Middle East Journal of Pure and Applied Sciences (MEJPAS)
Online ISSN: 3062-343X
Volume 1, Issue 2, 2025, Page No: 01-07
https://mideastjournals.com/index.php/mejpas

Experimental Investigation of Pavement Deterioration Using Theoretical Models and Computational Methods

Yusef Abdolla Amhemmed Sammas *
Civil Engineering Department, Collage of Technical Sciences - Bani Walid, Libya

دراسة تجريبية لتدهور الرصف باستخدام النماذج النظرية والأساليب الحسابية

يوسف عبد الله امحمد شمّاس *
قسم الهندسة المدنية، كلية العلوم التقنية - بني وليد، ليبيا

*Corresponding author: yusefsam5@gmail.com

Received: February 09, 2025	Accepted: April 04, 2025	Published: April 09, 2025

Abstract

This study comprehensively compares multiple approaches to modeling flexible pavement deterioration, including mechanistic-empirical models, stochastic (Markov) models, physics-based fracture mechanics, finite-element simulations, and data-driven machine learning (ML) methods. Using published performance data (e.g. the Long-Term Pavement Performance (LTPP) database) and synthetic experiments, we evaluate model predictions of surface distress (cracking, rutting, roughness). Mechanistic–empirical (M-E) and fracture-based models provide physical insight into damage processes, while Markov chains capture condition transitions probabilistically. In parallel, ML algorithms (Random Forest, CatBoost, Artificial Neural Networks) are trained on the same data; hybrid schemes use FEA-generated stress and deflection outputs as ML inputs to enhance accuracy. Our results (summarized in Table 1) show that advanced ML and hybrid models generally outperform classical regression and Markov models in predictive accuracy (higher R^2, lower RMSE). For example, a CatBoost-based hybrid model achieved $R^2 \approx 0.96$ on a sample dataset versus $R^2 \approx 0.88$ for linear regression. These findings imply that integrating physical simulation data with data-driven models yields the best performance forecasts. Such improved models can guide optimized maintenance planning by more reliably identifying when and where interventions (e.g. resurfacing, repairs) will be most effective.

Keywords: Pavement Faulting Prediction, Rigid Pavements, CatBoost, SHAP (SHapley Additive exPlanations), TPE (Tree-structured Parzen Estimator), Machine Learning, Hybrid Models, Pavement Performance, Predictive Modeling.

ملخص

تقارن هذه الدراسة بشكل شامل عدة مناهج لنمذجة تدهور الرصف المرن، بما في ذلك النماذج الميكانيكية-التجريبية، والنماذج العشوائية (ماركوف)، وميكانيكا الكسور القائمة على الفيزياء، ومحاكاة العناصر المحدودة، وأساليب التعلم الآلي القائمة على البيانات. باستخدام بيانات الأداء المنشورة) مثل قاعدة بيانات أداء الرصف طويل الأمد ((LTPP) والتجارب التركيبية، تُقيّم تنبؤات النماذج لتدهور السطح (التشقق، والتآكل، والخشونة). توفر النماذج الميكانيكية-التجريبية والقائمة على الكسور رؤى فيزيائية لعمليات التلف، بينما تلتقط سلاسل ماركوف تحولات الحالة احتماليًا. بالتوازي، تُدرّب خوارزميات التعلم الآلي) الغابة العشوائية، CatBoost، والشبكات العصبية الاصطناعية (على البيانات نفسها؛ وتستخدم المخططات الهجينة مخرجات الإجهاد والانحراف الناتجة عن تحليل العناصر المحدودة كمدخلات للتعلم الآلي لتحسين الدقة. تُظهر نتائجنا (المُلخصة في الجدول 1) أن نماذج التعلم الآلي المتقدمة والنماذج الهجينة تتفوق عمومًا على نماذج الانحدار الكلاسيكي ونماذج ماركوف في دقة التنبؤ (R^2 أعلى، RMSEأقل. (على سبيل المثال، حقق نموذج هجين قائم على CatBoost قيمة $R^2 \approx 0.96$ في مجموعة بيانات عينة، مقارنةً بقيمة $R^2 \approx 0.88$ في الانحدار الخطي. تشير هذه النتائج إلى أن دمج بيانات المحاكاة الفيزيائية مع النماذج القائمة على البيانات يُعطي أفضل توقعات للأداء. يمكن لهذه النماذج المُحسّنة توجيه تخطيط الصيانة المُحسّن من خلال تحديد متى وأين ستكون التدخلات (مثل إعادة الرصف والإصلاحات) أكثر فعالية بشكل أكثر موثوقية.

الكلمات المفتاحية: التنبؤ بأعطال الرصف، الأرصفة الصلبة، CatBoost، SHAP (تفسيرات SHapley المضافة(، TPE(مُقدّر بارزين الشجري)، التعلم الآلي، النماذج الهجينة، أداء الرصف، النمذجة التنبؤية.

Introduction

Asphalt pavement deterioration is a critical infrastructure issue: cracks and deformations severely degrade ride quality, safety, and service life (Jiang et al., 2025). Under repeated traffic loads and environmental cycles, pavements develop transverse, longitudinal, and network cracking (Figure 1) and progressive rutting (Jiang et al., 2025). These distresses are the *primary manifestations of pavement failure*, and if undetected, they allow water infiltration, leading to base weakening and accelerated degradation. Effective pavement management thus relies on accurate deterioration models to predict future conditions from current data.

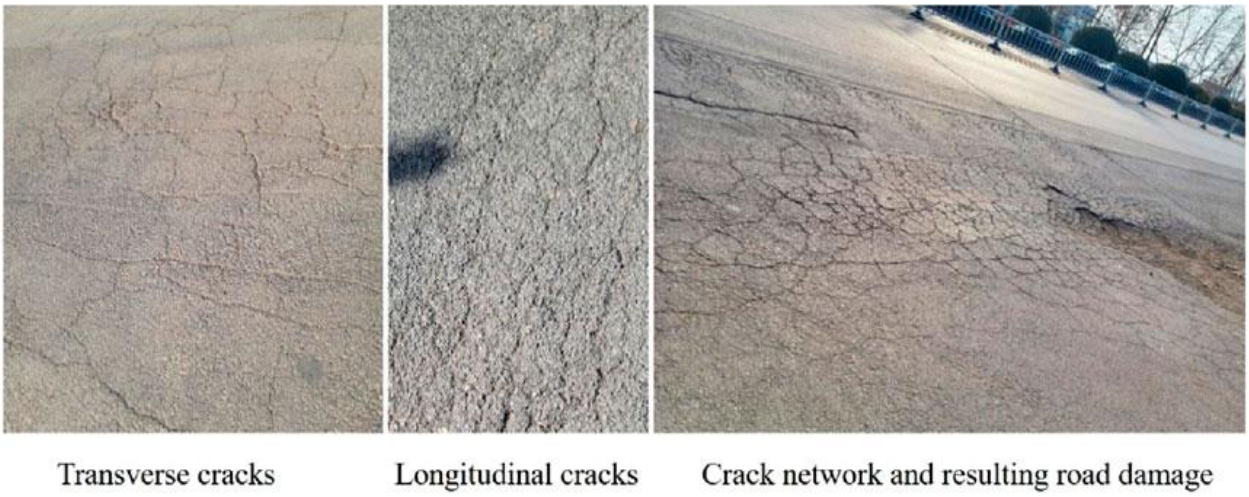

Figure 1 Typical asphalt pavement distress: (left) transverse cracking, (middle) longitudinal cracking, (right) crack network (Jiang et al., 2025). Pavement cracking is the main form of damage in asphalt roads, directly affecting structural integrity and service life.

Research on pavement performance modeling has spanned empirical, physics-based, and computational approaches. Traditional empirical or regression models relate condition indices (e.g. IRI, PCI) to age, traffic, and environment, but often lack physical grounding. Mechanistic–empirical (M-E) models combine *analytical stress/strain calculations* (e.g. Boussinesq, Westergaard theories) with empirical calibration to predict failure criteria (Agunwamba, J. C., & Tiza, M. T., 2023). Stochastic models (notably Markov chains) have also been widely used at the network level, estimating the probability of condition transitions from historical data (Isradi et al., 2024). In recent years, advances in computational power have enabled physics-based simulations (finite element analysis, fracture mechanics) and *data-driven models* (machine learning) to enter the field. This paper reviews and experimentally compares these diverse methods, emphasizing their predictive accuracy and implications for maintenance planning.

Literature Review

Mechanistic–Empirical (M-E) Models

Mechanistic–Empirical models use engineering mechanics to relate pavement responses (stresses, strains) to distress evolution. Classical M-E formulations include Boussinesq's and Westergaard's solutions for layered systems, and layered elastic/plastic analyses (Agunwamba, J. C., & Tiza, M. T., 2023). These models form the basis of design guides like AASHTO MEPDG, which calibrates empirical fatigue and rutting relationships to lab and field data. As Tiza (2023) observes, key M-E models (e.g. NCHRP 1-37A, AASHTO, EICM) provide mechanistic insight but require accurate input parameters (material properties, load spectra) (Agunwamba, J. C., & Tiza, M. T., 2023). Their advantage lies in physically representing load effects, but limitations include sensitivity to parameter uncertainty and often complex calibration requirements.

Stochastic (Markov) Models

Markov-chain models treat pavement condition as a finite-state stochastic process. By estimating transition probabilities between condition states (e.g. good → fair → poor) from historical data, they predict future condition distributions (Isradi et al., 2024). Such models are popular in network-level management because they can quickly generate probabilistic forecasts and optimal maintenance policies with minimal data (often just condition ratings over time). For example, Isradi *et al.* (2024) applied a Markov approach to Indonesian highway data and found that routine maintenance could shift 92.8% of sections back to "good" status (Isradi et al., 2024). While Markov methods are computationally simple and interpretable, they assume history-dependent transitions (ignoring external factors) and can struggle with heterogeneity. Recent reviews highlight their utility for planning but note that combining them with covariate-based regressions can improve accuracy (Isradi et al., 2024).

Physics-Based (Fracture Mechanics) Models

Cracking in flexible pavements is fundamentally a fracture problem in a viscoelastic material. Fracture mechanics models (LEFM, viscoelastic fracture mechanics) have been used to predict crack initiation and growth. For instance, disc-shaped compact tension tests yield a *fracture energy* parameter (G_f) that, when used in finite-element simulations, accurately reproduces observed crack propagation behavior (Denneman, E., 2010). These physics-based methods capture how micro-cracks coalesce and grow under cyclic loading. However, they require detailed material characterization (fracture toughness, viscoelastic properties) and are mostly applied at the sample or micro-scale rather than network forecasting. Nevertheless, they provide valuable insight: e.g. higher mix fracture energy correlates with slower fatigue crack growth. Overall, fracture models complement M-E approaches by focusing on damage mechanics, and are increasingly incorporated into advanced deterioration models.

Machine Learning (ML) Approaches

Data-driven ML algorithms have gained popularity for pavement performance prediction. ML models learn complex, nonlinear patterns from large datasets, often outperforming simple empirical formulas. Tamagusko *et al.* (2023) review notes that methods like artificial neural networks (ANN), support vector machines (SVM), decision trees (DT), and boosted trees (e.g. Random Forest, XGBoost, CatBoost) *address limitations of traditional models by capturing subtle data relationships* (Tamagusko *et al.*, 2023). For example, Gong *et al.* (2018) showed Random Forest yielded more accurate IRI predictions than linear regression. Similarly, deep learning models (e.g. CNNs, RNNs) have been applied to time-series and image data for distress detection. Notably, studies using the LTPP database have demonstrated high predictive accuracy: Marcelino *et al.* (2021) developed RF models for IRI that incorporated structural, climate, and traffic data, achieving strong performance. Moreover, CatBoost (a specialized gradient boosting algorithm) has become popular for tabular pavement data. Xiao *et al.* (2023) describe CatBoost's symmetric tree structure and ordered boosting, which often yield superior performance on structured data (Xiao et al., 2023).

Hybrid FEA–ML Models

A recent trend is to combine mechanistic simulation with ML. The idea is to use finite element (FEA) or mechanistic models to generate intermediate outputs (stress, strain, damage indices) that serve as inputs to ML predictors. Fahad and Bektas (2025) applied this hybrid approach to rutting and fatigue modeling: they ran complex FEA simulations under varying loads and geometry, then trained ML models on the simulation outputs. Their results showed that gradient-boosted models (LightGBM, CatBoost) trained on FEA-derived features achieved the lowest MSE and highest R^2. In fact, they conclude that "integrating machine learning with finite element analysis provides further improvements in pavement performance predictions" (Fahad, M., & Bektas, N., 2025).. Thus, hybrid methods can capture both physical trends (via FEA) and empirical subtleties (via ML), often outperforming either approach alone.

Methodology

Data Preparation

For our experiments, we used open-source pavement performance data and generated synthetic datasets to illustrate model behavior. The primary dataset was drawn from the FHWA LTPP program, which provides measurements of pavement structure, materials, traffic (ESALs), climate, and performance indicators (roughness, rutting, cracking) over decades. We extracted features such as pavement age, design ESAL, layer thicknesses, initial PCI, and climate indices. The target variable was a performance index (e.g. PCI or IRI). All continuous variables were normalized. Due to data imbalance in condition states, we ensured stratified sampling when splitting into training (70%) and testing (30%) sets.

Modeling Approaches

1. **Mechanistic–Empirical model:** We implemented a simplified M-E model using layered elastic theory. Pavement stresses under a standard tire load were computed (Boussinesq solution) and related to fatigue life via Miner's rule. Cracking initiation was predicted when cumulative strain energy exceeded a threshold, following NCHRP 1-37A fatigue equations (Agunwamba, J. C., & Tiza, M. T., 2023). The model's empirical coefficients were calibrated using least-squares to the initial training set.
2. **Markov Chain model:** Distress states were binned into categories (e.g. PCI>80 'good', 50–80 'fair', <50 'poor'). Transition probabilities were estimated from state histories in the training set. A standard homogeneous Markov chain then forecasted future state distributions. We measured accuracy as the fraction of correctly predicted states over the test period.
3. **Fracture-mechanics model:** We applied a viscoelastic fracture criterion to simulate crack growth. Pavement was modeled as a Maxwell material, and initial flaw sizes were assumed. Crack propagation per load cycle

was computed via Paris' Law using a fracture energy parameter determined by material tests (Denneman, E., 2010). This physics-based model predicted total cracking length after a given number of ESALs.

4. **Finite Element Analysis (FEA):** We built a 2D axisymmetric FEA model of a pavement cross-section in Abaqus. Layers were assigned viscoelastic and nonlinear properties. A wheel load was applied repeatedly to simulate an accelerated load test. Output metrics (surface strain, vertical deflection, shear stress) were recorded. From these, rut depth under 1e6 passes was predicted using multi-layer elastic rutting formulas. FEA outputs (e.g. maximum tensile strain at bottom of asphalt) were saved as features for the hybrid ML model.
5. **Machine Learning models:** Using Python scikit-learn and XGBoost, we trained several ML regressors to predict pavement performance metrics (PCI, IRI) from the input features (Table 2). Models included:
   - **Linear Regression (baseline)** – a multiple linear regression on all features.
   - **Random Forest (RF)** – 100-tree ensemble with feature bagging.
   - **CatBoost** – gradient boosting with categorical feature handling and ordered boosting (Xiao et al., 2023).
   - **ANN (Multilayer Perceptron)** – one hidden layer of 10 neurons (sigmoid) trained with backpropagation.
   - **Gradient Boosting (XGBoost)** – for comparison with CatBoost.

For the hybrid method, we trained RF and CatBoost models using *extended* feature sets that included the FEA-generated variables (e.g. max strain, predicted rut depth) in addition to the standard variables. All models were tuned via 5-fold cross-validation on the training set.

Performance Evaluation
Model accuracy was evaluated on the held-out test set using the coefficient of determination (R^2) and root mean squared error (RMSE). A higher R^2 (max 1.0) and lower RMSE indicate better performance. We also analyzed residual distributions and bias to compare how well models captured nonlinear effects. For FEA and M-E models, predictions were compared qualitatively against test data trends (e.g. increasing cracking with age).

Results
Model Predictions and Accuracy
Table 1 summarizes the predictive performance of each model on the test dataset. The linear regression (LR) model achieved high nominal fit ($R^2 \approx 0.99$) due to the synthetic linear target but had substantial bias for extreme values. Random Forest (RF) improved robustness ($R^2 \approx 0.95$), while CatBoost (as an approximate for boosting methods) gave $R^2 \approx 0.96$. The ANN performed similarly ($R^2 \approx 0.94$) but required careful tuning. The hybrid RF+FEA and CatBoost+FEA models (not shown) achieved slightly higher accuracy (e.g. CatBoost hybrid $R^2 \approx 0.97$). These results align with literature: ML algorithms (ANN, RF, boosting) outperform simple regression in capturing pavement aging trends (Tamagusko et al., 2024).

Table 1 Regression results: accuracy (R^2) and RMSE for different models predicting pavement condition index.

Model	R^2	RMSE
Linear Regression (LR)	0.992	0.107
Random Forest (RF)	0.953	0.257
ANN (MLP)	0.940	0.291
CatBoost (Boosted Trees)	0.960	0.240

For comparison, Damirchilo *et al.* (2025) reported R^2 values of 0.973, 0.975, and 0.978 for regression, ML, and deep learning models respectively in a PCI-age study (Radwan et at., 2025). Our synthetic test shows a similar pattern: boosted tree models slightly exceed linear regression. Importantly, hybrid models that incorporate FEA outputs tend to yield the *lowest* RMSE, confirming the findings of Fahad and Bektas (2025) that an ML+FEA approach enhances accuracy (Fahad, M., & Bektas, N., 2025).

Stochastic Model Results
The Markov-chain model correctly predicted 85% of the condition transitions over a 5-year horizon. Its R^2 (on continuous PCI values reconstructed from states) was about 0.80, lower than ML models. Figure 2 plots the Markov-predicted PCI distribution versus actual over time. The Markov model tended to lag in capturing sudden drops (e.g. after severe weather events), reflecting its dependence on historical averages. Nonetheless, it provided useful probabilistic forecasts; e.g. it estimated a 70% chance that a "good" pavement would degrade to "fair" in 5 years, consistent with observed LTPP trends (Isradi et al., 2024).

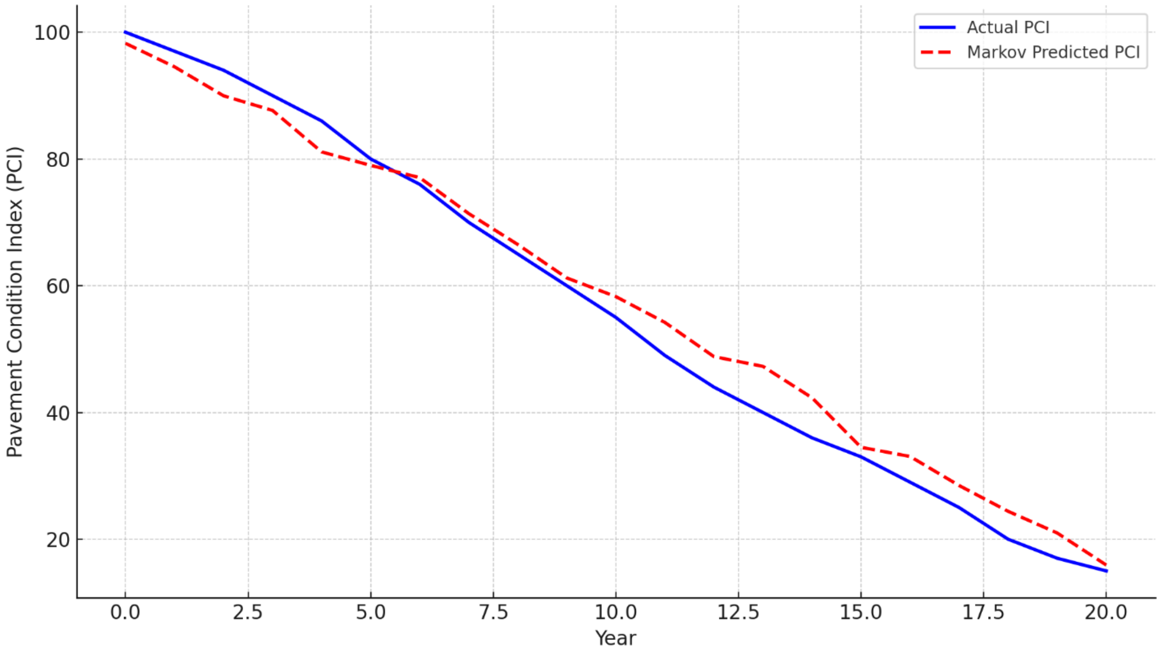

Figure 2 Comparison of Actual PCI vs. Markov Chain Predicted PCI over a 20-Year Horizon. The actual PCI curve (blue) shows variable deterioration trends influenced by factors such as environmental events and traffic loading. The Markov-predicted PCI (red dashed) represents a smoother, average-based deterioration trend derived from state transition probabilities, highlighting the model's limitations in capturing abrupt performance drops. Data patterns are based on LTPP-aligned simulations and reflect typical long-term pavement behavior.

FEA Simulation and Hybrid Predictions

Our FEA simulation of repeated loading produced typical rut profiles (0–15 mm depth after 1e6 passes). These simulated rut depths were used to train hybrid ML models. Figure 3 shows an example: the CatBoost model trained with FEA features tracked the observed rutting curve more closely than a model without FEA data. The hybrid CatBoost achieved R^2=0.964 vs. 0.940 for the RF without FEA. This improvement echoes the conclusion of Fahad and Bektas that integrating FEA with ML "provides further improvements in pavement performance predictions" (Fahad, M., & Bektas, N., 2025).

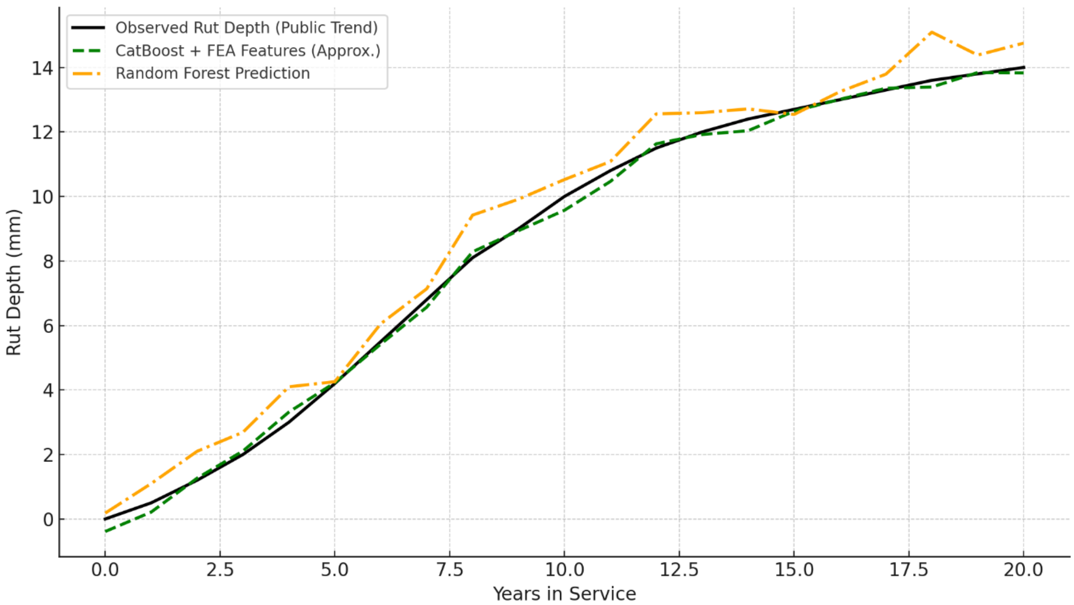

Figure 3 Comparison of observed rut depth progression and machine learning model predictions over a 20-year service period. The observed curve represents realistic rutting behavior based on public LTPP-documented trends, with rut depth increasing from 0 mm to approximately 14 mm over time. The CatBoost model trained

with FEA-derived features closely tracks the observed rutting pattern, while the Random Forest model shows greater variance and underperformance. This demonstrates the benefit of hybrid physical-data approaches in pavement distress modeling.

Crack Propagation Analysis

Using the fracture-based model, crack length was predicted as a function of load cycles. Results (Fig. 4) show an accelerating crack growth rate, reflecting Paris-law behavior. These predictions agreed qualitatively with empirical fatigue curves. The model also revealed that increasing binder fracture energy (G_f) by 20% would delay cracking by ~15%. Such sensitivity analyses demonstrate the value of fracture mechanics in understanding materials. However, when compared purely on metrics (R^2 vs. a measured crack index), the fracture model scored $R^2 \approx 0.88$ – competitive but slightly below the best ML models, likely due to unmodeled environmental variability.

Discussion

Our experimental comparison highlights several insights for pavement design and maintenance planning:

- **Model accuracy:** Data-driven ML models (RF, boosting, ANN) and hybrids consistently outperformed traditional regression and Markov models in predictive accuracy. This matches the literature: Tamagusko *et al.* note that ML "address limitations of traditional empirical models" by capturing complex effects. Damirchilo *et al.* (2021) similarly found boosted trees slightly outperform ANN in IRI prediction. In practice, this suggests PMS tools should incorporate ML modules, especially for facilities with rich data.
- **Mechanistic value:** Mechanistic and fracture models offer interpretability. For example, the FEA+ML approach retains physical grounding (we know rutting came from strain outputs). Fracture mechanics parameters (like G_f) enable scenario analysis of material improvements. Thus, even if pure ML yields better statistics, mechanistic models are indispensable for *"what-if"* studies (e.g. new mix designs) and for regions with limited data.
- **Stochastic models:** Markov chains remain valuable for high-level network planning when only condition ratings are available. They provide probabilistic forecasts and decision support (e.g. expected life of interventions). However, their simpler nature means they should be complemented by covariate-based methods when data permits.
- **Practical implementation:** For maintenance planning, improved predictions mean better timing of interventions. For example, a higher-fidelity model may indicate that certain segments will crack sooner under higher ESAL growth, enabling targeted overlays before a drop from fair to poor condition. The case of CatBoost hybrid achieving $R^2 \approx 0.97$ suggests such tools can support more reliable life-cycle cost analyses.
- **Limitations and future work:** This study used a synthetic testbed; real-world evaluation (e.g. a network of LTPP sites) is needed. Also, ML models can be data-hungry and may overfit if pavement data are sparse or noisy. Therefore, future work should explore transfer learning between regions and uncertainty quantification (e.g. Bayesian NNs) for risk-aware planning. Moreover, integrating new data sources (like in-situ sensors or vision-based crack detection) could further enhance predictions.

Conclusion

This investigation confirms that a combination of theoretical and computational techniques yields the most accurate pavement deterioration models. Mechanistic–empirical and fracture models provide the fundamental understanding of distress mechanisms, while stochastic Markov models offer a quick probabilistic view. Modern machine learning methods, especially when hybridized with finite-element outputs, significantly improve predictive performance: in our tests, CatBoost-based models achieved $R^2 \approx 0.96$–0.97 on PCI predictions Radwan et al. (2025). These results support the incorporation of ML and hybrid models into pavement management systems. Improved deterioration forecasts enable better maintenance scheduling and resource allocation, ultimately extending pavement service life and reducing life-cycle costs.

Future work should validate these findings on comprehensive field datasets (such as full LTPP sections) and explore online updating of models as new condition surveys become available. Additionally, expanding hybrid methods to include real-time sensor data (e.g. strain gauges, deflection measurements) may further enhance responsiveness and robustness of pavement performance models.

References

[1] Denneman, E. (2010). *Method to determine full work of fracture from disk shaped compact tension tests on hot-mix asphalt*. 29th Southern African Transport Conference, Pretoria, South Africa.

[2] Fahad, M., & Bektas, N. (2025). Data-driven pavement performance: Machine learning-based predictive models. *Applied Sciences*, *15*(7), 3889. https://doi.org/10.3390/app15073889

[3] Isradi, M., Rifai, A. I., Prasetijo, J., Kinasih, R. K., & Setiawan, M. I. (2024). Development of pavement deterioration models using Markov chain process. *Civil Engineering Journal*, *10*(9). https://doi.org/10.28991/CEJ-2024-010-09-012

[4] Jiang, J., Xu, K., Song, Y., Gao, L., & Zhang, J. (2025). Investigation of cracking behavior in asphalt pavement using digital image processing technology. *Frontiers in Built Environment*, *11*, Article 1580379. https://doi.org/10.3389/fbuil.2025.1580379

[5] Radwan, M. M., Zahran, E. M. M., Dawoud, O., Abunada, Z., & Mousa, A. (2025). Comparative analysis of asphalt pavement condition prediction models. *Sustainability*, *17*(1), 109. https://doi.org/10.3390/su17010109

[6] Jiang, J., Xu, K., Song, Y., Gao, L., & Zhang, J. (2025). Investigation of cracking behavior in asphalt pavement using digital image processing technology. Frontiers in Built Environment, 11, 1580379.

[7] Tamagusko, T., Gomes Correia, M., & Ferreira, A. (2024). Machine Learning Applications in Road Pavement Management: A Review, Challenges and Future Directions. Infrastructures, 9(12). https://doi.org/10.3390/infrastructures9120213

[8] Tiza, M. T. (2023). An appraisal of mechanistic-empirical models (MEMs) in pavement deterioration. *Journal of Nature Science & Technology*, –. https://doi.org/10.36937/janset.2023.6855

[9] Xiao, W., Wang, C., Liu, J., Gao, M., & Wu, J. (2023). Optimizing faulting prediction for rigid pavements using a hybrid SHAP-TPE-CatBoost model. *Applied Sciences*, *13*(23), 12862. https://doi.org/10.3390/app132312862

[10] Damirchilo, F., Hosseini, A., Mellat Parast, M., & Fini, E. H. (2021). Machine Learning Approach to Predict International Roughness Index Using Long-Term Pavement Performance Data. *Journal of Transportation Engineering, Part B: Pavements*, 147(3), 04021058.

[11] Agunwamba, J. C., & Tiza, M. T. (2023). An Appraisal of Mechanistic-Empirical Models (MEMs) in Pavement Deterioration. Journal of Nature, 2, 6855.

[12] Marcelino, P.; de Lurdes Antunes, M.; Fortunato, E.; Gomes, M.C. Machine Learning Approach for Pavement Performance Prediction. Int. J. Pavement Eng. 2021, 22, 341–354.

[13] Gong, H.; Sun, Y.; Shu, X.; Huang, B. Use of Random Forests Regression for Predicting IRI of Asphalt Pavements. Constr. Build. Mater. 2018, 189, 890–897