TECHNOLOGY DIGEST

TD-18-032
©NOVEMBER 2018

TIMELY TECHNOLOGY TRANSFER

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

Neural Network Analysis for Rail Flaw Prediction

Abe Meddah, Matt Witte, and Ivan Aragona

Summary

Transportation Technology Center, Inc., (TTCI) investigated an application of the neural network (NN) technique to predict bolt hole crack (BHC), vertical split head (VSH), and crushed head (CH) rail flaws in the 136 RE rail. The proposed analysis for rail flaw prediction demonstrated that the NN model can accurately predict rail flaw types by relating a combination of flaw-related input variables to the flaw type output.

The data to develop the NN model was obtained from a Union Pacific (UP) database for track defects. The NN model was developed to capture the existing non-linear relationships between a set of 10 input variables pertaining to rail flaw development and rail defect type outputs. To build an accurate NN model, input variables were selected based on availability and the potential underlying physical relationship they have with the rail defect outputs. The input variables contain relevant information such as seasonality, defect length, track type, annual and accumulated annual gross tonnage (million gross tons or MGT), manufacturing date and track classes.

In this preliminary effort, only the 136 RE rail related flaws were investigated. The data subsets used contained a combination of data records pertaining to the three modeled defect types. The data subsets contained 2,000 data records, 50 percent of which were randomly selected for training the NN model. The other 50 percent were used to validate the model and assess performance accuracy. The validation data were fresh datasets that were not used in training the models. Training and validation data was grouped into three season groups: fall-spring combined, summer and winter. Models were developed for each defect type and group.

In the validation dataset, and for combined fall-spring seasons, the NN model was able to identify the BHC 95.9 percent, the VSH 83.7 percent, and the CH 59.1 percent. For the summer season, the BHC was successfully identified 92.3 percent, the VSH 84.2 percent, and the CH 82 percent. For the winter season, the NN model was accurate 94.9 percent for the BHC, 78 percent for the VSH, and only 22.9 percent for CH. Also investigated were the relative effects each input variable to the NN model had on the model performance. Each input variable was removed, in turn, from the training data; the NN model was retrained and then redeployed on the same validation dataset. Results show that the Defect Length variable is the most significant input variable to the model and without it the prediction accuracy is the worst for all investigated defect types. The input variable with the least useful predictive information was the variable month.

INTRODUCTION

With onboard and wayside railway technologies increasing in the railroad industry, railroads are increasingly collecting track condition data to monitor track health and optimally plan for maintenance needs. Track condition data, although abundant, may not be exploited in ways that help maintenance personnel identify the root causes behind some of common track defects. TTCI investigated the development of rail defects using neural network (NN) analysis. Those techniques are capable of learning about data features and recognizing patterns in large datasets. As with an earlier NN analysis applied to investigate the causes of vertical split rim (VSR) wheel failure,[1] NN could also be applied to predict rail defect types given a combination of input variables with solid underlying physical relationship to the rail defects.

In this preliminary investigation, TTCI engineers developed NN models to predict the following rail defect types:

- Bolt hole crack (BHC)
- Vertical split head (VSH)
- Crushed head (CH)

Figures 1, 2 and 3 show examples of typical BHC, VSH, and CH rail defect types, respectively.

Figure 1. Bolt hole crack (BHC) defect [2]

Figure 2. Vertical split head (VSH) defect [2]

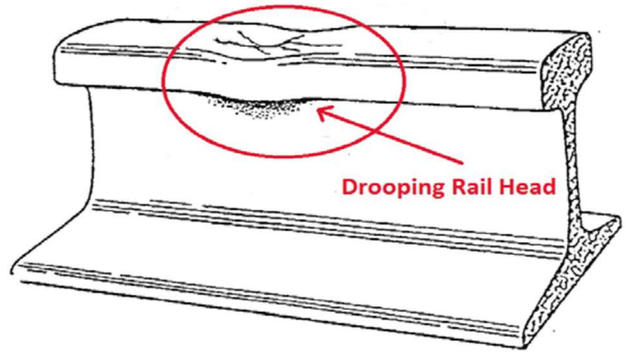

Figure 3. Crushed head (CH) defect [2]

DATABASE AND DATA STRUCTURE

The data used in this preliminary investigation was obtained from UP. Figure 4 shows the distribution of 10 common defect types collected between January 1997 and August 2015.

In assessing the merit of the NN technique, the amount of data must be adequate for the model development yet not necessitate significant pruning and processing. The defect types BHC, VSH, and CH address distinct defect types that represent a substantial portion of common defects while providing appropriate data sample size.

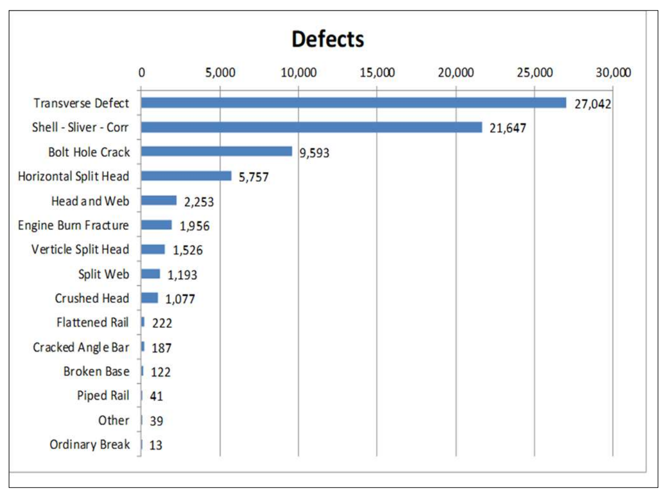

Figure 4. Rail defect types in UP database

Before processing, the data for BHC, VSH and CH was grouped by rail type and season of the year. Three distinct seasons: fall and spring combined, summer, and winter were used. Figure 5 shows an example flow chart of how the data was sorted and grouped for the BHC defect type. The data for VSH and CH was processed and grouped in the same way. In this investigation, only the track flaws associated with 136 RE rail type were considered.

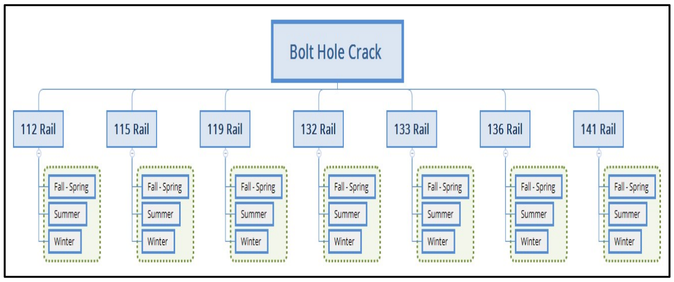

Figure 5. Data grouping by rail type and seasons

NN TRAINING AND VALIDATION DATA

The neural network models were developed to predict the following three rail defects: bolt hole crack; vertical split head; and crushed head.

A separate model was developed for each defect type. For instance, the model developed for BHC was trained to identify the BHC defects when it is deployed on validation data containing a combination of BHC, VSH and CH defect types. And the same was repeated with VSH and CH defects.

Training and validation subsets contained data pertaining to all three defect types. The NN models were trained to recognize patterns pertaining to each defect type. For each defect type, about 50 percent of the data records were randomly selected for training the model. For evaluating the model performance, the other 50 percent of data, independent of the training data, was used to validate the model. A total of 2,000 data records were used for training and validating the NN models.

Ten input variables to NN model were used:

- Month in which data was collected (season)
- Track type
- Left and right track sides
- Defect length
- Rail weight
- Manufacturing date
- Annual million gross tons (MGT)
- Accumulated MGT
- Route class
- Track class

Output target:

- Defect type (i.e., BHC, VSH, and CH)

It should be noted that the NN models were developed to predict one defect type output at a time. Figure 6 shows a graphical representation of neural network architecture consisting of an input layer comprising the input variables to the model, two hidden layers and an output layer for one predicted rail defect type.

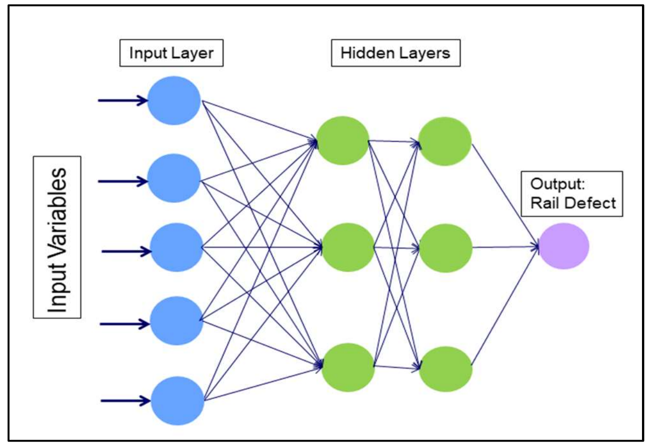

Figure 6. NN simplified architecture

MODEL PERFORMANCE RESULTS

Table 1 shows a summary of NN model performance for each defect on fresh datasets used solely for validation. The BHC model performed satisfactorily for all the seasons with greater than 90 percent confidence. The VSH model performance was greater than 80 percent confidence for fall-spring and summer seasons and 78 percent for the winter model. The CH summer model performed best with 82 percent confidence and the winter model worst with just 22.9 percent confidence. The CH winter model, in comparison, presented a poor predictive performance on the validation data. The CH winter model appears to have just memorized the training data patterns. It was incapable of adequately generalizing when deployed on the unseen validation data. It is possible that this performance could be the result of poor data quality in the winter season datasets. Also a possibility is that the CH Winter model input variables need to be optimized. Or, there may be a seasonality to the defect type. This is explored in the next section. The BHC model consistently performed better than the VSH and CH models. The BHC flaw may have been easier to discriminate, given the flaw sizes and orientation compared to the VSH and CH flaws.

Table 1. Result summary

	Fall-Spring Model Accuracy	Summer Model Accuracy	Winter Model Accuracy
Bolt hole crack (BHC)	95.9	92.3	94.9
Vertical split head (VSH)	83.7	84.2	78
Crushed head (CH)	59.1	82	22.9

RELATIVE EFFECTS OF MODEL INPUT VARIABLES

At the outset, all the input variables available in the database appeared, to varying degrees, to have underlying physical relationship with the rail defect types and how each defect developed. To assess the predictive information of each input variable to the model, each was investigated for the relative effects that each input variable has on the model performance.

Each input variable was removed, in turn, from the training data, the fall-spring NN model then retrained without it, and, once developed, redeployed on the same validation dataset initially used. If there is no change in the prediction accuracy when a variable is removed, then the variable is not significant. Figure 7 shows a summary of the NN model performance after individual input variables were removed in turn.

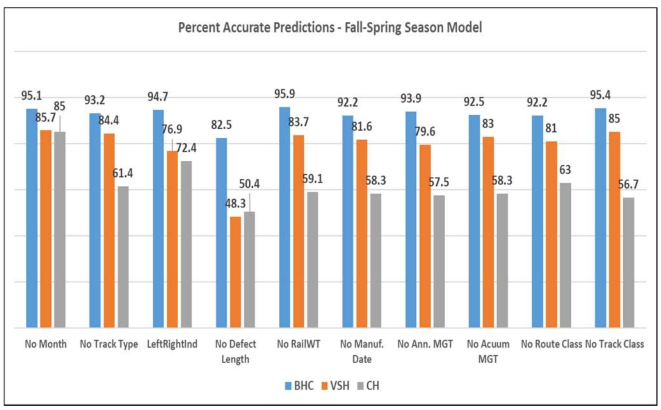

Figure 7. Input variable relative effects

The results indicate that the defect length input variable carries the most predictive information for all three defect types. Without the defect length variable in the training datasets, the NN model prediction accuracy was the worst for every defect type. This is mainly because a defect's size is a strong indicator of the defect type relative to its location and orientation within the rail section. After removing the Defect Length variable, the fall-spring model predictive performance was as follows:

- BHC: Confidence dropped from 95.9 to 82.5 percent
- VSH: 83.7 to 48.3 percent
- CH: 59.1 to 50.4 percent

The input variable with the least useful predictive information was the variable month. The model performance improved for the VSH and remained nearly unchanged for BHC. For CH defect, the NN model performance drastically improved when month was removed. The model accuracy went from 59.1 percent confidence up to 85 percent. It is possible that variable month may be redundant with seasonality and its presence in training data was patternless noise that did not contribute to learning.

Neither Rail MGT nor Annual MGT variables were significant to the model performance, possibly because of a change in traffic over the years. Significance was about the same as for Manufacturing Date or Track Type. Variable Rail Weight did not contribute at all since the investigation was solely conducted for rail type 136 RE. Removing that variable from the training set did not, expectedly, have any effects on the model performance. The individual effects of the other variables, when removed in turn, were not as significant as the variable Defect Length.

CONCLUSION

The analysis showed the rail Defect Length was the most significant input variable to NN model, and without it, the model performed the worst for the three defect types analyzed. Variable month was the least significant to the model performance. Winter CH model performed better with month variable excluded as input.

WAY FORWARD

As this investigation shows, NN pattern recognition capabilities can be successfully deployed for rail flaw prediction. For better understanding of rail flaw presence and formation, other pertinent rail flaw-related condition data, such as fastening systems, crosstie types and track geometry, also should be explored for rail flaw predictions. Using such historic condition data with the NN technique could prove useful in identifying beforehand rail flaw types that are likely to develop so diagnosis warning can be aptly provided.

References

1. Meddah, A, Stone, D and Cummings, S.. "Neural Network Analysis of the Causes of Vertical Split Rim Wheel Rim Failure," *Technology Digest* TD-13-014, June 2013, AAR/TTCI, Pueblo, CO.
2. Federal Railroad Administration, Office of Railroad Safety, *Track Inspector Rail Defect Reference Manual.* Revision 2, July 2015, U.S. Dept. of Transportation.