"The SVM is good at solving data with a small sample, nonlinearity, and high dimensions and thus has been used for pattern recognition, regression analysis, and time series prediction for track degradation." **[1]**

Support Vector Machine (SVM)

Benefits:

1. SVM models can consider many factors for track degradation because the complexity of SVM is irrelevant to the dimension of input variables.
2. SVM models can still achieve high prediction accuracy when the sample amount (inspection data) is limited.
3. SVM models are good at predicting track geometry data with nonlinearity and high dimensions.
4. The SVM can assign new examples to one category or another by transforming the input space into a high-dimensional space using an optimal classification surface.

Concerns:

1. SVM models have low interpretability and are difficult to understand the physical mechanism behind the prediction results. In addition, it is also difficult to select appropriate kernel functions for SVM models.
2. To overcome this shortcoming, SVM models are often combined with other methods to optimize model parameters, increase computational efficiency, and expand the scope of application.

Research indicates that "Falamarzi et al. have compared the prediction results (gauge) of the SVM model and that of the ANN model and found that the prediction accuracy of the SVM is higher in curved sections while lower in straight sections." **[1]**

"The choice of tree-based classification techniques is motivated by the fact that they are comparatively easy to interpret and have empirically shown to provide optimal results for small and structured datasets" **[2]**

"With the class labelling, the dataset was found to be highly imbalanced in nature particularly for the maintenance activity and trigger's status prediction. Our preliminary analysis also presents the high bias of the model towards majority classes, which results in poor classification ability. To mitigate data imbalance and to improve the classification precision, the classes having the large data representation must be under-sampled." **[2]**

Random Forest (RF)

Benefits:

1. High predictive performance on structured datasets
2. Effective on structured, multi-source maintenance data
3. No need for data normalization or feature scaling
4. Built-in feature selection and interpretability

Concerns:

1. Sensitive to class imbalance
2. Misclassification on minority classes
3. Requires tuning of tree number

[1] Liao, Y, Han, L, Wang, H & Zhang, H 2022, 'Prediction Models for Railway Track Geometry Degradation Using Machine Learning Methods: A Review', *Sensors*, vol. 22, no. 19, p. 7275, doi: https://www.mdpi.com/1424-8220/22/19/7275.

[2] 'Predictive maintenance using tree-based classification techniques: A case of railway switches' 2019, *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 35–54, doi: https://www.sciencedirect.com/science/article/abs/pii/S0968090X18309057.