time for large networks
  - No output value for probability – only classification
  - “Black Box” model
    - Limited ability to explain what is happening in the model and the possible causal relationships.
Artificial Neural Network

Inputs
- Rail age
- Percent Grade
- Number of Wheel Passes
- Bias

Hidden Neurons

Outputs
- Failure
- Non-Failure

Prediction
Neural Network Classification Model

Simple ANN Model:

• Built using same dataset as logistic regression model
• Learning method used was backpropagation
  – Most common learning method of ANNs
• Results:
  – 67.7% correct classification.
  – Optimal number of hidden neurons: 76
    • ANN restricts total number of neurons to not over-fit the model
  – Tested against unseen data and model was found to be robust:
    • 66.6% correct for validation sample.
Hybrid Logit/Neural Network Models

Hybrid classification models:

- Research in economic and medical applications have shown that hybrid models show a higher level of classification accuracy as compared to other techniques.
- Logistic regression analysis can be used as input information for neural network learning.
  - Two types of hybrid models:
    1) Pre-selection of input variables (Logit-ANN)
    2) Probability output as an additional input (PLogit-ANN)
- Main advantage of a hybrid model:
  - Information from Logit allows for fewer cases to be used on learning and therefore more cases are used on optimization.
Hybrid Logit/Neural Network Models

Hybrid models used for broken rail prediction:

1) Pre-selection of variables (Logit-ANN)
   - 23 variables selected as significant from step-wise Logit model
     • 67.5% correct classification.
     • Optimal number of hidden neurons: 71
     • 66.6% correct for validation sample.

2) Additional input variable of probability (PLogit-ANN)
   - Step-wise Logit model used to produce probability of failure.
     • 67.9% correct classification.
     • Optimal number of hidden neurons: 77
     • 66.4% correct for validation sample.
Summary of Developed Models

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Accuracy of Initial Classification Model</th>
<th>Accuracy for Testing Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Logit</td>
<td>66.3%</td>
<td>66.3%</td>
</tr>
<tr>
<td>Simple Logit w/ Transformation</td>
<td>68.5%</td>
<td>67.7%</td>
</tr>
<tr>
<td>Two Term Interaction Logit</td>
<td>71.1%</td>
<td>57.1%</td>
</tr>
<tr>
<td>Two Term Interaction Logit w/ Transformation</td>
<td>72.3%</td>
<td>67.3%</td>
</tr>
<tr>
<td>Practical Logit Model</td>
<td>64.7%</td>
<td>64.1%</td>
</tr>
<tr>
<td>ANN</td>
<td>67.7%</td>
<td>66.6%</td>
</tr>
<tr>
<td>Logit-ANN</td>
<td>67.5%</td>
<td>66.3%</td>
</tr>
<tr>
<td>Plogit-ANN</td>
<td>67.9%</td>
<td>66.4%</td>
</tr>
</tbody>
</table>

• Results:
  – Neural network (ANN) & hybrid neural networks performed as well as statistical models (Logit) for service failure prediction.
  – Two-term Logit models were not robust for validation data.
  – Practical Logit model is simple, understandable, and useable.
Practical Logistic Regression Model

Prospective statistical service failure model:

\[ P = \frac{e^U}{1+e^U} \]

where, \( P \) = probability that a service failure will occur

\[ U = -0.270 - 0.0454S - 1.35R - 0.0106A + 0.00899T \]
\[ + 0.0232L + 1.61I + 0.823G + 1.63B \]

where, 
- \( S \) = Rail weight (in pounds per yard)
- \( R \) = Rail type (1 if welded, 0 if bolted)
- \( A \) = Rail age (in years)
- \( T \) = Annual Traffic (in million gross tons)
- \( L \) = Weight of car (in tons)
- \( I \) = Presence of an ultrasonic defect (1 if present, else 0)
- \( G \) = Presence of a geometric defect (1 if present, else 0)
- \( B \) = Presence of a bridge within 200’ (1 if present, else 0)
Application of Practical Model

• The “practical” Logit model was transformed into a prospective prediction model.
  – Prospective model based on actual rate of service failures.
  – This model can be used to determine the probability of a broken rail or expected number of broken rails on a line.

• The models developed in this study can be used to assist railroads in prevention of broken rails.
  – Short-term maintenance planning:
    • Assist roadmasters on a local scale to identify locations that have a high probability of experiencing a broken rail.
  – Long-term maintenance planning:
    • Model can be used to determine the number of expected service failures on various lines to help determine maintenance priorities.
Presentation Outline

• Economic Impact
  – Cost associated with broken rail derailments and service failures
  – Cost associated with train delay
  – Cost of typical preventive measures

• Prediction of broken rails
  – Description of broken rail data
  – Development of new statistical and neural network prediction models
  – Presentation of practical model for broken rail prediction

• Conclusions & future work
Conclusions

- Economic impact of broken rail events is significant:
  - Annual direct cost of $83 million due to broken rail service failures and derailments on US Class I railroads.
  - Annual indirect costs exceed $850 million for broken rail prevention techniques.

- Statistical, artificial neural network, and hybrid modeling techniques were used to develop prediction models.
  - ANN and hybrid models did not show a significant improvement over the stand-alone statistical models.
  - “Practical” statistical model developed is simple & useable
    - Included only the top 8 predictor variables.
    - ~ 65% accuracy achieved as well as robust for unseen data.
Future Work

• Future research may be possible to understand the economic impact of “hard-to-quantify” values:
  – Lading loss in a derailment.
  – Loss of business from accidents and train delays.
  – Costs associated with delaying privately owned cars.

• Future work may also be possible to refine and improve our broken rail prediction models.
  – May be possible my attempting to understand the remaining unexplained model variance (~30% misclassified)
  – Evaluate additional possible predictive factors for broken rails:
    • Service failure rates, defect rates, temperature fluctuation, flat wheel incidence, inspection frequency, etc.
Acknowledgments

- Current & past University of Illinois students:
  - Daniel Peltier, Yung-Cheng (Rex) Lai, Rapik Saat, Mark Dingler, Billy Lau, John Zeman, Tyler Dick, & Avery Grimes

- Industry assistance:
  - BNSF Railway: Larry Milhon, Mike Mischke, Paul Gabler, Manal Bishr, Dennis Morgart, Mark Stehly, & Ed Chapman
  - CN Railway: David Ferryman
  - Union Pacific Railroad: Dale Mellor
  - CSX Transportation: Chuck Gullakson & Michael Hill
  - TTCI: Tom Guins
  - RSI-AAR Tan Car Safety Project: Todd Treichel

- Support from CN Fellowship and BNSF Research Grant