Proceedings of the 4th International Conference on Functional Materials and Civil Engineering
DOI: 10.54254/2755-2721/2026.38583

Physics-Constrained Machine Learning Based Long-Term Performance Prediction Method for Asphalt Pavements

Yu Huang

School of Economics and Management, Shenyang Aerospace University, Shenyang, China
huangyu02581@outlook.com

Abstract. Long-range pavement forecasting should incorporate nonlinear dependencies between traffic, structure, climate, and material aging processes without creating unrealistic degradation paths. In this work, a physics-constrained machine-learning architecture is proposed, utilizing the Federal Highway Administration Long-Term Pavement Performance (LTPP) data set. Multivariate time series consisting of annual records of International Roughness Index (IRI), rutting depth, cracking severity, traffic loads, structure characteristics, material properties, and climate factors are used for training of the two-layer Long Short-Term Memory (LSTM) neural network. The loss function consists of prediction error and soft constraints on deterioration continuity, untreated section monotonicity, cumulative traffic damage, climate aging, and physically realistic condition boundaries. The forecasting performance is examined at 5, 10, 15, and 20-year horizons for random forest, XGBoost, and unphysical LSTM methods. For 20-year forecasts, IRI, rut depth, and cracking RMSE errors were 0.284±0.031 m/km, 2.18±0.26 mm, and 5.73±0.61 percentage points, respectively. Deterioration direction violations were reduced to 2.7±0.8% and the deviation of the timing of the serviceability threshold violation was 1.34±0.29 years. The results indicate that embedding engineering deterioration knowledge improves both long-horizon accuracy and physical plausibility.

Keywords: Asphalt pavement, Physics-constrained machine learning, Long-term performance prediction, Pavement deterioration, LTPP

1. Introduction

Predictive performance of asphalt pavement conditions in the long term is important for maintenance scheduling, rehabilitation planning, and life-cycle management. Deterioration is a sequential phenomenon driven by the interaction between pavement structure, material properties, climate factors, repeated axle loads, and maintenance practices that cause sudden but rational recoveries in pavement conditions. Review of machine learning models for roughness suggests that although they can capture nonlinear associations which are hard to include in deterioration equations, generalization and stability problems are still prevalent [1]. The literature on prediction models based on LTPP database and Gaussian process or locally weighted regression demonstrates that pavement age, past distress, rutting, and structural condition are rich predictors of IRI [2].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.38583

Innovations in ensemble learning have made possible better performance and better interpretability in predicting conditions of asphalt pavements based on LTPP characteristics and feature attribution analysis [3]. Long-horizon studies suggest that machine learning can be a useful tool for pavement maintenance decision making if combined with information about traffic volume, climatic conditions, structural condition and treatments [4]. Recurrent networks are particularly appropriate for longitudinal analysis due to dependence of current pavement condition on its deterioration history, recent LTPP evidence shows that LSTM-based models provide better prediction of PCI and IRI compared to feed forward ones [5].

Nevertheless, mere accuracy is not sufficient for long-term predictions. Predictions produced by recursive neural model can fall out of feasible range, produce spurios improvements in sections without treatment, and behave inconsistently depending on traffic growth and climatic exposure. Such predictions can be consistent enough in terms of aggregate errors, but are not appropriate for engineering decisions. In this study, a physics-constrained LSTM-based model is developed. It uses known deterioration patterns as soft constraints for training process, rather than replacing data-driven learning with a mechanistic equation. The proposed framework performs joint IRI, rutting, and cracking predictions for 5-20 year horizons, distinguishes the quality of predictions in terms of accuracy and engineering consistency, using specific violation and threshold-timing measures. Such distinction is needed for pavement management since timing of pavement intervention depends not only on prediction accuracy, but also on physical consistency of the deterioration path under uncertainty.

2. Literature review

2.1. Asphalt pavement performance prediction

The study of pavement performance recently has shifted toward using heterogeneous longitudinal data as opposed to regression models under single condition. The machine learning interpretive studies indicate that roughness and rutting are regulated by interactions between traffic, structural, climatic and historical condition variables that have differing impacts within the feature space [6]. Studies conducted for LTPP confirm the existence of systematic geographic linkages among heavy trucks, climatic regions, roughness and cracking, implying that environmental and loading impacts should be considered jointly [7]. Thus, long-term prediction requires building the model that will take into account the history of the section as well as different structures and exposures. Tree models provide for an adequate representation of nonlinear relations; however, they do not account for the sequence dependency due to which cumulative impact causes the impact on future condition. This aspect becomes especially relevant when the forecast horizon expands.

2.2. Machine learning and physics-constrained modeling

Recently, machine learning in physics-informed pavement response analysis has evolved from mere statistical fitting to penalty modeling of physically unrealistic responses. Physics-informed neural networks have been employed for preserving the well-known pavement response relationships without losing flexibility in data-driven modeling [8]. Physics-guided models of pavement rutting illustrate the possibility of increasing the stability and reasonability of the predicted outcomes using engineering knowledge without losing accuracy [9]. More recent physics-informed models of LTPP-based pavement cracking clearly illustrate the benefits of incorporating deterioration rules in the learning process of deep learning algorithms for managing infrastructure assets [10]. These papers

DOI: 10.54254/2755-2721/2026.38583

serve as motivation for employing weaker engineering constraint rules in long-term field performance predictions in terms of continuity, no improvement without repair, sensitivity to loading, environmental aging effects, and realistic range of outputs for regularizing an LSTM algorithm without developing a full-fledged mechanical deterioration rule. These constraint rules are particularly appropriate for heterogeneous field databases where one deterministic rule is not applicable to all structures of pavements.

3. Physics-constrained long-term prediction method

3.1. Dataset and temporal variable construction

LTPP yearly data sets were matched based on section identification number and inspection year. Those sections with at least ten years of condition data and related traffic, structural, and climatic data were kept. This filtered data set included 438 asphalt sections and 6,914 section-year data entries. Inputs were previous year's IRI, rut depth, cracking, deflection, age, cumulative ESALs, annual increase in ESALs, asphalt and base thicknesses, asphalt modulus, mean annual temperature, precipitation, freeze index, and freeze-thaw cycle count. Maintenance years were noted instead of being considered as improvement years.

Continuous inputs were normalized using training period statistics, and those climate missing values less than one annual observation were interpolated using adjacent LTPP climate data. There were three outputs, which were next year's IRI, rut depth, and cracking representing functional performance, permanent deformation, and damage progression respectively. Sequence lengths were not fixed but depended on data availability, and the padded years in the sequence were masked in the recurrent loss function.

3.2. Physics-guided LSTM prediction model

The predictor is composed of stacked LSTM layers with 128 and 64 units, a 64 unit shared dense layer, and three separate output layers. The dynamic features from the past are fed into the recurrent encoder, whereas the invariant or slowly varying structural features are concatenated with the final hidden layer. The principal innovation is a soft-constraint objective that allows measurement variability but penalizes trajectories inconsistent with untreated pavement deterioration as shown in Equation (1):

$$L = \sum_{k=1}^{3} \frac{\alpha_k}{N} \sum_{i,t} \left(\hat{y}_{i,t}^{(k)} - y_{i,t}^{(k)}\right)^2 + \lambda_c \sum_{i,t,k} (1 - M_{i,t}) \left[\max\left(0, \left|\Delta \hat{y}_{i,t}^{(k)}\right| - c_k\right)\right]^2 \tag{1}$$

$$+\lambda_m \sum_{i,t,k} (1 - M_{i,t}) \left[\max\left(0, -\Delta \hat{y}_{i,t}^{(k)}\right)\right]^2 + \lambda_E \sum_{i,t,k} \left[\max\left(0, -\frac{\partial \hat{y}_{i,t}^{(k)}}{\partial E_{i,t}}\right)\right]^2$$

$$+\lambda_A \sum_{i,t,k} \left[\max\left(0, -\frac{\partial \hat{y}_{i,t}^{(k)}}{\partial A_{i,t}}\right)\right]^2 + \lambda_b \sum_{i,t,k} \left\{\left[\max\left(0, l_k - \hat{y}_{i,t}^{(k)}\right)\right]^2 + \left[\max\left(0, \hat{y}_{i,t}^{(k)} - u_k\right)\right]^2\right\}$$

DOI: 10.54254/2755-2721/2026.38583

Here, $M_{i,t}$ marks maintenance, $E_{i,t}$ is cumulative ESAL, $A_{i,t}$ is cumulative environmental-aging exposure, and l_k and u_k define feasible target ranges. The continuity threshold c_k is estimated from the 99th percentile of untreated annual change in the training data. Constraint weights are chosen based on validation errors and direction violations; thus, improvement in physical consistency cannot be achieved merely by means of too much smoothing. Environmental aging is a combination of cumulative positive degree-days, precipitation aging, and freeze-thaw cycles, enabling the use of a constraint that measures accumulated climate degradation rather than age alone. As illustrated in Figure 1, the proposed framework integrates multi-source pavement information with a physics-guided LSTM and recursive forecasting strategy to generate long-term deterioration trajectories that remain both data-responsive and engineering-consistent.

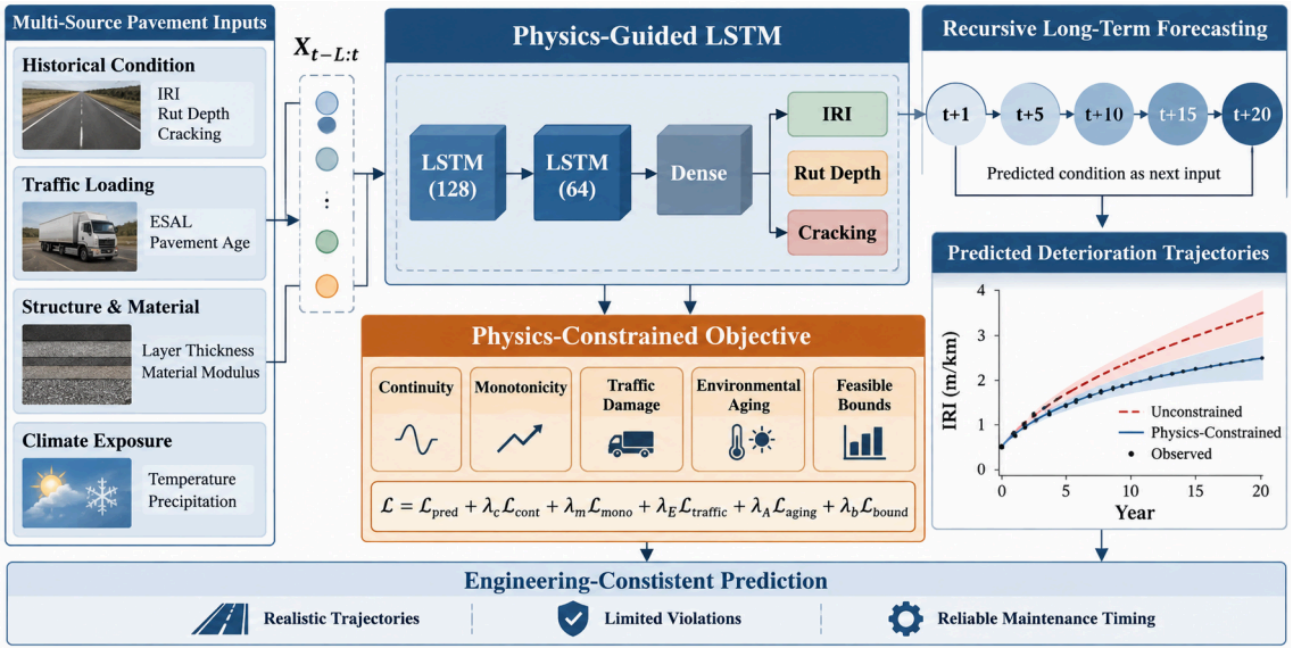

Figure 1. Physics-constrained LSTM framework for long-term asphalt pavement performance prediction

3.3. Training and long-term forecasting

Training was done using AdamW with learning rate of 0.001, batch size 64, gradient clipping 1.0, and maximum 300 epochs. For early stopping, no improvement in the validation objective for 25 consecutive epochs was needed. The method of teacher forcing was employed during one-step training; in evaluation, the predicted condition at year t+1 was fed back as input into the model for each of the subsequent years while traffic and climate covariates followed their actual trajectories in the validation period. Predictions were made at 5-year, 10-year, 15-year, and 20-year intervals. Ten seeds were kept to capture the neural model variation. Recursive predictions were stopped at the time of major rehabilitation because the post-rehabilitation state is considered as a new state of the pavement structure.

DOI: 10.54254/2755-2721/2026.38583

4. Experimental design and validation

4.1. Sample selection and comparative experiment

Segments containing significant reconstruction during an evaluation time frame were omitted unless the intervention year and type of intervention could be clearly identified. Chronological data was divided into 70%, 15%, and 15% training, validation, and testing periods for each valid segment, respectively, while 62 entire segments were withheld for use in cross-sectional testing. The Random Forest, XGBoost, and the baseline LSTM were all supplied with exactly the same predictor variables.

4.2. Multi-horizon and cross-section validation

Calculations of MAE, RMSE, and R^2 values were done separately for IRI, rutting, and cracking for every prediction horizon. Consistency of engineering calculations was checked for deterioration direction violations, frequency of unrealistic annual change in years without treatment, feasibility bound violations, and absolute error in prediction of the year of IRI crossing 2.7 m/km or cracking crossing 20%. Sections for held-out test were taken from different traffic volumes, pavement types, and climatic zones. Resampling was done via paired bootstrap with 1,000 sections' replicas, uncertainty intervals were calculated, and Holm correction was used for comparisons to the unconstrained LSTM model.

5. Results

5.1. Multi-horizon pavement performance prediction accuracy

Prediction error increased with horizon for every model, but accumulated substantially more slowly for the physics-constrained LSTM. Its IRI RMSE increased from 0.146±0.018 m/km at 5 years to 0.284±0.031 m/km at 20 years, whereas standard LSTM error increased from 0.169±0.020 to 0.372±0.044 m/km. At the 10- and 15-year horizons, the proposed model produced IRI RMSE values of 0.191±0.022 and 0.238±0.027 m/km, respectively. Twenty-year rut-depth RMSE reached 2.18±0.26 mm*, compared with 2.91±0.34 mm for standard LSTM, while cracking RMSE decreased from 7.42±0.79 to 5.73±0.61 percentage points*. Corresponding 20-year R^2 values were 0.842±0.021, 0.806±0.028, and 0.781±0.032 for IRI, rutting, and cracking (see Table 1).

Table 1. Twenty-year prediction accuracy and engineering consistency

Model	IRI RMSE (m/km)	Rut RMSE (mm)	Crack RMSE (p.p.)	Direction Violations (%)	Threshold-Year Error (yr)
Random Forest	0.438±0.052	3.64±0.41	9.31±0.96	14.8±2.3	3.42±0.58
XGBoost	0.401±0.047	3.31±0.38	8.56±0.88	12.9±2.0	3.06±0.51
Standard LSTM	0.372±0.044	2.91±0.34	7.42±0.79	10.7±1.8	2.63±0.46
Physics-Constrained LSTM	0.284±0.031*	2.18±0.26*	5.73±0.61*	2.7±0.8*	1.34±0.29*

** $p<0.01$ versus standard LSTM after Holm adjustment.*

DOI: 10.54254/2755-2721/2026.38583

5.2. Effect of physics constraints on long-term stability

The unconstrained LSTM generated 10.7±1.8% deterioration-direction violations and 6.2±1.4% unrealistic annual changes at 20 years. Physics constraints reduced these values to 2.7±0.8%* and 1.5±0.6%*, respectively. Removing the monotonicity term increased direction violations to 7.9±1.5%, demonstrating that accuracy gains alone did not account for the improved trajectories. Removing traffic- and aging-sensitivity penalties increased 20-year IRI RMSE to 0.319±0.036 m/km and cracking RMSE to 6.41±0.69 percentage points. Feasible-bound violations declined from 3.8±1.0% for standard LSTM to 0.4±0.2%*. The complete constraint set therefore provided complementary control over direction, rate, and absolute deterioration magnitude.

5.3. Cross-section generalization and engineering consistency

For the 62 unseen sections, the suggested model showed an IRI RMSE value of 0.307±0.039 m/km, a rut depth RMSE value of 2.39±0.31 mm, and cracking RMSE value of 6.08±0.72 percentage points. Error in threshold years was still 1.51±0.33 years* in comparison with 2.88±0.54 years of standard LSTM. Drop-off in performance from the normal test set to unseen sections for IRI RMSE was 8.1±2.7%, which is much less than the drop-off of 18.6±4.1% in the case of the unconstrained network. Direction violations for unseen sections were 3.2±0.9%*, whereas for the standard LSTM, direction violations were 11.9±2.1%. These results indicate that the physics-constrained model learned deterioration patterns that transferred across sections rather than merely reproducing section-specific historical trajectories.

6. Conclusion

Asphalt road prediction through the physics-constrained LSTM model achieves improved long-term accuracy by merging engineering regularization with sequence learning. In the range from 5 to 20 years ahead, it demonstrated smaller errors for IRI, rutting, and cracking than tree-based models and LSTM models lacking constraints, while significantly minimizing the likelihood of self-improving road condition, unrealistic changes year by year, and violations of feasible range. Physical constraints have been also shown effective in controlling long-horizon drift and in estimating the serviceability threshold value on the cross-section test.

References

[1] Tamagusko, T., & Ferreira, A. (2023). Machine learning for prediction of the International Roughness Index on flexible pavements: A review, challenges, and future directions. *Infrastructures, 8*(12), 170.

[2] Kaloop, M. R., El-Badawy, S. M., Hu, J. W., & Abd El-Hakim, R. T. (2023). International Roughness Index prediction for flexible pavements using novel machine learning techniques. *Engineering Applications of Artificial Intelligence, 122*, 106007.

[3] Yang, S., Wang, Y. D., Hu, X., & Liu, J. (2022). An efficient and explainable ensemble learning model for asphalt pavement condition prediction based on LTPP dataset. *IEEE Transactions on Intelligent Transportation Systems, 23*(11), 22084–22093.

[4] Mansour, E., Dhasmana, H., Mousa, M. R., & Hassan, M. M. (2024). Machine learning-based technology for asphalt concrete pavement performance decision-making in hot and humid climates. *Construction and Building Materials, 442*, 137625.

[5] Chen, L., Li, H., Wang, S., Shan, F., Han, Y., & Zhong, G. (2025). Improved model for pavement performance prediction based on recurrent neural network using LTPP database. *International Journal of Transportation Science and Technology, 19*, 128–138.

DOI: 10.54254/2755-2721/2026.38583

[6] Wu, J., Chen, J., Liu, Q., & Ma, X. (2024). An interpretation framework of machine learning models: Prediction of pavement long-term performance. *International Journal of Pavement Engineering, 25*(1), 2387737.
[7] Kargah-Ostadi, N., Vasylevskyi, K., Ablets, A., & Drach, A. (2024). Physics-informed neural networks to advance pavement engineering and management. *Road Materials and Pavement Design, 25*(11), 2382–2403.
[8] Deng, Y., Wang, H., & Shi, X. (2024). Physics-guided neural network for predicting asphalt mixture rutting with balanced accuracy, stability and rationality. *Neural Networks, 172*, 106085.
[9] Pasupunuri, S. K., Thom, N., & Li, L. (2024). Roughness prediction of jointed plain concrete pavement using physics informed neural networks. *Transportation Research Record, 2678*(11), 1733–1746.
[10] Taheri, A., Sobanjo, J., & Elwardany, M. (2026). Pavement cracking prediction models based on deep learning physics-informed neural network. *Transportation Research Record, 2680*(4), 241–260.