# Modeling the Useful Residual Life of Railroad Grease

Dr. Doug Timmer, Thania Martinez, Dr. Robert Jones, Dr. Constantine Tarawneh

University of Texas – Pan American

# Project Description

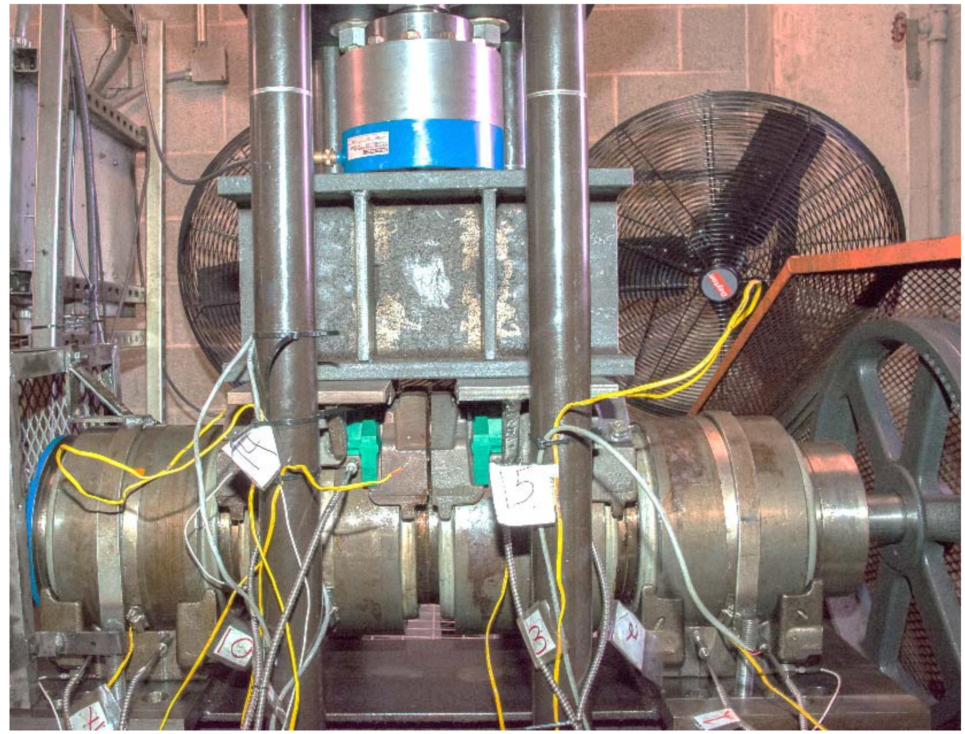
- The degradation of grease used to lubricate railroad bearings is believed to occur due two processes:
  - Mechanical processes occurring within the bearing,
  - Oxygen diffusion.
- Appropriate lubrication of the bearings is critical during railroad service operation.
- This study focuses on the development of empirical models that can accurately predict the residual useful life of railroad bearing grease.
- Employed Modeling Techniques:
  - Linear Regression Analysis
  - Regression Trees
  - Split Plots

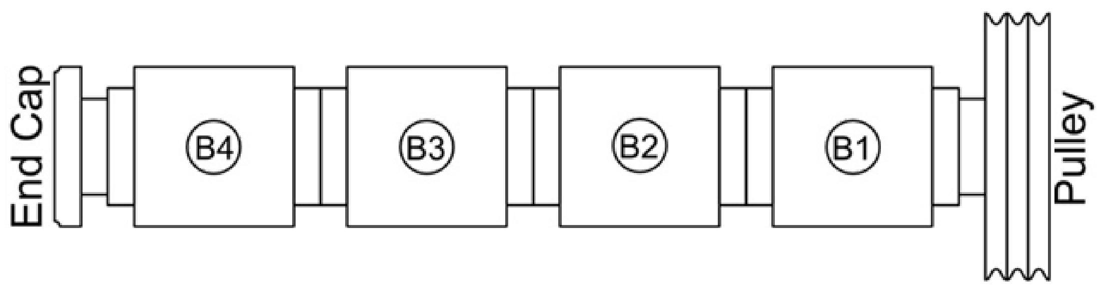
# Project Description (cont.)

- The data set used in the development of the model consists of more than 100 samples of grease taken from the railroad bearings which were observed in a laboratory setting.

# Laboratory Bearing Tester

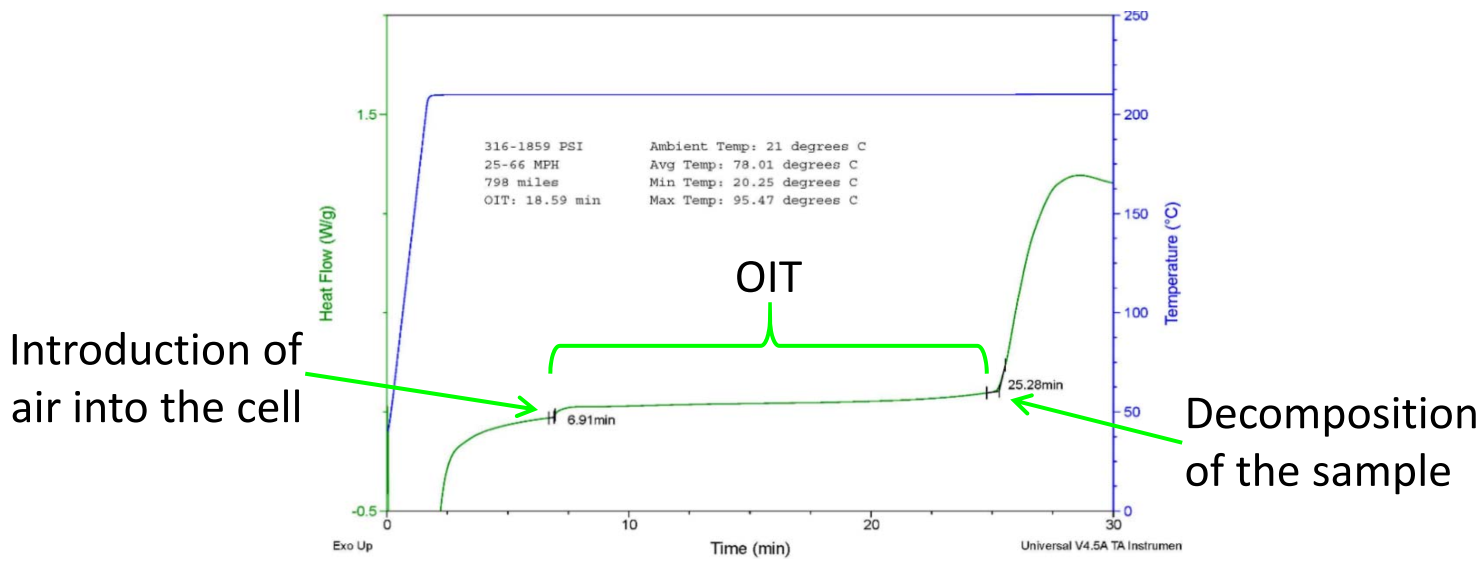
- Four bearings on the axle are subjected to the following experimental variables:
    - Load Conditions
    - Rotational Speed
    - Mileage
    - Temperature





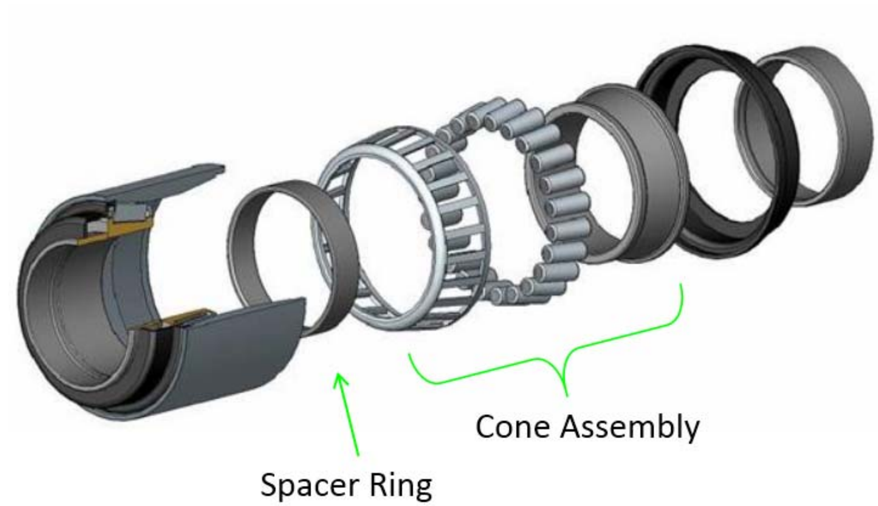
# Oxidation Induction Time

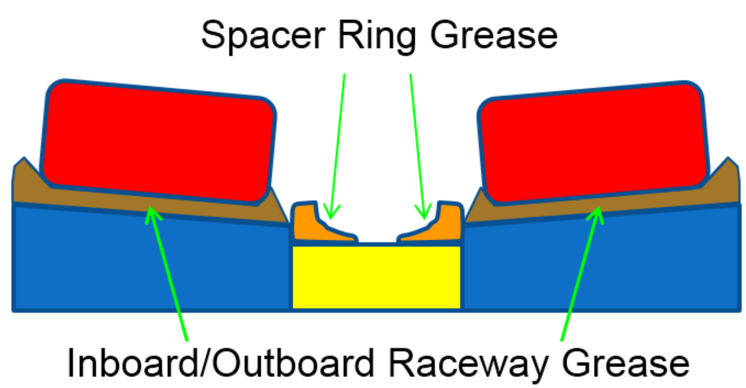
- **Oxidation Induction Time (OIT)** is a test performed in a Differential Scanning Calorimeter (DSC) which measures the level of thermal stabilizers in the material.
- The DSC produces a graph of heat flow vs time.
- The time elapsed between the introduction of air into the cell and the decomposition of the sample reveals the time to oxidation which is then recorded as OIT.



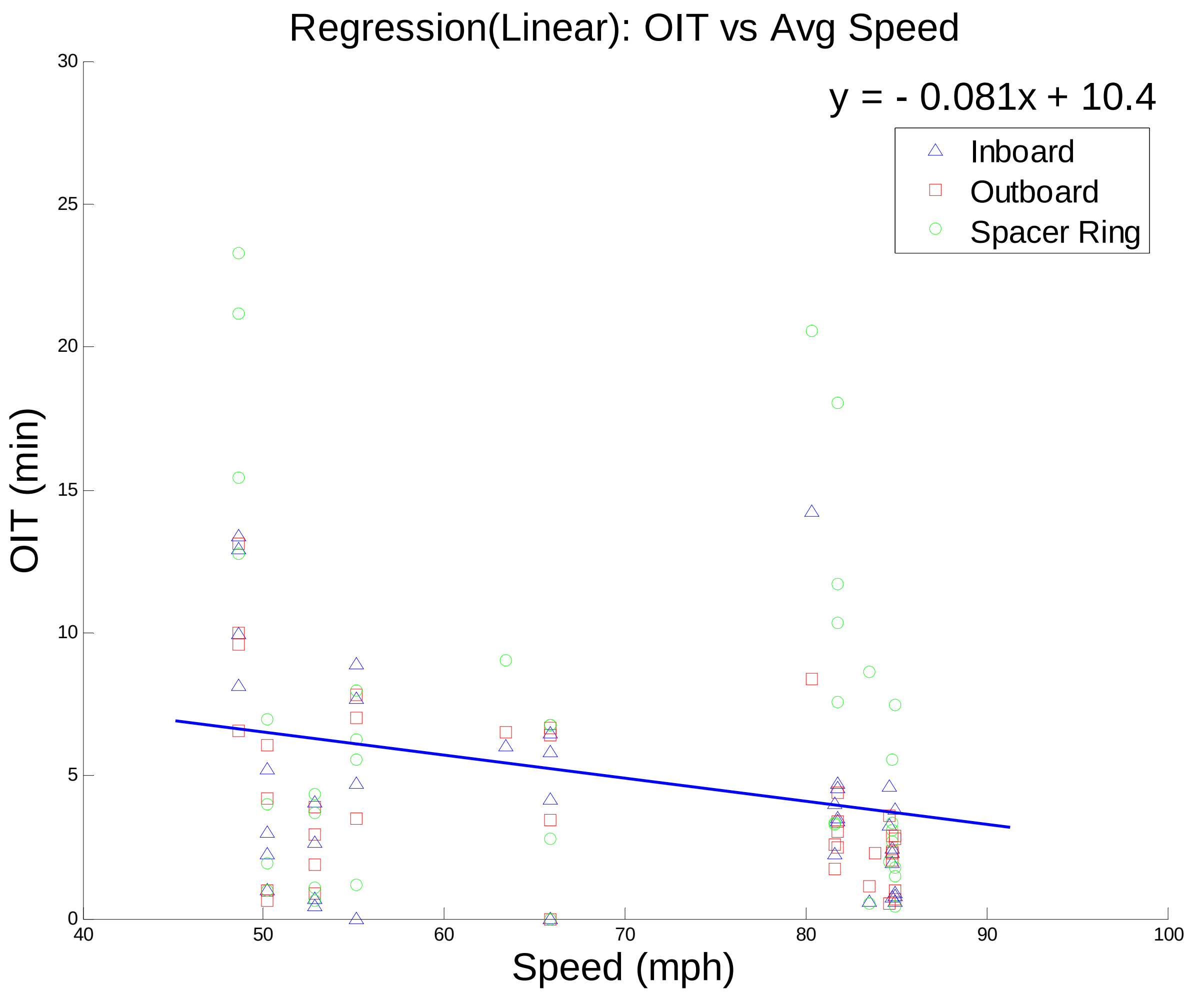
# Bearings

- Three samples come from each bearing, giving a total of twelve possible samples from each axle.
- Grease is sampled from the three critical locations of the bearing:
  - Inboard Cone Assembly Raceway
  - Outboard Cone Assembly Raceway
  - Spacer Ring Area



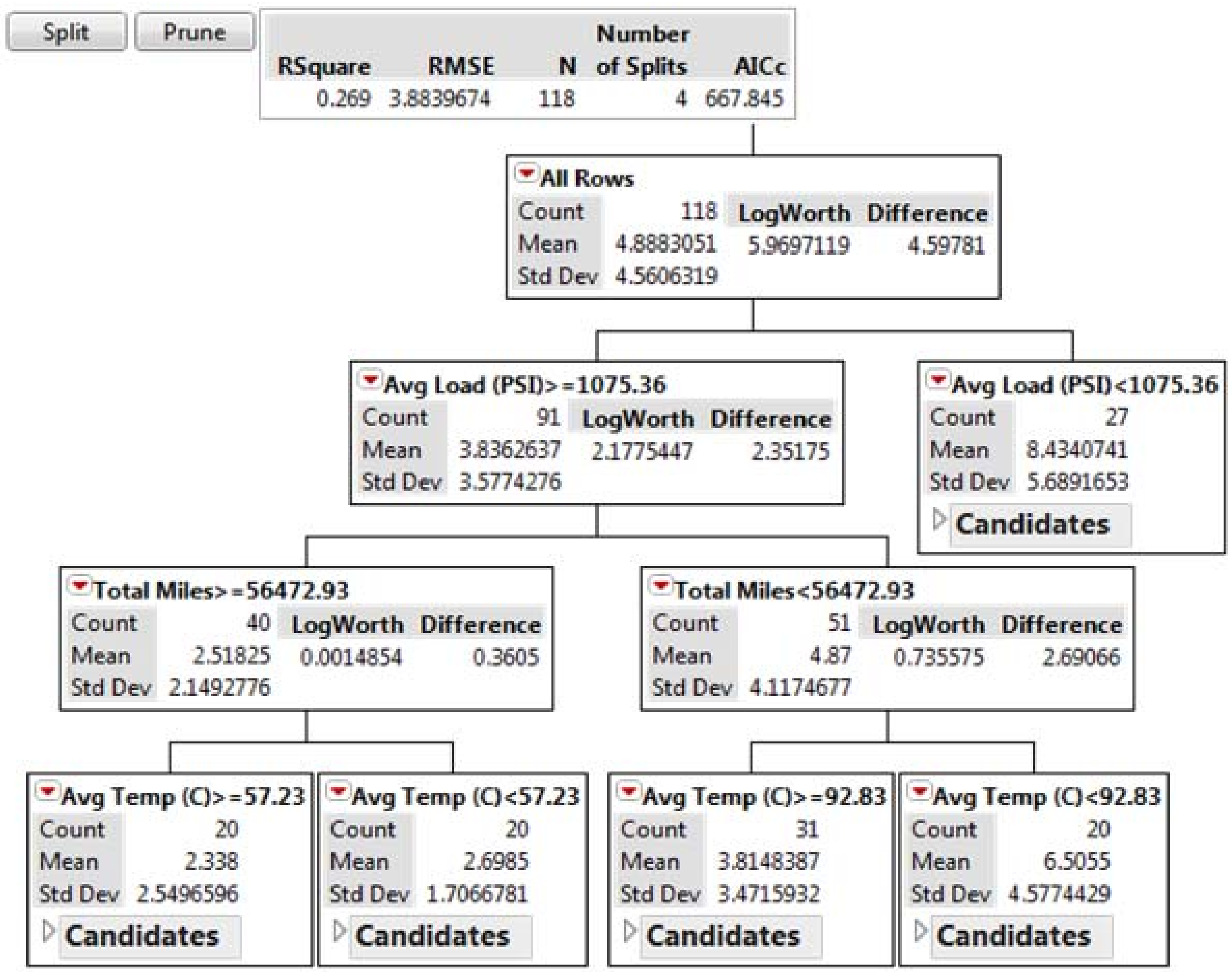
# Linear Regression Plot for OIT vs Speed



# Regression Tree

Min size split 20

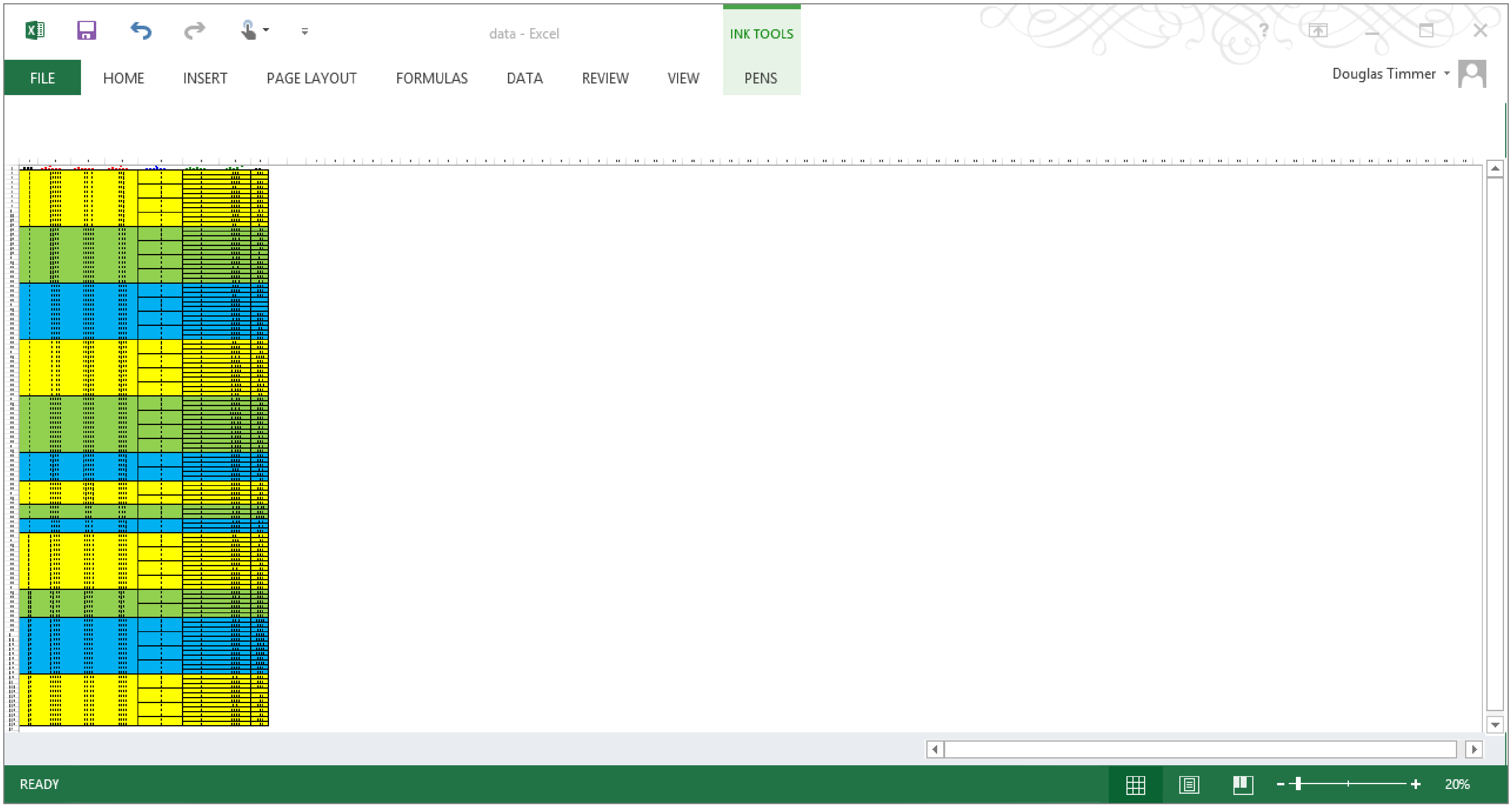
- Ad



# Experimental Design

- Split, Split-plot Design
  - Whole plot: axle-setup
  - Sub plot: each bearing on axle
  - Sub, sub plot: sample location within each bearing
- Single replicate
- Unbalanced design

# Unbalanced Data



# Parameter Estimation

- Restricted Maximum Likelihood (REML)
- Implemented in Matlab
- Degrees of Freedom are approximate due to unbalanced data

# Representation of Bearing Location in Regression Model

- The bearing location was recorded as a nominal value (1, 2, 3, 4)
- Modeled using three indicator variables

| | Dummy Variables | | |
|---|---|---|---|
| **Bearing** | X4 | X5 | X6 |
| **1** | 0 | 0 | 0 |
| **2** | 1 | 0 | 0 |
| **3** | 0 | 1 | 0 |
| **4** | 0 | 0 | 1 |

# Representation of Grease Location in Regression Model

- The grease location was recorded as a nominal value (1, 2, 3)
- Modeled using two indicator variables

| | Dummy Variables | |
|---|---|---|
| **Grease** | X7 | X8 |
| **1** | 0 | 0 |
| **2** | 1 | 0 |
| **3** | 0 | 1 |

# Initial Model

| | Term | Coef | se(Coef) | t-statistic | approx p-val | |
|---|---|---|---|---|---|---|
| wp terms (approx error df = 6) | Intercept | 1.0737 | 3.7641 | 0.2853 | 0.7850 | |
| | load | 1.5122 | 7.1931 | 0.2102 | 0.8405 | |
| | mileage | -13.2989 | 4.3152 | -3.0819 | 0.0216 | ** |
| | speed | 1.7393 | 4.5440 | 0.3828 | 0.7151 | |
| | load*mileage | -9.3852 | 7.4705 | -1.2563 | 0.2557 | |
| | load*speed | -4.6215 | 7.1840 | -0.6433 | 0.5438 | |
| | mileage*speed | 16.6564 | 7.5960 | 2.1928 | 0.0708 | * |
| sp terms (approx error df = 24) | x4 | 2.8713 | 0.9978 | 2.8776 | 0.0083 | ** |
| | x5 | 0.6385 | 1.0011 | 0.6378 | 0.5296 | |
| | x6 | 1.6466 | 0.9917 | 1.6604 | 0.1099 | |
| ssp terms (approx error df = 72) | x7 | 2.9221 | 0.6333 | 4.6141 | 0.0000 | ** |
| | x8 | -0.3409 | 0.5432 | -0.6276 | 0.5323 | |
| | temperature | -8.7969 | 3.2085 | -2.7417 | 0.0077 | ** |
| | load*temperature | -3.9283 | 5.1582 | -0.7616 | 0.4488 | |
| | mileage*temperature | -7.8488 | 3.6546 | -2.1476 | 0.0351 | ** |
| | speed*temperature | 2.7801 | 4.1258 | 0.6738 | 0.5026 | |

Analysis of coded variables

| | Obs | Approx DF | Approx Error DF |
|---|---|---|---|
| WP | 13 | 13 | 6 |
| SP | 40 | 27 | 24 |
| SSP | 118 | 78 | 72 |

| Variance Componer | Estimate |
|---|---|
| wp | 15.33 |
| sp | 1.87 |
| ssp | 5.59 |

| Variance Ratios | Estimates |
|---|---|
| eta1 | 2.742 |
| eta2 | 0.334 |

# Model 2

| | | Term | Coef | se(Coef) | t-statistic | approx p-val | |
|---|---|---|---|---|---|---|---|
| wp terms | (approx error df = 9) | Intercept | 0.6999 | 1.5747 | 0.4445 | 0.6672 | |
| | | mileage | -8.0003 | 3.1156 | -2.5678 | 0.0303 | ** |
| | | speed | 1.6769 | 1.8711 | 0.8962 | 0.3935 | |
| | | mileage*speed | 6.0180 | 3.7871 | 1.5891 | 0.1465 | |
| sp terms | (approx error df = 26) | x4 | 1.9142 | 0.7417 | 2.5808 | 0.0159 | ** |
| ssp terms | (approx error df = 75) | x7 | 2.9134 | 0.5294 | 5.5032 | 0.0000 | ** |
| | | temperature | -6.6029 | 2.4734 | -2.6696 | 0.0093 | ** |
| | | mileage*temperature | -6.2692 | 2.4603 | -2.5481 | 0.0129 | ** |

| Variance Components | Estimate |
|---|---|
| wp | 13.680 |
| sp | 2.052 |
| ssp | 5.497 |

| Variance Ratios | Estimates |
|---|---|
| eta1 | 2.4885 |
| eta2 | 0.3733 |

**Analysis of coded variables**

| | Obs | Approx DF | Approx Error DF |
|---|---|---|---|
| **WP** | 13 | 13 | 9 |
| **SP** | 40 | 27 | 26 |
| **SSP** | 118 | 78 | 75 |

# Model 3

| | | Term | Coef | se(Coef) | t-statistic | approx p-val | |
|---|---|---|---|---|---|---|---|
| wp terms | (approx error df = 11) | Intercept | 1.8964 | 1.3727 | 1.3815 | 0.1697 | |
| | | mileage | -3.9511 | 1.9252 | -2.0523 | 0.0423 | ** |
| sp terms | (approx error df = 26) | x4 | 1.8173 | 0.7385 | 2.4608 | 0.0208 | ** |
| ssp terms | (approx error df = 75) | x7 | 2.7981 | 0.5002 | 5.5940 | 0.0000 | ** |
| | | temperature | -5.0227 | 1.9157 | -2.6219 | 0.0106 | ** |
| | | mileage*temperature | -3.5885 | 3.0580 | -1.1735 | 0.2443 | |

| Variance Components | Estimate |
|---|---|
| wp | 14.714 |
| sp | 2.089 |
| ssp | 5.500 |

| Variance Ratios | Estimates |
|---|---|
| eta1 | 2.6753 |
| eta2 | 0.3799 |

**Analysis of coded variables**

| | Obs | Approx DF | ıprox Error DF |
|---|---|---|---|
| **WP** | 13 | 13 | 11 |
| **SP** | 40 | 27 | 26 |
| **SSP** | 118 | 78 | 75 |

# Final Model

| | | Term | Coef | se(Coef) | t-statistic | approx p-val | |
|---|---|---|---|---|---|---|---|
| wp terms | (approx error df = 11) | Intercept | 2.3872 | 1.1846 | 2.0152 | 0.0690 | |
| | | mileage | -3.8116 | 1.7281 | -2.2057 | 0.0496 | ** |
| sp terms | (approx error df = 26) | x4 | 1.7551 | 0.7521 | 2.3336 | 0.0276 | ** |
| ssp terms | (approx error df = 76) | x7 | 2.7443 | 0.4986 | 5.5040 | 0.0000 | ** |
| | | temperature | -3.7388 | 1.5485 | -2.4145 | 0.0160 | ** |

**Analysis of coded variables**

| | Obs | Approx DF | Approx Error DF |
|---|---|---|---|
| WP | 13 | 13 | 11 |
| SP | 40 | 27 | 26 |
| SSP | 118 | 78 | 76 |

| Variance Components | Estimate |
|---|---|
| wp | 11.311 |
| sp | 2.231 |
| ssp | 5.599 |

| Variance Ratios | Estimates |
|---|---|
| eta1 | 2.0202 |
| eta2 | 0.3984 |

# Final Model

- $\widehat{OIT} = 2.3872 - 3.8116 * mileage' + 1.7551 * x_4 + 2.7443 * x_7 - 3.7388 * temperature'$
- Where
  - $mileage' = \frac{mileage - 53396}{45687}$
  - $temperature' = \frac{temperature - 80.16}{32.71}$
  - $x_4$ is 1 if bearing 2 location, 0 for other bearing locations
  - $x_7$ is 1 if grease sampling location is the spacer ring and 0 for the inner or outer raceway

# Future Research

- Model Diagnostics
  - Residual analysis
  - R^2
  - VIF
- Model Refinement
  - Why is bearing 2 statistically different?
  - Is temperature a covariate (function of load, mileage and speed)?
  - Developing second response variable related to length of grease molecule
  - Alternative Model: neural network or ensemble of neural networks

# Acknowledgements

- University Transportation Center for Railway Safety (http://www.utrgv.edu/railwaysafety) for their support of this research
- The Matlab code was provided by Dr. Marcus Perry, Associate Professor of Statistics, Culverhouse College of Commerce, University of Alabama