Proceedings of the Seventh International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 15, Paper 8.4
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.15.8.4

Railway Track Risk-Based Degradation Modelling: A Data-Driven Approach

N. G. D. D. Santos[1], B. Alkali[1], A. K. Jain[2], O. Niculita[1]

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

Abstract

Smart monitoring technologies and artificial intelligence are increasingly deployed in railway systems, generating large volumes of operational data from in-service infrastructure. However, translating these heterogeneous and indirect measurements into actionable maintenance decisions remains challenging, particularly because railway track degradation processes are partially observable, non-stationary, and influenced by complex interactions between traffic loading, environment, and maintenance actions.

This paper presents a degradation-aware modelling framework for railway infrastructure that integrates wayside condition monitoring data with probabilistic state-space degradation modelling. The framework is formulated as a decision-relevant analytical engine that can be coupled to existing railway digital twin environments, rather than as a standalone or geometric digital twin. Asset condition is represented through latent degradation states inferred from axle-level dynamic response measurements, operational variables, and environmental factors using a Dynamic Multi-Regression Hidden Markov Model.

The inferred degradation states are linked to state-conditioned hazard models and Remaining Useful Life estimation, enabling monitoring data to be translated into uncertainty-aware risk metrics suitable for maintenance planning. A case study focusing on rolling contact fatigue demonstrates the ability of the proposed framework to identify degradation regimes and decision-relevant risk windows consistent with operational practice. The results highlight the value of probabilistic degradation

modelling as an enabling component within digital twin-based railway asset management systems.

Keywords: smart monitoring, artificial intelligence, probabilistic modelling, predictive maintenance, railway infrastructure, railway track, dynamic multi-regression hidden Markov model, remaining useful life.

1. Introduction

Smart monitoring technologies and artificial intelligence (AI) are increasingly deployed across railway systems, enabling continuous acquisition of high-resolution operational data from in-service infrastructure. Wayside condition monitoring (WCM) systems, in particular, provide axle-resolved measurements of dynamic wheel-rail and track-support interactions as trains pass fixed locations, offering insight into the mechanical loading environment experienced by railway assets [1; 8]. These developments present significant opportunities to improve maintenance planning and infrastructure reliability.

Despite this potential, translating WCM outputs into actionable maintenance decisions remains challenging. The measurements primarily capture dynamic response rather than direct evidence of physical damage. Consequently, degradation mechanisms such as rolling contact fatigue, ballast deterioration, and loss of support remain partially observable and evolve in a non-stationary manner under the combined effects of traffic loading, environmental exposure, and maintenance interventions [4; 5; 12]. Conventional threshold-based approaches therefore provide limited support for predictive and risk-based decision-making [7; 11].

Digital twins are increasingly used in railway asset management to integrate data and decision processes [6,8]. However, most implementations remain descriptive, emphasising monitoring and visualisation rather than explicitly modelling degradation dynamics and uncertainty [2]. This paper therefore proposes a degradation-aware analytical framework designed for integration with existing railway digital twin environments.

By integrating WCM data with a Dynamic Multi-Regression Hidden Markov Model, the framework infers latent degradation states and links them to hazard rates and Remaining Useful Life estimates. These outputs can be fed into digital twin dashboards to provide near-real-time asset condition awareness and risk-based maintenance triggers. A case study on rolling contact fatigue illustrates the operational relevance of the approach.

2. Degradation-Aware Digital Twin Model Formulation

Figure 1 illustrates the representation of the degradation-aware modelling framework based on a Dynamic Multi-Regression Hidden Markov Model (DMR-HMM). High-frequency monitoring data are first subjected to data role assignment and 24-hour temporal aggregation. The resulting observation and covariate vectors drive a probabilistic state-space model in which latent degradation states evolve under covariate-dependent transition dynamics. Inference via forward-backward smoothing and Viterbi decoding yields posterior state probabilities and degradation trajectories, which are subsequently linked to state-conditioned Weibull hazard models and probabilistic Remaining Useful Life estimation. The framework supports uncertainty-aware prognostics and maintenance decision-making.

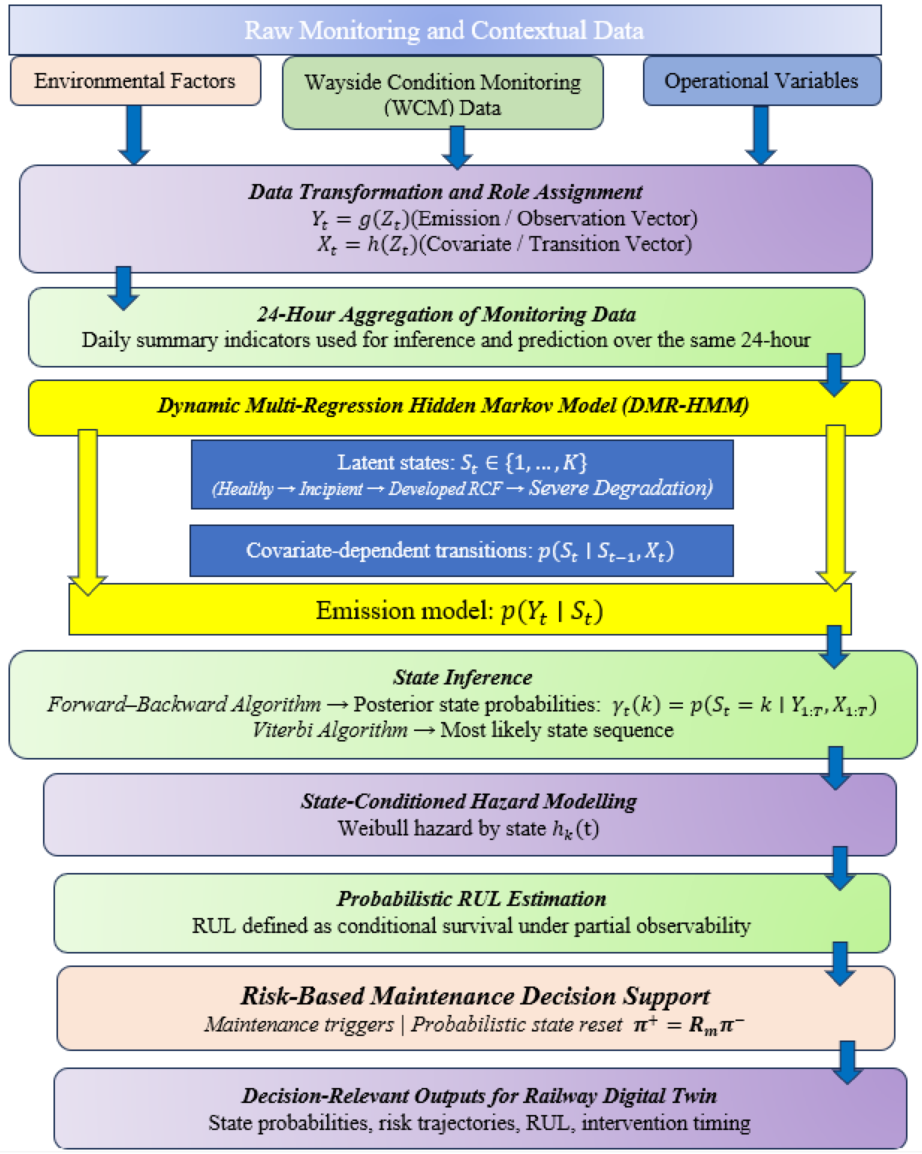

Figure 1: degradation-aware modelling framework based on a Dynamic Multi-Regression Hidden Markov Model (DMR-HMM

The degradation-aware digital twin is formulated as a probabilistic state-space model that links heterogeneous monitoring data to latent degradation dynamics. Let

$$\boldsymbol{Z}_t = \{Z_t^{(w)}, Z_t^{(o)}, Z_t^{(e)}\} \tag{1}$$

denote the raw data available at time t prior to any modelling role assignment, where $Z_t^{(w)}$represents axle-level dynamic response measurements, $Z_t^{(o)}$denotes operational variables, and $Z_t^{(e)}$includes environmental factors such as temperature and rainfall. The vector Z_t therefore acts as a data container, rather than a direct input to the degradation model.

$$Y_t = g(Z_t), \qquad X_t = h(Z_t), \tag{2-3}$$

where Y_t is an observation (emission) vector providing noisy evidence of asset condition, and X_t is a covariate (transition) vector governing degradation dynamic. Environmental variables may appear in both Y_t and X_t, reflecting their dual role as sources of response variability and as physical drivers of degradation [3; 9].

Asset condition is represented by a latent degradation state

$$S_t \in \{1,2, \dots, K\}, \tag{4}$$

represents the unobserved condition regime of the asset at time t. Each state corresponds to a qualitatively distinct degradation regime, ordered by increasing severity. In this study, the states represent: (1) healthy condition, (2) incipient damage, (3) developed squat/RCF damage, and (4) severe degradation approaching intervention limits [5;12; 13].

Conditional on S_t, the observation vector follows the emission distribution

$$p(Y_t|S_t = k) \sim N(\mu_k, \Sigma_k), \tag{5}$$

while degradation evolution is governed by covariate-dependent transitions

$$P(\, S_t = j \mid S_{t-1} = i, X_t\,) = \frac{exp(\alpha_{ij}+\beta_{ij}\top X_t)}{\sum_{m=1}^{K} \exp\,(\alpha_{im}+\beta_{im}\top X_t)} \tag{6}$$

Uncertainty in initial condition is captured by

$$\pi_k = p(S_1 = k) \tag{7}$$

reflecting pre-existing damage prior to the first observation.

The *complete joint distribution of states and observations factorises* as

$$p(S_{1:T}, Y_{1:T}|X_{1:T}, \Theta) = p(S_1) \prod_{t=2}^{T} p(S_t \mid S_{t-1}, X_t) \prod_{t=1}^{T} p\,(\,Y_t|\, S_t) \tag{8}$$

which constitutes the core mathematical definition of the digital twin, explicitly linking $Z_t \to (Y_t, X_t) \to S_t$.

In addition to the most likely state sequence obtained via the Viterbi algorithm, inference in the DMR-HMM yields the smoothed posterior probability of each latent degradation state at time t. The posterior probability of occupying state k at time t is defined as

$$\gamma_t(k) = p(S_t = k \mid Y_{1:T}, X_{1:T}, \Theta) \tag{9}$$

where $Y_{1:T}$ denotes the sequence of observed feature vectors, $X_{1:T}$ the corresponding covariates, and $\Theta = \{\pi, \alpha_{ij}, \beta_{ij}, \mu_k, \Sigma_k\}$ the estimated model parameters.

The posterior probabilities $\gamma_t(k)$ are computed using the forward-backward algorithm, which combines information from past and future observations to provide a smoothed estimate of state occupancy [3; 10]. While the Viterbi algorithm yields the single most likely state sequence, posterior probabilities provide a full probabilistic description of inferential uncertainty by allocating probability mass across all states at each time step [9]. Such smoothed state posteriors are widely used in state-space and Hidden Markov Models to support uncertainty-aware decision-making and downstream prognostic analysis.

To translate probabilistic state inference into maintenance-relevant decision metrics, each latent degradation state k is associated with a state-conditioned hazard function:

$$h_k(t) = \frac{\beta_k}{\eta_k}\left(\frac{t}{\eta_k}\right)^{\beta_k - 1} \tag{10}$$

where β_k and η_k are the shape and scale parameters of the Weibull distribution for state k. Railway track degradation processes such as rolling contact fatigue and loss of support are dominated by wear-out mechanisms, in which failure risk increases with accumulated damage and exposure to repeated loading cycles. Such behaviour is naturally represented by Weibull hazard functions with shape parameter β>1, which are widely adopted in mechanical and railway reliability modelling to characterise degradation-driven failures [5; 7; 12; 14]. By associating a Weibull hazard model with each latent degradation state, the proposed framework enables a physically interpretable escalation of failure risk as assets transition into more severe degradation regimes, providing a well-established link between inferred degradation behaviour and maintenance-relevant risk metrics.

Given partial observability and uncertainty in future state evolution, Remaining Useful Life is defined as a conditional survival quantity rather than a deterministic prediction. The RUL at time τ is expressed as:

$$\mathrm{RUL}(\tau) = \mathbb{E}\left[\inf\{t > \tau : S_t \in \mathcal{F}\} \mid \mathcal{I}_\tau\right] \tag{11}$$

where F denotes failure or intervention states and $\mathcal{I}_\tau$ is the information available at time τ. RUL is therefore expressed as a probabilistic quantity reflecting uncertainty in both current degradation state and future evolution [11].

Maintenance interventions are incorporated within the model as probabilistic state re-initialisation events, defined by

$$\pi^+ = \mathrm{R}_m \, \pi^- \tag{12}$$

Where $\pi^-$and $\pi^+$are pre- and post-maintenance state distributions, and R_m is a maintenance-specific reset matrix. This formulation avoids unrealistic as-new assumptions and allows residual uncertainty to propagate into subsequent risk and RUL predictions [7].

This decision layer enables intervention timing and prioritisation to be based on sustained risk accumulation and uncertainty, rather than solely on instantaneous condition indicators.

2.1.Study Area and Units of Analysis

The study focuses on a 4-mile track section in the Linlithgow area (Glasgow-Edinburgh direction) of the Edinburgh-Glasgow Main Line (EGM1, ERL 1100), analysed over February 2023. The section accommodates intensive mixed passenger and freight traffic at line speeds of up to 100 mph, generating sustained dynamic loading conditions representative of fatigue-driven degradation processes.

To ensure consistency between observation, inference, and prediction horizons, high-frequency wayside condition monitoring (WCM) data are aggregated over 24-hour intervals, such that each day is represented by a single set of summary indicators. These daily aggregated observations are used to infer latent degradation states and to generate predictions over the corresponding 24-hour period, aligning the temporal resolution of condition assessment with the decision-relevant timescale of maintenance planning.

The location has been identified as a recurrent squat defect hotspot, with multiple historical events recorded in the Rail Defect Management System (RDMS), indicating persistent susceptibility to rolling contact fatigue. The unit of analysis is defined at the track-segment level, consistent with the spatial resolution of the WCM system and RDMS reporting. While WCM data provide indirect measurements of dynamic loading and support conditions, RDMS records are treated as failure-event observations for hazard modelling and Remaining Useful Life (RUL) estimation.

3. Results Summary Based on 24-Hour Aggregated Sample Data

3.1. Latent Degradation State Inference

Figure 2 shows the evolution of latent degradation states for Track ID 1100 during February 2023, combining the most-likely state sequence obtained from the Viterbi algorithm with the corresponding smoothed posterior probabilities $\gamma_t(k)$. The results indicate prolonged periods in which the track remains within a single degradation regime, interrupted by short transition intervals that reflect progressive changes in condition.

During stable operating periods, posterior probabilities are strongly concentrated on one state, demonstrating reliable inference despite the indirect and noisy nature of the WCM measurements. In contrast, transition intervals are characterised by a temporary spread of posterior mass across neighbouring states, indicating gradual degradation progression rather than abrupt regime switching.

From a maintenance perspective, this behaviour is informative: stable regimes support confidence in normal operation, while periods of increased uncertainty signal emerging risk that may warrant closer monitoring.

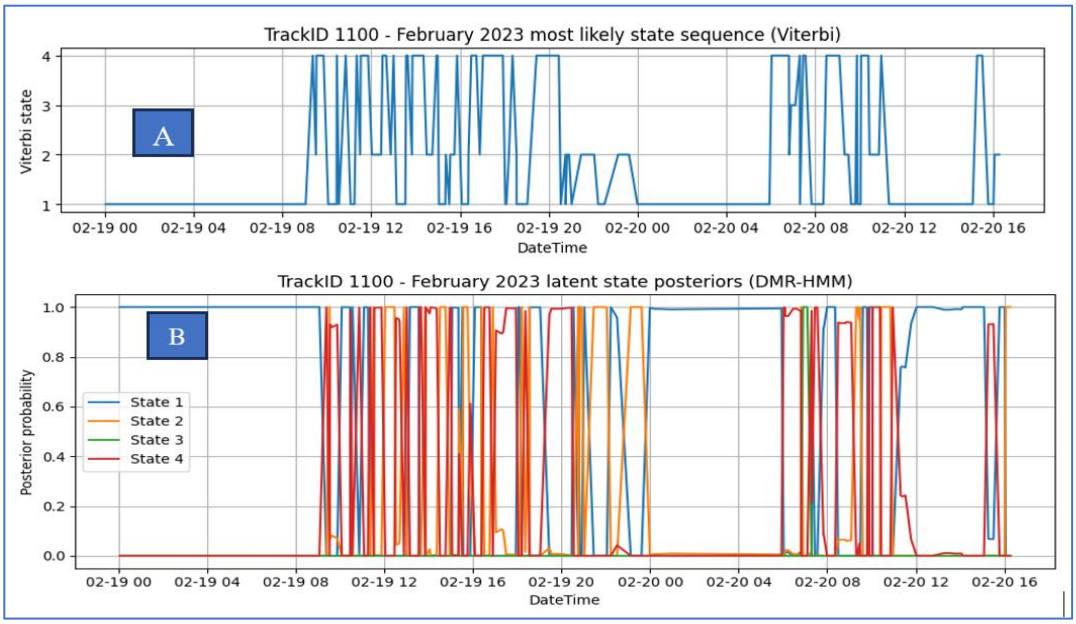

Figure 2: Latent Degradation State Inference results for Track ID 1100 for February 2023.

3.2. Decision-Relevant Risk Windows

Figure 3 presents the time-varying degradation risk for Track ID 1100 during February 2023, defined as the combined posterior probability of the track occupying severe latent degradation states. Two decision thresholds are applied to translate probabilistic risk into actionable guidance. Exceedance of the WATCH threshold (Q75) indicates periods where targeted inspection is advisable, whereas exceedance of the ACTION threshold (Q90) identifies intervals where maintenance intervention is assured. These risk windows correspond to persistent increases in latent degradation probability.

Conversely, periods dominated by healthy degradation states are associated with consistently low risk levels.

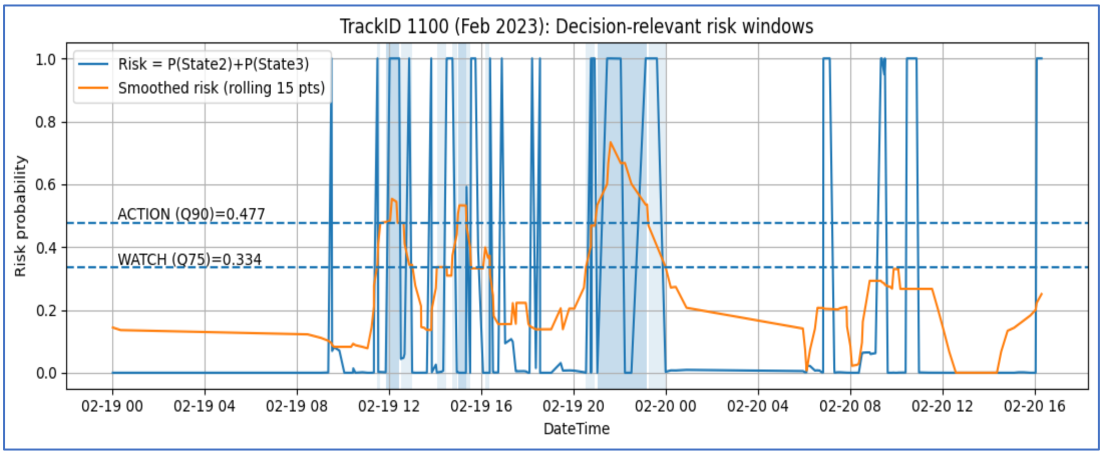

Figure 3: Decision-Relevant Risk Windows (Track ID 1100 – February 2023)

3.3. Transition sensitivity analysis

Figure 4 summarises the sensitivity of latent degradation state transitions to key operational covariates. Total train weight and maximum axle load emerge as the dominant drivers of transitions between degradation regimes, exerting a substantially stronger influence than train speed.

Axle count and train length show a moderate contribution, consistent with cumulative fatigue effects associated with repeated loading. In contrast, speed exhibits a comparatively limited influence once loading-related variables are accounted for, indicating that its effect is secondary rather than primary.

Overall, these results suggest that degradation progression at the monitored location is governed predominantly by loading-related factors, with kinematic variables playing a lesser role. This finding is consistent with fatigue-driven degradation mechanisms and supports prioritising load-related indicators in predictive maintenance and decision-support applications.

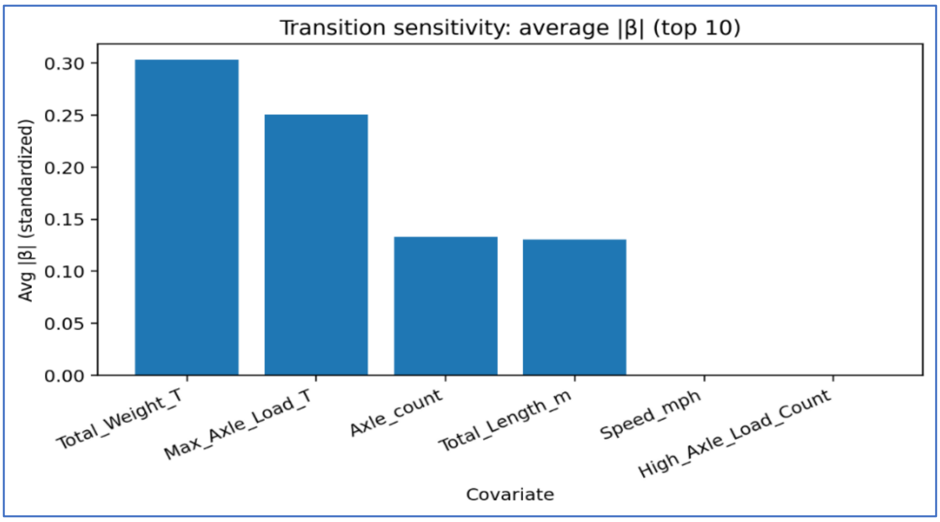

Figure 4: Sensitivity of degradation transition dynamics to operational covariates.

3.4. Latent Degradation Severity versus Axle Count

Figure 4 Expected latent degradation severity E[state] versus raw axle count for Track ID 1100 (four-state model). Each point represents a passage-level estimate, while the dashed line denotes a weak overall linear trend. Elevated degradation severities are observed primarily at moderate axle counts, indicating that axle count alone is insufficient to explain degradation severity and that other loading and operational factors play a significant role.

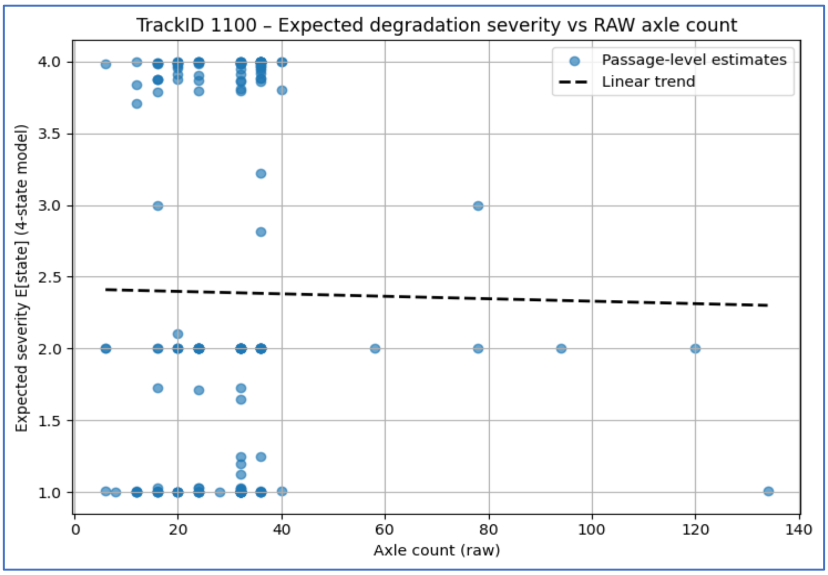

Figure 5: Expected latent degradation severity versus axle count for Track ID 1100 (four-state model).

3.5. Expected RUL Mixture over Latent States

figure 6 presents the expected Remaining Useful Life (RUL) estimated as a mixture over latent degradation states, with each time-step weighted by the posterior state probabilities inferred by the DMR-HMM. The resulting RUL profile shows sharp downward excursions separated by extended periods of relative stability. These transient RUL drops coincide with temporary increases in the probability of higher-severity degradation states, while subsequent recovery indicates that such escalations are not persistent. Overall, the behaviour reflects uncertainty in both current condition and future degradation evolution, supporting a risk-aware approach to maintenance planning rather than deterministic life prediction.

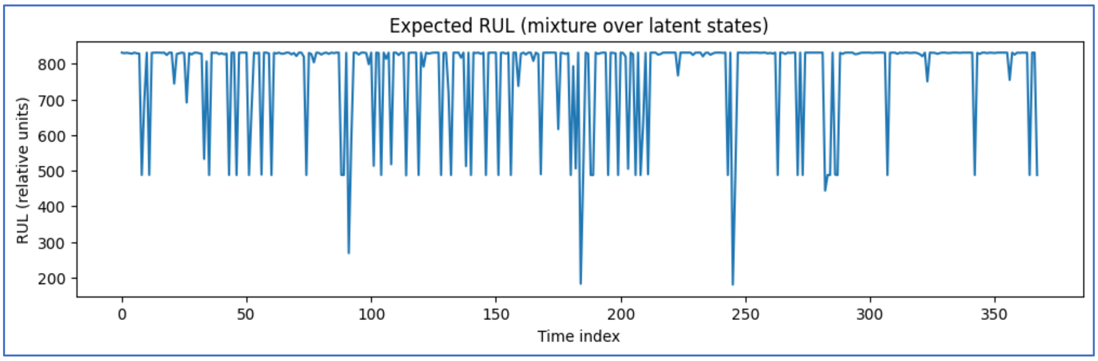

Figure 6: Expected Remaining Useful Life (RUL) estimated as a mixture over latent degradation states.

4 Discussion and Implications

The results show that track degradation evolves through persistent latent regimes with short transition phases, providing a stable and interpretable representation of condition under partial observability. Transition dynamics are driven primarily by loading-related variables, confirming the dominant role of cumulative contact stresses.

Risk windows derived from sustained threshold exceedance translate probabilistic inference into actionable maintenance signals, while mixture-based RUL estimates capture uncertainty in future degradation trajectories. Together, these findings demonstrate the suitability of the proposed framework for risk-informed maintenance planning within railway digital twin environments.

5 Conclusions and Recommendations for Future Work

5.1 Conclusions

This paper proposed a degradation-aware digital twin framework that integrates wayside condition monitoring data with a DMR-HMM to support risk-based railway maintenance under uncertainty. The results show that latent degradation regimes can be reliably inferred from indirect measurements and that degradation escalation is driven primarily by traffic loading. By translating probabilistic state inference into risk windows and mixture-based RUL estimates, the framework supports maintenance decisions based on sustained risk rather than isolated condition changes.

5.2. Recommendations for Future Work

Future work should incorporate inspection and intervention data for validation, extend the framework to multiple track sections for network-level analysis, and investigate integration with existing asset management systems for near real-time decision support.

Acknowledgements

we gratefully acknowledge Network Rail Scotland for providing access to data and operational insights that made this research possible. Their support and collaboration were essential in enabling the practical evaluation of the proposed degradation-aware digital twin framework within a real railway operating context.

References

[1] Boller, C., Chang, F.-K. and Fujino, Y. (2016). Encyclopedia of Structural Health Monitoring. Chichester: John Wiley & Sons.

[2] Boje, C., Guerriero, A., Kubicki, S. and Rezgui, Y. (2020). Towards a semantic construction digital twin: Directions for future research. Automation in Construction, 114, 103179.

[3] Cappé, O., Moulines, E. and Rydén, T. (2005). Inference in Hidden Markov Models. New York: Springer.

[4] Esveld, C. (2001). Modern Railway Track. 2nd edn. Zaltbommel: MRT-Productions.

[5] Grassie, S.L. (2005). Rolling contact fatigue on the railhead: A review. Wear, 258(7–8), pp.1310–1326.

[6] Grieves, M. and Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In: Kahlen, F.-J., Flumerfelt, S. and Alves, A. (eds.) Transdisciplinary Perspectives on Complex Systems. Cham: Springer, pp.85–113.

[7] Jardine, A.K.S., Lin, D. and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7), pp.1483–1510.

[8] Kaewunruen, S., Lian, Q., Chen, Y. and Yin, Z. (2021). Digital twin aided sustainability and vulnerability audit for railway infrastructure. Sustainability, 13(15), 8128.

[9] Murphy, K.P. (2012). Machine Learning: A Probabilistic Perspective. Cambridge, MA: MIT Press.

[10] Rabiner, L.R. (1989). A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), pp.257–286.

[11] Sankavaram, C., Kodali, A., Pattipati, K., Singh, S., Salman, M. and Buckley, S. (2013). Prognostics and health management: A review. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, 43(4), pp.466–480.

[12] Zerbst, U., Lundén, R., Edel, K.-O. and Smith, R.A. (2009). Introduction to the damage tolerance behaviour of railway rails – a review. Engineering Fracture Mechanics, 76(17), pp.2563–2601.

[13] Network Rail (2020). Rail Defect Management System (RDMS): Guidance Note.

[14] Nelson, W. (2004). Applied Life Data Analysis. Wiley.