Center for Multidisciplinary Research Excellence in Cyber-Physical Infrastructure Systems (MECIS)

The University of Texas Rio Grande Valley
Center for Multidisciplinary Research Excellence in Cyber-Physical Infrastructure Systems (MECIS)

Autonomous Train Maintenance: Machine Learning Model for Bearing Mileage Prediction

National Science Foundation

NSF Award No. 2112650

Elian Cantu, Diego Cantu, Ping Xu, Ph.D., Constantine Tarawneh, Ph.D., Heinrich Foltz, Ph.D.

Abstract

A single flaw in a railcar bearing can lead to catastrophic damage, posing serious risks to infrastructure, goods, animals, and people. While traditional train inspections aim to anticipate bearing failure, they are often insufficient because railway workers are unable to capture the real-time conditions of the bearing. This project leverages the powerful tool of machine learning (ML), assisted by onboard sensors, to forecast the remaining mileage of a bearing based on its historical performance, creating a proactive solution to railway safety.

Introduction & Background

- Level 1 analysis is used to detect a defect within the bearing as depicted in Figure 1

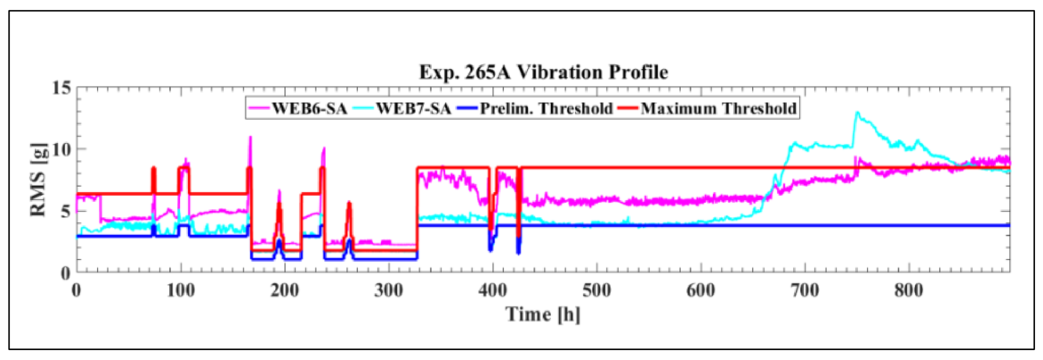

Figure 1: Experiment 265A Level 1 Analysis

- In Figure 1, the root mean square (RMS) values of vibration are plotted against thresholds that indicate statistical levels of the bearing's health over time
- RMS values above the maximum threshold likely indicate defects, so it is a crucial feature to determine the remaining mileage of a bearing
- An ML model, named the LSTM, will be employed to forecast RMS values based on preprocessed datasets
- The model benefits from its ability to manage long-term dependencies in sequential data
- The basic building block of an LSTM network is an LSTM cell, illustrated in Figure 2.

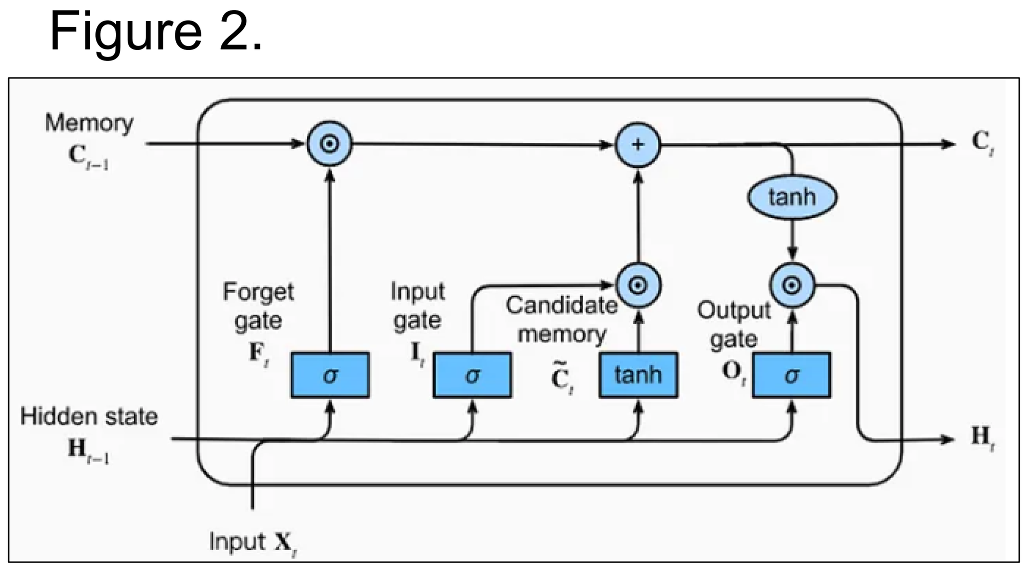

Figure 2: LSTM Cell Architecture

Methodology

The entire procedure consists of utilizing preprocessed datasets to train ML models and form RMS predictions that can later be used to approximate the remaining bearing mileage.

- Step 1: Bearing data is collected from onboard sensors by the University Transportation Center of Railway Safety (UTCRS) every 10 minutes
- Step 2: The data is then transmitted to a central processing unit (CPU) and formatted to a .csv file, so Python can read and load [2]
- Step 3: One or multiple features may be extracted from the dataset for training; however, the model's prediction variable is strictly the RMS feature
- Step 4: Tailor the rest of the model's parameters to enable the optimal training conditions
- Step 5: Redo step 4 until the model predicts RMS values with the least amount of error which in this case are mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE)
- Step 6: Estimate the remaining bearing mileage by ML algorithms with the current bearing features

Data and Results

- Preliminary results were recorded to assess the model's performance, as presented in Table 1
- Experiment 1: Used only RMS as the input feature
- Experiment 2: Used RMS and speed as input features
- Experiment 3: Used RMS, speed, and load as input features
- Three trials were conducted, with each experiment predicted future RMS values
- MSE, MAE, and RMSE quantify the error between the predicted RMS values and the actual values, providing a measure of the model's accuracy

Study 1B	Evaluation Metric	Trial 1	Trial 2	Trial 3
Experiment 1 (Features: RMS)	MAE $[g]$	0.01769	0.02504	0.02133
	MSE $[g^2]$	0.00050	0.00082	0.00065
	RMSE $[g]$	0.02237	0.02869	0.02555
Experiment 2 (Features: RMS and Speed)	MAE $[g]$	0.02413	0.02417	0.02226
	MSE $[g^2]$	0.00077	0.00078	0.00069
	RMSE $[g]$	0.02780	0.02806	0.02627
Experiment 3 (Features: RMS, Speed, and Load)	MAE $[g]$	0.02301	0.02912	0.03030
	MSE $[g^2]$	0.00073	0.00103	0.00111
	RMSE $[g]$	0.02704	0.03220	0.03336

Table 1: Experiment 265A Univariate vs Multivariate Forecasting Results

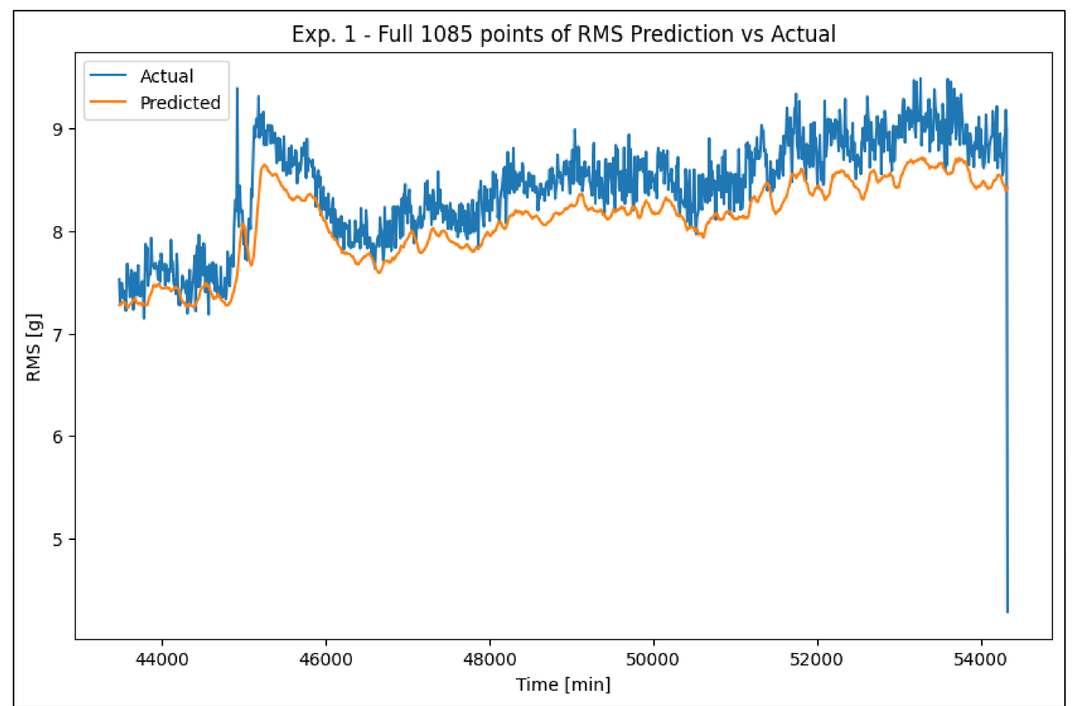

Figure 3: Experiment 1 – Full 1085 points of RMS Prediction vs Actual Values

- Figure 3 represents the plot of the prediction versus actual values for Experiment 1.

Conclusions & Future Work

In the future, the relationship of RMS and mileage over time will be studied using data from previously conducted experiments where the bearing was tested at a constant speed.

The ultimate goal is to determine the remaining mileage using ML algorithms, considering varying bearing conditions including mileage, RMS, speed, and load for a given experiment.

The final step will involve testing the model in real time to validate its performance against a blind experiment.

Acknowledgments

The authors would like to acknowledge funding provided by the National Science Foundation CREST Center for Multidisciplinary Research Excellence in Cyber-Physical Infrastructure Systems (MECIS) under NSF Award No 2112650.

References

[1] Calzone, Ottavio. "An Intuitive Explanation of LSTM." *Medium*, Medium, 10 Apr. 2022, medium.com/@ottaviocalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c.
[2] Martinez, Sergio M. "Vibration-Based Machine Learning Models for Condition Monitoring of Railroad Rolling Stock." ScholarWorks @ UTRGV, August 2023.
[3] R. Han, L. Yang, L. Tan, M. Wang, X. Ma and Z. Wang, "Joint Predictive Replacement Management and Spare Part Planning of High-speed Railway Bearing Under Bayesian Framework," *2022 13th International Conference on Reliability, Maintainability, and Safety (ICRMS)*, Kowloon, Hong Kong, 2022, pp. 99-103, doi: 10.1109/ICRMS55680.2022.9944605

8th Annual STEM Ed Conference, South Padre Island, Texas. February 13 – 15, 2025