Proceedings of the Seventh International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 15, Paper 8.1
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.15.8.1

Pantograph Failure Analysis and Short-Term Forecasting for EMU Fleet Using Hybrid and Machine Learning Models

B. Alkali[1], M. M. Rahman[1], J. P. Gutierrez[1], A. K. Jain[2], C. Mcneil[3] and J. Nelson[3]

[1]School of Science and Engineering, Glasgow Caledonian University, United Kingdom
[2]Department of Mechanical Engineering, Birla Institute of Technology and Science, India
[3]Siemens Mobility Limited, Glasgow, United Kingdom

Abstract

Pantograph systems are critical to the reliable operation of electric multiple unit (EMU) trains because they maintain electrical contact with the overhead catenary. Failures in these systems can lead to maintenance issues, service disruptions, and train cancellations. This study presents a data-driven framework for pantograph failure analysis and short-term forecasting for the EMU fleet operating in Scotland. A five-year historical maintenance data was analysed, together with operational failure records, to ascertain the pantograph failure pattern at both fleet and unit levels. The methodology combines data preparation, time-series construction, multi-step forecasting, and quantile-based risk classification. Four forecasting models were compared: Autoregressive Integrated Moving Average (ARIMA)-Artificial Neural Network (ANN), Long Short-Term Memory (LSTM)-Support Vector Regression (SVR), LSTM, and Decision Tree Regressor. The results show that the ARIMA-ANN hybrid model achieved the best overall performance, with the lowest forecasting errors among the compared methods, indicating that the pantograph failure series contains both linear and nonlinear characteristics. Based on the best-performing model, fleet-level and unit-level forecasts were generated for the next 15 operational periods and level into low, medium, and high-risk. The proposed framework provides technically interpretable and practically useful decision support for proactive railway maintenance planning and service delivery.

Keywords: pantograph, forecasting, predictive maintenance, railway rolling stock, ARIMA-ANN, LSTM, decision tree regressor.

1 Introduction

Electric multiple unit (EMU) trains depend on an efficient current collection system to acquire electrical power from overhead catenary lines. The pantograph is an essential component that maintains uninterrupted electrical contact between the train and the overhead contact wire, ensuring a stable power supply to the traction system. Failures in pantograph equipment can disrupt electricity collection and lead to operational disruptions, such as train delays or service cancellations. Majority of the electric multiple unit fleet operating within the electrified railway network in Scotland use a roof-mounted pantograph which are installed on the Pantograph Trailer Open Standard Lavatory Wheelchair (PTOSLW) vehicle.

The pantograph considered in this paper is designed as a single-arm and has an air-operated current collector with secondary head suspension, consisting of several key assemblies including the base frame, lower arm, upper arm, control rod, cross tube, pan head and pneumatic control equipment. The general configuration of the pantograph used in the fleet is shown in Figure 1. The components work together to maintain a stable contact force between the pantograph head and the overhead contact wire during operation, ensuring efficient power transmission to the traction system.

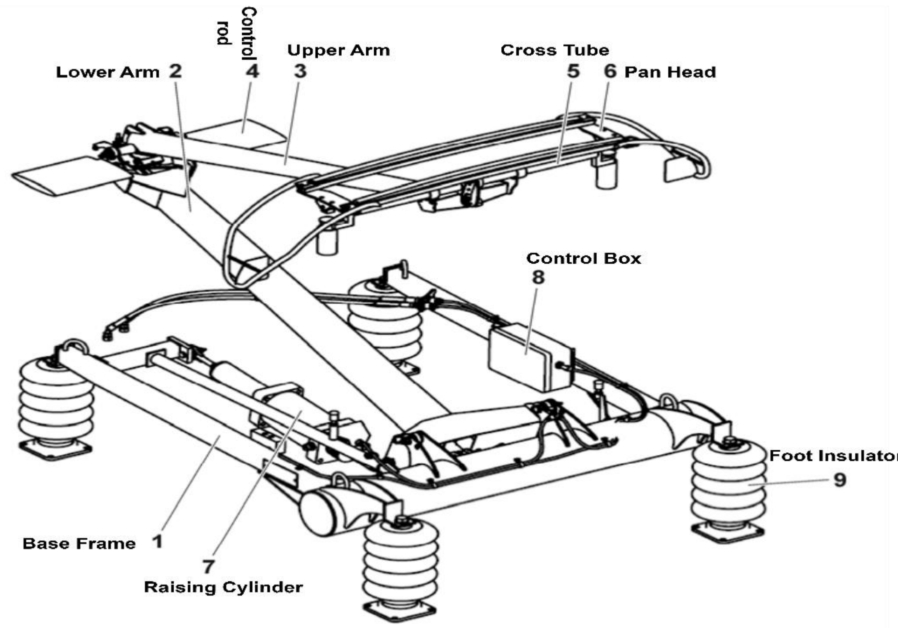

Figure 1. Pantograph electric multiple unit

Figure 1 shows that the pantograph is the roof of the PTOSLW vehicle and is pneumatically operated to elevate and lower the collector head. The pan head remains in contact with the overhead contact wire, while the articulated arm structure enables the pantograph to accommodate vertical fluctuations of the overhead line during train operation.

The maintenance dataset from February 2011 to January 2026, pantograph downtime recorded over the period is approximately 868 hours across the fleet. Operational records from February 2022 to February 2026 indicate 24 pantograph-related technical failures were encountered, resulting in 143 minutes of service delay, which include 19 full-service cancellations and 17 partial-service cancellations. These results are calculated based on an average operating time of approximately 6000 hours per year. The results show that a pantograph technical failure occurs approximately every 1000 operating hours, approximately every 42 days, which is consistent with the pattern of anomalies detected in practice.

In this paper, an interactive pantograph failure analysis with short-term forecasting is conducted to aid the assessment of maintenance requirements and operational decision-making. The forecasting provides us with individual unit and fleet-level insights from historical maintenance data, failure descriptions, and fleet information, along with their risk levels which are classified as Low, Medium, and High category. To have a better oversight of the risk levels, a traffic light system is used: where green indicates Low risk, orange indicates medium risk, and red indicates high risk respectively. Machine learning models were fitted to the dataset to predict pantograph failure trends over 15 railway operational periods. The Supervised ML methods were compared with hybrid models (Autoregressive Integrated Moving Average (ARIMA) Artificial Neural Network (ANN)) and a Long Short-Term Memory (LSTM), Support Vector Regression (SVR) model as well as a decision tree model. The resulting framework provides technical interpretability and supports maintenance planning and service delivery.

2 Fleet data performance

Exploratory data analysis (EDA) was performed on the pantograph failure and maintenance datasets to understand fleet failure behaviour and component replacement patterns. The failure dataset was first analysed to identify the primary reasons for pantograph-related maintenance events. The most frequently recorded activities include pantograph checks after overhead line (OHL) trips, panhead replacement, and carbon strip damage. These results suggest that a significant proportion of pantograph maintenance activities are associated with interactions between the pantograph and the overhead contact system, as well as with normal wear of key current-collection components. Understanding these failure patterns provides useful insights for maintenance prioritisation and operational reliability improvement.

The historical maintenance dataset was analysed to investigate component replacement process. The analysis indicates that carbon strips, carbon carrier assemblies, and fastening elements such as washers and bolts represent the largest share of maintenance replacements. This pattern reflects the mechanical wear and electrical contact degradation experienced by pantograph components during regular train service operations. These insights help support maintenance planning and resource allocation for pantograph systems.

3 Methodology

Pantograph failure forecasting is considered an intermittent-demand prediction problem, as the series exhibits many random failures and many zero-demand periods where demand is sparse, highly variable, and difficult to model using conventional forecasting approaches [1, 5]. Intermittent-demand studies show that such series are difficult to forecast and require appropriate error metrics, such as MAE, RMSE, and sMAPE [2, 6]. A hybrid and machine learning-based forecasting models, including ARIMA-ANN, LSTM-SVR, LSTM, and Decision Tree Regressor, to address both linear and nonlinear temporal behaviour. The ARIMA-ANN structure follows a hybrid linear-nonlinear forecasting approach, while neural network forecasting is used for intermittent demand in time series analysis [3]. The hybrid deep learning and machine learning structures such as LSTM-SVR have shown strong capability for modelling nonlinear sequential patterns and improving forecasting performance in complex prediction problems [4, 7]. Tree-based forecasting approaches have also been recognised as useful benchmark models in time-series forecasting studies [8].

A data-driven forecasting framework for analysing pantograph failures utilising historical maintenance records from the fleet is presented in this paper. The aim is to forecast short-term pantograph rate of occurrence of failure at both the fleet and individual unit levels by evaluating hybrid statistical-machine learning, hybrid deep learning-machine learning, deep learning, and tree-based regression models.

The comprehensive methodology encompasses data preparation, time-series construction, model development, multi-step forecasting, risk classification, and model evaluation. The modelling workflow framework is illustrated in Figure 2.

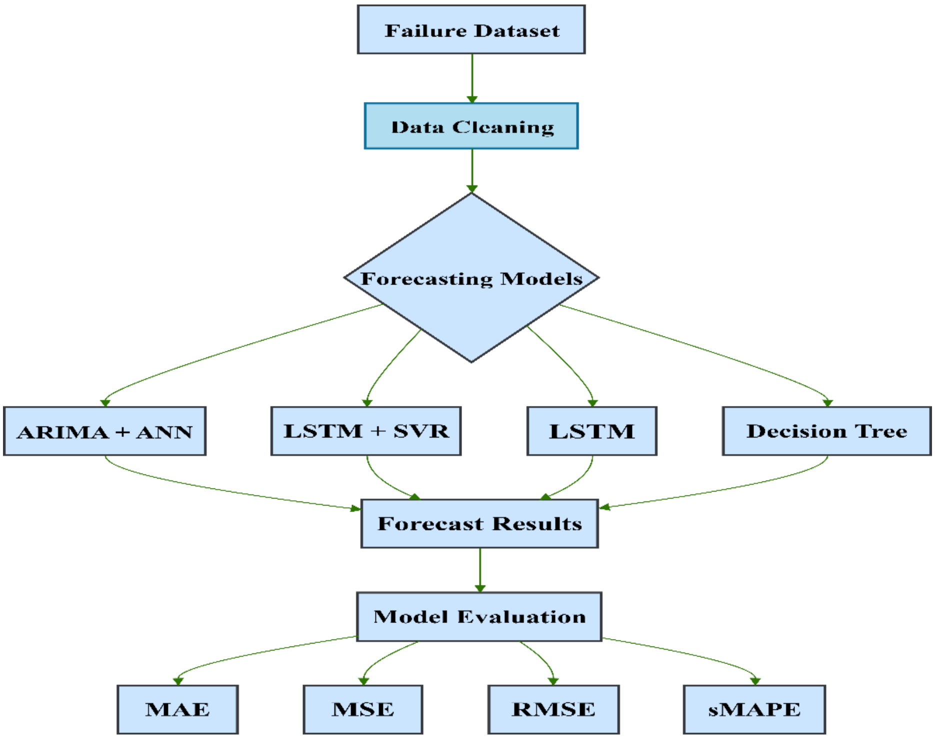

Figure 2: Flow chart of pantograph failure forecasting model

3.1 Data preparation and time-series construction

The raw dataset consists of pantograph maintenance records of train unit, maintenance period, and part description. Part description text approach was used and normalised by removing special characters and standardising equivalent labels for consistency. The failure pattern was identified by aggregating the number of pantographs-related maintenance events for each train unit across all maintenance periods.

Let y_t denote the observed number of pantograph failures in the period t where, $y_t =$ number of failures in the period t, $t = 1,2,3 \ldots \ldots T$ and $T =$ total number of observed periods. The failure time series for each train unit is therefore represented as

$$Y = \{y_1, y_2, .., y_T\} \tag{1}$$

To preserve the intermittent structure of the data, missing combinations of train unit and maintenance period were set to zero. A sliding-window transformation was then applied to convert the sequence into supervised learning samples. For a lookback window of length L, the input vector is defined as

$$X_t = [y_{t-L}, y_{t-L+1}, \ldots\ldots, y_{t-1}] \tag{2}$$

and the prediction target is y_t. The lookback window was fixed at $L = 12$ periods, while the forecasting horizon was set to $H = 15$ future periods.

3.2 ARIMA-ANN hybrid model

The first forecasting approach combines autoregressive integrated moving-average (ARIMA) and artificial neural network (ANN) models to capture both linear and nonlinear temporal patterns in the pantograph failure series. The ARIMA component models the linear structure of the series as

$$y_t = c + \sum_{i=1}^{p} \emptyset_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{3}$$

Where, $c =$ constant term, $\emptyset_i =$ autoregressive coefficient at lag i , $\theta_j =$ moving average coefficient at lag j, ε_t = random error term, p = autoregressive order, $q =$ moving average order. Here, ARIMA (1,0,1) was used. After fitting the ARIMA model, the residual sequence

$$e_t = y_t - \hat{y}_t^{ARIMA} \tag{4}$$

Where, e_t residual at period t, $\hat{y}_t^{ARIMA}$ = ARIMA prediction at period t. The ANN component was then trained on the residual sequence to model nonlinear information not captured by ARIMA. The residual forecast is defined as

$$\hat{e}_t = f(e_{t-L}, e_{t-L+1}, \ldots\ldots, e_{t-1}) \tag{5}$$

Where, $f(\cdot)$ denotes the nonlinear ANN mapping. The final hybrid forecast is then obtained by combining the two components:

$$\hat{y}_t = \hat{y}_t^{ARIMA} + \hat{e}_t \tag{6}$$

3.3 LSTM–SVR hybrid model

The second model combines long short-term memory (LSTM) and support vector regression (SVR). The LSTM network first learns temporal patterns from the historical failure sequence, and the resulting features are then passed to an SVR model for final regression. Given an input sequence X_t. The LSTM updates its internal memory through gated operations. The forget gate equation (7), the input gate equation (8), the candidate cell state equation (9), the update cell state equation (10), the output gate equation (11), the hidden state equation (12):

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{7}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{8}$$

$$\tilde{C}_t = \tanh\,(W_c[h_{t-1}, x_t] + b_c) \tag{9}$$

$$C_t = f_t \odot C_{t-1}, + i_t \odot \tilde{C}_t \tag{10}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{11}$$

$$h_t = o_t \odot \tanh\,(C_t) \tag{12}$$

Where, x_t input at time t, h_t hidden state, C_t cell state, (W_f, W_i, W_c, W_o = weight matrices), ($b_f, b_i, b_c, b_o =$ bias vectors), $\sigma(\cdot)$ sigmoid activation function, $\odot =$ element-wise multiplication. The temporal features extracted by the LSTM are then used as inputs to the SVR model, which estimates the regression function

$$\hat{y} = g(z) = w^T \emptyset(z) + b \tag{13}$$

Where, z = feature vector is extracted by LSTM, $\emptyset(z)$ = transformed feature space, w = weight vector, b = bias term. This hybrid structure allows LSTM to learn sequential representations, while SVR performs nonlinear regression on the learned features.

3.4 LSTM model

The third forecasting approach is a standalone LSTM model. Using the same sliding-window input sequence X_t, the LSTM directly learns a nonlinear mapping from historical failure observations to future failure demand:

$$\hat{y}_t = f(X_t) \tag{14}$$

The model uses stacked LSTM layers followed by a dense output layer. This structure enables the model to capture short- and long-term temporal dependencies without requiring manual feature extraction.

3.5 Decision Tree Regressor

The fourth model is a Decision Tree Regressor, which was included as a machine learning baseline. For each rolling window, a small set of summary features and recent lag values was extracted. The feature vector is represented as:

$$z_t = [\, \mu_t, \sigma_t, m_t, y_{t-1}, y_{t-2}, y_{t-3}] \tag{15}$$

Where, μ_t = mean of the lookback window, σ_t = standard deviation of the lookback window, m_t = maximum value in the lookback window, $y_{t-1}, y_{t-2}, y_{t-3}$ = recent lagged observations. The Decision Tree Regressor learns a piecewise nonlinear mapping:

$$\hat{y}_t = f(z_t) \tag{16}$$

3.6 Risk analysis

In addition to numerical forecasting, a simple risk analysis layer was introduced to support maintenance decision-making. The predicted failure values were converted into qualitative risk categories using a quantile-based classification scheme derived from the historical failure distribution of each train unit. The 33rd and 66th percentiles of the historical failure series for a given train unit. The forecast risk level for the future period $t + h$ was assigned as

$$R_{(t+h)} = \begin{cases} Low, & \hat{y}_{t+h} \leq Q_{0.33} \\ Medium, & Q_{0.33} < \hat{y}_{t+h} \leq Q_{0.66} \\ High, & \hat{y}_{t+h} > Q_{0.66} \end{cases} \tag{17}$$

Where, $R_{(t+h)}$ = risk category assigned to period $t + h$, $\hat{y}_{t+h}$ = forecasted failure count, $Q_{0.33}$ lower risk threshold, $Q_{0.66}$ high risk threshold. Low-risk periods correspond to relatively low expected maintenance requirement, medium-risk periods indicate moderate concern, and high-risk periods indicate elevated failure likelihood requiring closer maintenance attention.

3.7 Model evaluation

The predictive performance of the models was evaluated using four forecasting accuracy metrics: mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and symmetric mean absolute percentage error (sMAPE). Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{18}$$

Where, y_i = actual failure count, $\hat{y}_i$ = predicted failure count, n = number of test observations.
Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{19}$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{MSE} \tag{20}$$

Symmetric Mean Absolute Percentage Error (sMAPE):

$$sMAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{2|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i| + \varepsilon} \tag{21}$$

Where, $\varepsilon =$ a very small positive constant is introduced to avoid division by zero.

4. Analysis and results

This section presents the comparative forecasting performance of the developed models, followed by the short-term pantograph failure forecasts and associated risk interpretation. Firstly, the four forecasting models are compared using standard error metrics. Secondly, the best-performing model is analysed at the fleet level, and finally, its unit-level forecasting and risk classification capability are demonstrated.

4.1 Model comparison

The model results in Table 1 illustrates the forecasting performance of the four models using MAE, MSE, RMSE, and sMAPE. The calculated lower values indicate better predictive performance. Among the compared models, the ARIMA-ANN hybrid model achieved the best overall accuracy, with the lowest MAE (11.09), MSE (214.70), RMSE (14.65), and sMAPE (27.41) values. This indicates that the hybrid combination of linear ARIMA modelling and nonlinear ANN residual learning is more effective in capturing the temporal structure of pantograph failure than the other approaches considered in this study.

Model	MAE	MSE	RMSE	sMAPE
ARIMA-ANN	11.09	214.70	14.65	27.41
LSTM-SVR	16.67	396.79	19.92	37.77
LSTM	15.30	318.59	17.84	35.0
Decision Tree	22.88	681.60	26.11	47.20

Table 1: Forecasting performance comparison model

The LSTM model achieved the second-best performance in terms of MAE, MSE, and RMSE, suggesting that deep learning can also model the nonlinear behaviour of pantograph failures reasonably well. However, the LSTM-SVR hybrid model did not outperform the standalone LSTM model, indicating that the additional SVR stage did not improve predictive accuracy for this dataset. The Decision Tree Regressor produced the largest errors across all reported metrics, indicating that tree-based regression was less effective at capturing the sequential and intermittent characteristics of the pantograph failure series.

4.2 Fleet-level forecast of the best model

The graph in Figure 3 shows the fleet-level pantograph failure forecast produced by the ARIMA-ANN model. The historical failure trend is represented by the blue line, while the forecasted values for the next 15 periods are shown as coloured markers. The results indicate that fleet-level pantograph failure pattern is expected to fluctuate over the short-term horizon. Using quantile-based thresholds, the forecasted values are classified into low (green), medium (orange), and high (red) risk categories. The occurrence of several medium- and high-risk periods suggests that elevated maintenance requirement may persist in the near future, supporting the need for proactive inspection and maintenance planning.

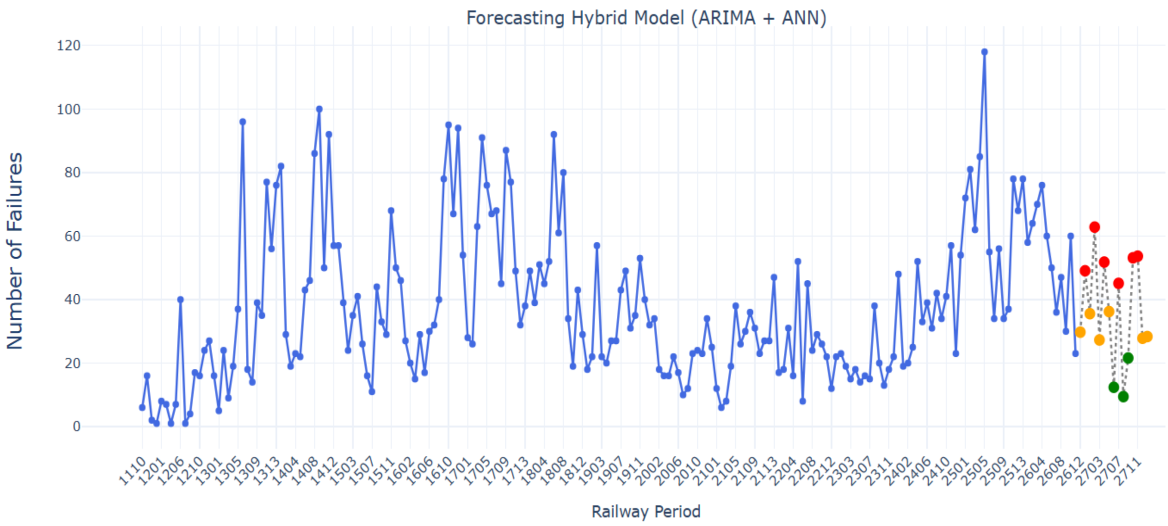

Figure 3: ARIMA-ANN forecast pantograph failures at fleet level

4.3 Unit-level forecast and risk analysis

Similarly, the graph in Figure 4 presents the 15-period forecast for a representative Unit A using the ARIMA-ANN model. The forecast indicates a variable short-term pattern, with predicted failure counts ranging from zero to moderate levels across the horizon. This suggests that maintenance requirement for the selected unit is not constant and may fluctuate over the upcoming periods.

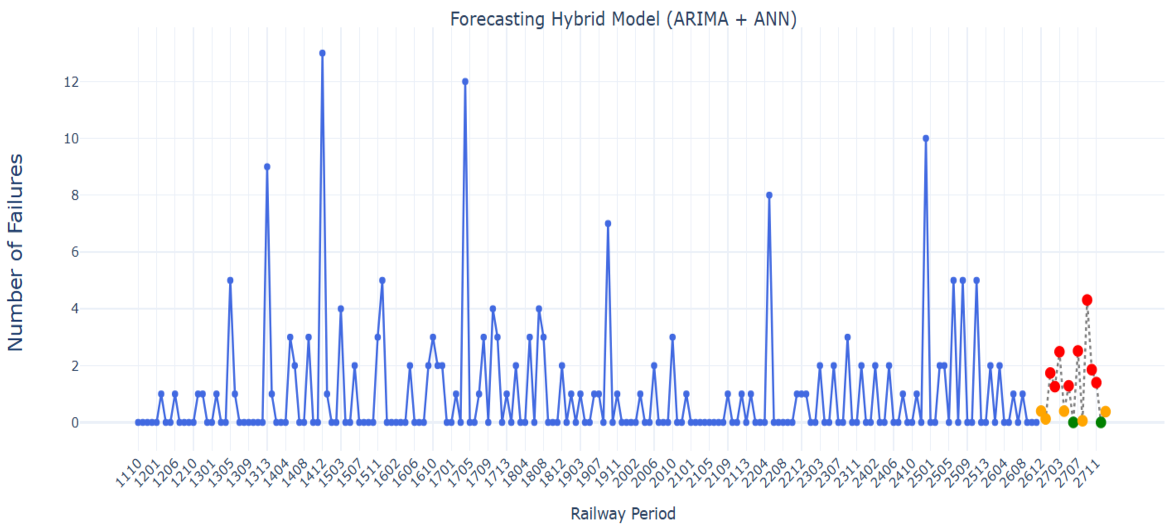

Figure 4: ARIMA-ANN forecast pantograph failures for Unit A

The resulting risk categories for each forecast period are summarised in Table 2. The table shows that the unit is expected to experience a mix of low, medium, and high-risk periods, with several higher-risk periods in the early and middle parts of the forecasting horizon.

Period	Risk level
1	MEDIUM
2	MEDIUM
3	HIGH
4	HIGH
5	HIGH
6	HIGH
7	HIGH
8	MEDIUM
9	LOW
10	HIGH
11	LOW
12	LOW
13	MEDIUM
14	LOW

Table 2: Forecast risk level

These results provide a more operationally meaningful view of the forecast by identifying which specific future periods may require closer monitoring, earlier inspection, or increased spare-parts readiness.

5 Conclusions and Contributions

This study developed a data-driven framework for analysing and forecasting pantograph failures in the EMU fleet using maintenance failure records and operational failure information. The findings show that pantograph failures exhibit intermittent, variable behaviour, making short-term prediction a challenging yet important task for maintenance planning. By comparing ARIMA-ANN, LSTM-SVR, LSTM, and Decision Tree models, we found that the ARIMA-ANN hybrid model delivered the best overall forecasting performance, achieving the lowest MAE, MSE, RMSE, and sMAPE values among the evaluated approaches. This suggests that pantograph failure patterns contains both linear and nonlinear temporal patterns, which are better captured by hybrid modelling than by machine learning or tree-based regression. The selected ARIMA-ANN model was then used to generate short-term forecasts at both fleet and unit levels for the next 15 periods.

The results presented in this paper demonstrate that the proposed framework can support proactive maintenance decision-making by identifying future failure and translating numerical forecasts into practical low, medium-, and high-risk levels. At the fleet level, the forecast indicated fluctuating short-term maintenance requirement, while at the unit level, the results highlighted periods requiring closer monitoring and possible early intervention.

Acknowledgements

The authors would like to thank Siemens Mobility Limited Glasgow, UK for funding this research project.

References

[1] Tian, X. and Wang, H., 2021. Forecasting intermittent demand for inventory management by retailers: A new approach. *Journal of Retailing and Consumer Services*, *62*, p.102662.
[2] Kim, S. and Kim, H., 2016. A new metric of absolute percentage error for intermittent demand forecasts. *International journal of forecasting*, *32*(3), pp.669-679.
[3] Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, pp.159-175.
[4] Guo, J., Xie, Z., Qin, Y., Jia, L. and Wang, Y., 2019. Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM. *Ieee Access*, *7*, pp.42946-42955.
[5] Pinçe, Ç., Turrini, L. and Meissner, J., 2021. Intermittent demand forecasting for spare parts: A Critical review. *Omega*, *105*, p.102513.
[6] Hyndman, R.J. and Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International journal of forecasting*, *22*(4), pp.679-688.
[7] Tanyıldızı Ağır, T., 2025. Daily Global Solar Radiation Prediction with Hybrid LSTM-SVM: The Case of Nusaybin. *Arabian Journal for Science and Engineering*, pp.1-14.
[8] Januschowski, T., Wang, Y., Torkkola, K., Erkkilä, T., Hasson, H. and Gasthaus, J., 2022. Forecasting with trees. *International Journal of Forecasting*, *38*(4), pp.1473-1481.