Data-Driven Machine Learning Prognostics of Buckling Failure Modes in Ballasted Railway Track

[1]Watcharapong Wongkaew, [2]Wachira Muanyoksakul, [3]Chayut Ngamkhanong
[4]Jessada Sresakoolchai, [5]Sakdirat Kaewunruen

[1]Chulalongkorn University Transportation Institute, Chulalongkorn University, Bangkok, Thailand, Email: watcharapong.w@chula.ac.th
[2]Department of Civil Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand, Email: 6678002021@student.chula.ac.th; chayut.ng@chula.ac.th
[3]Advanced Railway Infrastructure, Innovation and Systems Engineering (ARIISE) Research Unit, Chulalongkorn University, Bangkok,

ABSTRACT

This study presents a machine learning (ML) approach to predict buckling failures in ballasted railway tracks, crucial for enhancing railway safety and reliability. By analysing extensive data from advanced numerical studies, sophisticated ML models are trained to detect failure patterns indicating potential buckling modes. The study uniquely incorporates temperature and track parameters, recognising their impact on snap-through and progressive buckling. Various ML algorithms are evaluated for their effectiveness in simulated and real-world scenarios, demonstrating proficiency in identifying early signs of both snap-through and progressive buckling, leading to timely interventions. This capability not only improves railway safety but also aids in efficient maintenance scheduling and asset management. A case study in Thailand's railway system highlights the model's effectiveness and applicability in real-world scenario. Therefore, this research offers a novel method for predicting buckling failures, contributing significantly to safer and more efficient railway operations under varying environmental conditions.

METHODOLOGY

1.Oversampling Technique

The Synthetic Minority Over-Sampling Technique (SMOTE) is employed to rectify class imbalance within the dataset. This approach synthetically augments the minority class by interpolating new samples within the feature space. The process begins by randomly selecting a minority class sample "a," followed by identifying its k nearest neighbours within the same class.

Table 1 Proportion of data splitting after oversample

Data Split	Data Instances	Train-test-split ratio	Class Split		
			Non-buckling	Snap-through buckling	Progressive buckling
All Data	8,000	100%	3,693	3,693	3,693
Training and validation sets	8,862	80%	2,954	2,954	2,954
Testing set	2,217	20%	739	739	739

2. Machine Learning Techniques

In this paper, various machine learning methods are applied to assess their effectiveness in predicting the buckling modes of ballasted railway tracks under temperature variations. We employ Python 3.11 for constructing models manually and for visualizing the results. Additionally, the K10-Fold Cross Validation technique is used for both training and validation. 6 machine learning techniques are listed below.

1.Logistic Regression
2.K-Nearest Neighbor (KNN)
3.Decision Tree (DT)
4.Random Forest (RF)
5. Light Gradient Boosting (LGBM)
6. Extreme Gradient Boosting (XGB)

3. Performance Measures

To evaluate the performance of the models trained in this study, five statistical analyses were employed: **Accuracy** is a metric that captures the overall performance of the model, **F1-Score** which combines Precision and Recall by using harmonic mean, **Precision** , **Recall** , and **Cross Entropy often referred to "Log loss"** measures the performance by quantifying the difference between predicted probabilities and the actual class labels.

RESULTS

Table 2 displays the performance of each machine learning model

Model	Train set				
	Precision	Recall	F1 score	Cross Entropy	Accuracy
Logistic Regression	0.79	0.79	0.79	-	0.79
k-Nearest Neighbor	0.88	0.87	0.87	-	0.87
Decision Tree	1.00	1.00	1.00	-	1.00
Random Forest	1.00	1.00	1.00	-	1.00
LGBM	0.99	0.99	0.99	0.50	0.99
XGBoost	0.99	0.99	0.99	0.38	0.99

The preliminary model selected, as indicated in Table 2, is the XGBoost model. Based on this selection, the authors proceeded to fine-tune and optimize the XGBoost model, building upon the preliminary version. The XGBoost model contains numerous hyperparameters. Figures 1-2 illustrate the confusion matrices for the XGBoost model on the training and testing datasets, respectively.

Legend 0 – 2958

		True Label (Train Data)		
		Non-Buckling	Snap-Through Buckling	Progressive Buckling
Prediction Label (Model Prediction)	Non-Buckling	2929	30	2
	Snap-Through Buckling	10	2932	0
	Progressive	2	0	2958

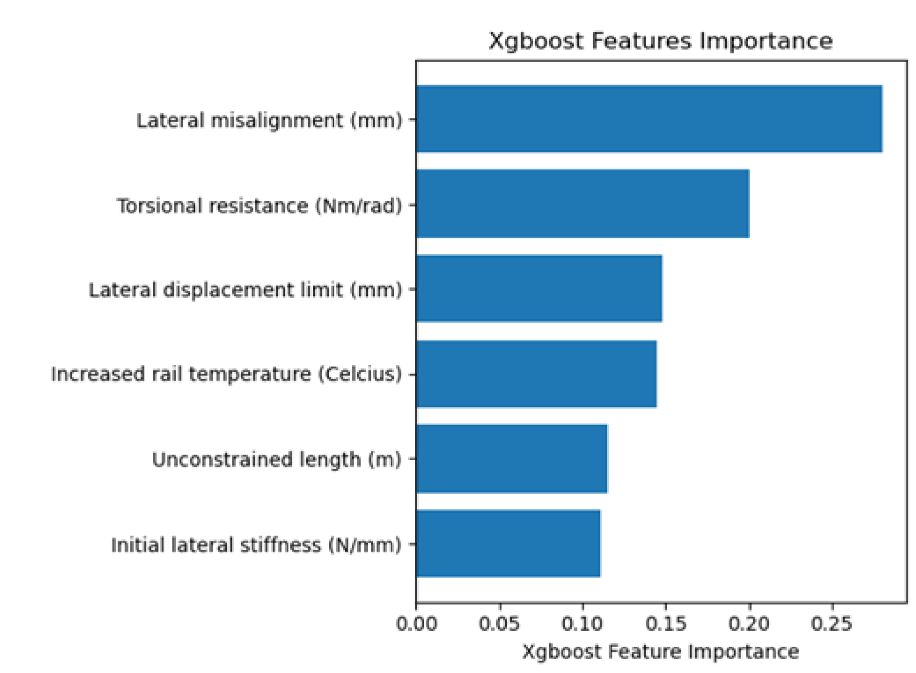
Figure 1 Train Data

Legend 0 – 729

		True Label (Test Data)		
		Non-Buckling	Snap-Through Buckling	Progressive Buckling
Prediction Label (Model Prediction)	Non-Buckling	710	19	3
	Snap-Through Buckling	25	726	0
	Progressive	3	1	729

Figure 2 Test Data

RESULTS



The importance of each parameter is assessed to gain insights into the behaviour of the model and to strategize the maintenance and inspection plan for ballasted tracks. Figure 3 illustrates the feature importance of the trained XGBoost model. It reveals that lateral misalignment, torsional resistance, and lateral displacement limit are the most significant factors, in that order. Lateral misalignment plays a crucial role in track buckling, especially in tangent track sections. This is due to the fact that lateral misalignment can introduce an initial curvature to a track section, providing the necessary eccentricity for axial compression forces to induce track buckling. Torsional resistance is another vital factor influencing track buckling behaviour, along with lateral stiffness that composes of lateral displacement limit and lateral resistance.

Buckling Mode Transition Diagram

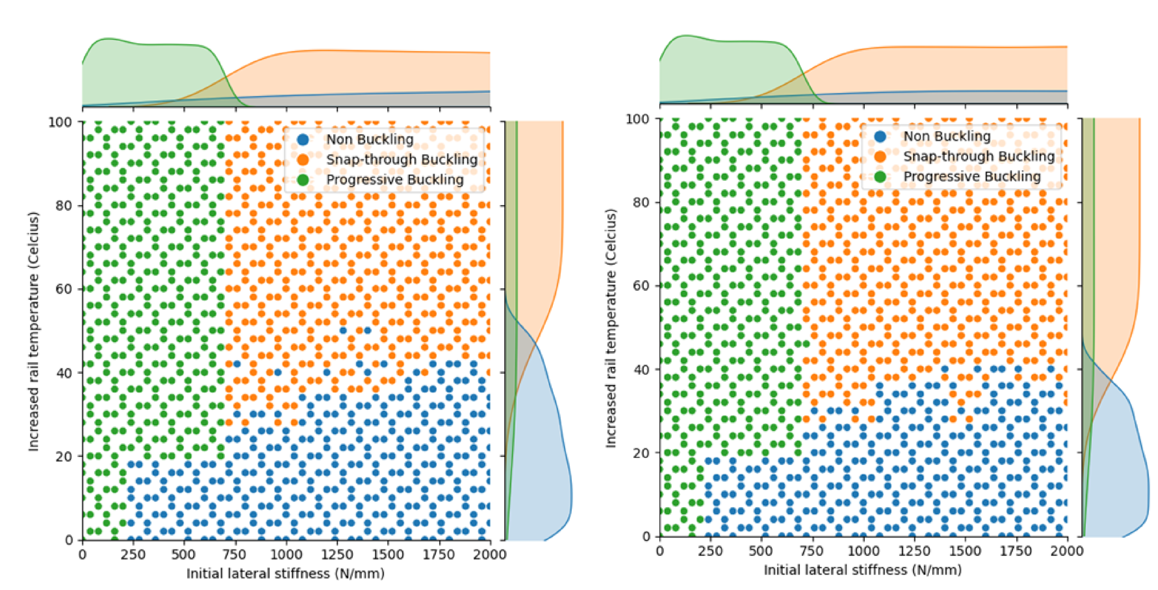


Figure 3 Buckling mode prediction considering initial lateral stiffness and sleeper's displacement limit a) 0.5 mm b) 1 mm c) 2 mm

Figure 3 illustrates the prediction of buckling modes, taking into account lateral resistance and the displacement limit of sleepers while keeping other parameters constant. These plots display a distinct boundary for each type of buckling. The boundary between the buckling and non-buckling areas corresponds to the maximum temperature that can be sustained before buckling occurs. It is evident that progressive buckling modes typically occur when lateral stiffness is low at lower temperatures. As the lateral displacement limit increases, leading to greater lateral resistance force, it becomes apparent that railway tracks are able to withstand higher temperatures, and the occurrence of snap-through buckling becomes more pronounced.

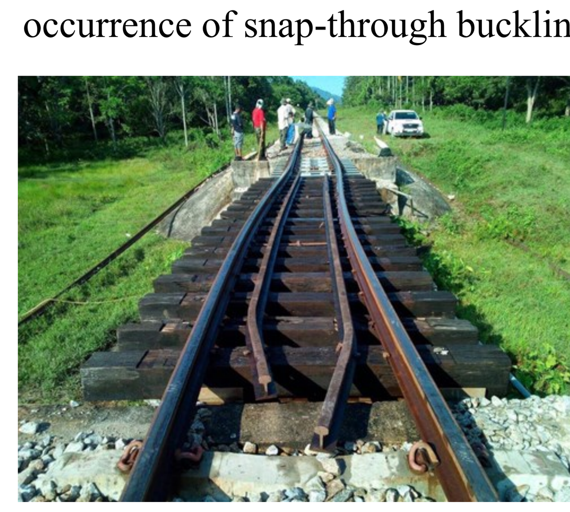


Figure 4 A case study in Thailand (Photo courtesy of State Railway of Thailand)

On April 20, 2017, historical data indicates that the air temperature was recorded at 36°C. Based on this, it was assumed that the rail temperature rose to 56°C and the rail was buckled as shown in Figure 4. The proposed model is utilised for analysis, setting several parameters based on above assumptions. Notably, the predicted temperature at which buckling occurs aligns closely with temperatures recorded in real-world field conditions with only lower bound difference of 3.9% illustrated in Figure 5

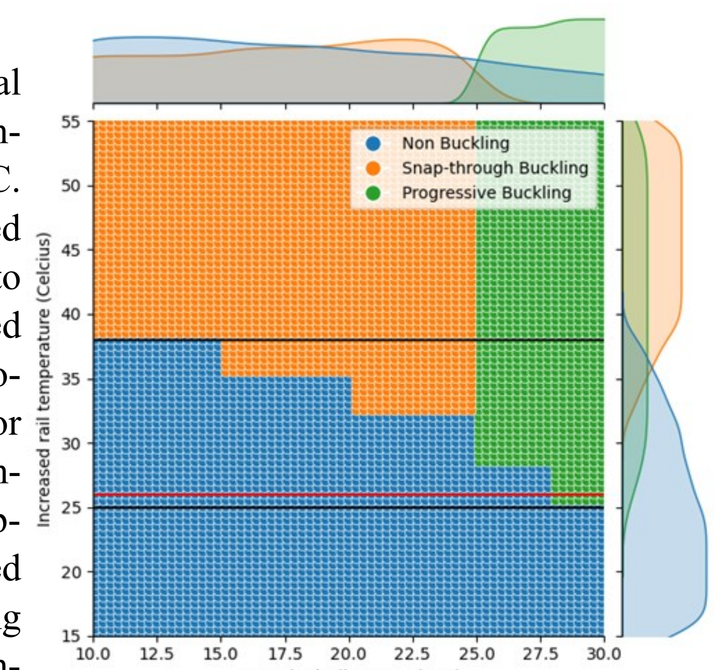


Figure 5 Results of a case study using ML model.

CONCLUSION

This study has demonstrated the substantial potential of machine learning (ML) in enhancing the predictability of buckling failure modes in ballasted railway tracks. Through the deployment of various ML algorithms, including Logistic Regression, Decision Tree, k-Nearest Neighbor, Random Forest and XGBoost, LGBM, we have successfully developed models that offer accurate predictions of track buckling behaviour under a range of conditions. The significant achievement of these models, especially the XGBoost algorithm which recorded an F1 score of 0.97, marks a pivotal advancement in the field of railway engineering. This outcome not only demonstrates the robustness of ML models in complex failure prediction but also opens new avenues for enhancing railway track safety and operational efficiency. We are the first to introduce a detailed analysis of the track buckling phases using ML models, taking into account a variety of different parameters.

This project is supported by National Council Research Thailand (NCRT: N42A650199). This Research is also funded by Thailand Science Research and Innovation Fund Chulalongkorn University (BCG66210020). The authors would like to acknowledge the support from Chulalongkorn University for the establishment of Advanced Railway Infrastructure, Innovation and Systems Engineering (ARIISE) Research Unit.

CHULA ENGINEERING
Foundation toward Innovation