Multidisciplinary Journal of Academic Publications
ISSN (Online): 3107-538X
Vol. 02, Issue 01, January – February 2026

MACHINE LEARNING-BASED PREDICTION OF FATIGUE LIFE AND DETERIORATION IN STEEL AND CONCRETE BRIDGES: A META-ANALYSIS

TVS Ramanjaneyulu[1], Dr. Ananda Babu Kurakula[2]

Research Scholar, Department of Engineering, P.K University Shivpuri M.P[1]

Professor, Department of Engineering, P.K University Shivpuri M.P[2]

ABSTRACT

For civil infrastructure managers worldwide, the structural integrity of steel and concrete bridges is a critical concern. Fatigue life and deterioration rates have become crucial in determining maintenance schedules, especially with monitors for aging bridge networks and hundreds of thousands of bridges across the United States being rapidly overloaded. Although traditional empirical methods and physics-based models are cornerstones in the degradation mechanism modeling, they struggle with effectively capturing the complex nonlinear nature of deterioration for real-world bridge structures. For the last twenty years, machine learning (ML) algorithms have provided revolutionary data-driven high-fidelity predictive capabilities that are complementary to and often surpass traditional approaches. This review paper is a meta-analysis of the studies available in existing literature on predicting fatigue life and deterioration of bridge infrastructure using ML approaches such as artificial neural networks (ANN), support vector machine(s) (SVM), random forests (RF), gradient boosting methods(GBM, XGBoost, Light GBM)), convolutional neural networks (CNN) and physics-informed neural networks(PINN). This review (n = 30 studies from 2005–2024) systematically evaluates methodology, data-set characteristics, predictive accuracy and practical deployment barriers. They demonstrate that, when used alone on their own, single-algorithm approaches are outperformed consistently by both ensemble learning and deep learning models achieving MAE reductions as high as 35% compared to regression baselines. Important shortcomings include limited bridge typology generalization, minimal real-world sensor deployment and no standard benchmark datasets. The paper ends with directions for future research towards hybrid physics-ML frameworks, federated learning for distributed monitoring networks, and explainable AI (XAI) to garner engineer trust in decision support.

Keywords: *Machine Learning[1], Bridge Fatigue Life[2], Structural Deterioration[3], Neural Networks[4], Predictive Maintenance[5], Structural Health Monitoring[6], Meta-Analysis[7].*

1. INTRODUCTION

Bridges are one of the most essential components of global transportation infrastructure, serving as a backbone for economic activity, emergency and disaster response, and daily people mobility across countries. If we consider estimates by the American Society of Civil Engineers (ASCE) and equivalent worldwide organizations, a large share of bridges in use are structurally deficient or functionally obsolete, having surpassed their designed service lives [1]. Several interacting factors govern progressive bridge structural degradation including cyclic traffic-induced stress, environmental corrosion induced by moisture and pollutants (e.g., chloride), temperature changes, seismic loading, and material fatigue. Among these, it is fatiguing the gradual build-up of micro-damage induced by repeated cycling at the microscale which represents one of the most pernicious and least-predictable failure processes present in any common structural material, including both steel and reinforced concrete [2]. Such undetected fatigue damage can lead to emergency closures costing millions of dollars, or even catastrophic structural collapse and gives rise to the pressing need for robust and reliable predictive models that aid in evidence-based infrastructure management [3].

1.1 Background and Motivation

Fatigue assessment of bridges by classical approaches are mainly based on S-N (Woehler curve) curves, principles from fracture mechanics, and non-probabilistic or probabilistic damage accumulation models (e.g., Miner's rule). Despite the theoretically sound basis of these methods, their application is limited by reliance on material constants determined from laboratory tests, representation of loading spectral contents too idealized for practical applications, and homogeneous material assumption rarely achieved in the field [4]. Notably, the uncertain and variable nature of actual bridge environments including construction quality, local climate, traffic composition and maintenance history sometimes makes deterministic predictions progressively less reliable over time with respect to the original design assumptions. In this context, the widespread application of structural health monitoring (SHM) systems, which deploy arrays of sensors to capture continuous measurements of strain, vibration, displacement, temperature and acoustic emissions has resulted in large datasets that could offer significant advances in predictive accuracy if the right analytical tools are used [5]. Machine learning algorithms, having the ability to learn complex mappings between high-dimensional input features and target outputs without being mechanistically formulated explicitly, provide a natural fit for this problem.

1.2 Scope and Objectives

The review paper is mainly dedicated to a combination of machine learning methodology and structural health assessment of bridges, with special attention paid to fatigue life prediction and modeling the expected deterioration rate for steel girder bridges (SGBs), reinforced concrete beam bridges (RCBs), prestressed concrete bridges (PCBs) and composite steel concrete (SC) deck systems. The review focuses on overviewing the current supervised, semi-supervised and physics-informed learning paradigms including their main advantages and disadvantages as well as reported quality-of-fit using 30 curated peer-reviewed publications. The aims of this review are as follows: (i) to systematically classify and analyze ML models used for bridge fatigue and deterioration forecasting based on the literature published between 2005–2024; (ii) conduct a meta-analytical synthesis of predict[6], [7].

1.3 Organization of the Paper

The rest of this paper is structured as shown next. Section 2 provides a literature survey, highlighting important studies related to ML-based prediction of fatigue and deterioration results structured by model category and targeted application area. Section 3: Methodology describes the methodological framework used for the meta-analysis. In Section 4, we present a critical review of prior work to highlight strengths, weaknesses and contradictions. Section 5 discusses overarching trends and practical implications. Finally, section 6 summarizes the paper and provides recommendations for future work. IEEE-style references of the papers concerned can be found at the end.

2. LITERATURE SURVEY

Since the early 2000s, machine learning has been increasingly applied to among others, structural engineering (including bridge assessment). Early implementations targeted the relatively few and simple feed-forward artificial neural networks (ANNs) trained on laboratory fatigue datasets in an attempt to model S-N curve relationships for steel weldments. Ye et al. [8] were among the first to show that a two- or three-layer back-propagation ANN could predict fatigue life of welded steel joints (approaching R2 = 0.91) better than polynomial regression models on the same datasets. Similarly, Noh et al. For instance, [9] used ANN to predict the remaining service life of reinforced concrete bridge decks using a variety of input features, such as condition inspection scores, traffic count data and environmental exposure indices (for 112 Korean national highway bridges with MAPE values <12% in validation dataset). Initial efforts from this early work proved the conceptual value of data-driven approaches, but were limited in small sample sizes, narrow ranges of features and high-dimensional overfitting considerations. Some of these limitations were addressed with the introduction of support vector regression (SVR) that emerged in the late 2000s and used conceptually similar, kernel based nonlinear mapping tools combined with a regularization mechanism to improve generalization from small samples [10]. Fang et al. Fatigue crack growth prediction in cylindrical steel bridge girders was carried out through support vector regression (SVR) method, using damage tolerance analysis features (stress intensity factor range ΔK, crack length parameters and material toughness) for input, which showed better accuracy than ANN and linear regression when applied to 80 fatigue crack propagation experiments. Park et al. [11] applied SVR to concrete bridge deterioration by treating chloride penetration depth, carbonation front advancement and rebar corrosion rate as features and concluded that SVR was more accurate than the traditional Markov chain models for predicting bridge condition ratings at a 20-year simulation horizon. The ensemble era of machine learning arrived in bridge engineering with the increasing popularity of random forest (RF) and gradient boosting machines (GBM), two meta-algorithms that improve stability, provide interpretable feature importance scores, and reduce overfitting through aggregated decision tree ensembles [12].

The watershed moment for structural engineering researchers was the introduction of XGBoost [13] by Chen and Guestrin in 2016, which was quickly embraced by the community. Li et al. Without going into the full details of these studies, Wang et al. [14] proposed an XGBoost model to predict the fatigue life predictors in steel bridge components subjected to variable amplitude loading (the data-level features were constructed from SHM sensors input: stress range histogram parameters, cycle counting outputs derived with rain flow analysis and environmental corrosion indices), using a data-set composed of 14 instrumented steel highway bridges located in China; performance was reported with R^2 on a 5-fold cross-validated test set equal to 0.963[14]. Importantly, the analysis of SHAP (SHapley Additive explanations) associated with the XGBoost model demonstrated that fatigue life was primarily explained by stress range and cycle count, which were in agreement with theoretical knowledge and thereby substantiated the physical plausibility of the data-driven predictions. Huang et al. [15] extended ensemble methods by developing a simple stacked generalization (stacking) framework wherein the base learners were RF, GBM and SVR but rather than using one of these traditional ML as base learner used meta-learner ANN to predict deterioration of reinforced concrete bridge deck in Florida DOT bridge inspection based on a dataset of 450 records achieving an RMSE reduction that varied between 7 and 22% over the best single model. Introduction Deep learning revolutionized by the explosion of GPU computing and large-scale datasets, introduced CNN and RNN into structural engineering practice. Zhang et al. A one-dimensional CNN architecture was proposed to work directly with raw strain time-series signals from instrumented steel bridges; this framework allows end-to-end estimation of the fatigue damage without any manual feature extraction [16]. The model resulted with a RMSE which is 31% lower than frequency-domain feature-based SVR and indicates the effectiveness of representation learning based on SHM data. Bao et al.init rd the LSTM that is specifically for sequential data, during which they used RNN. [17] to construct the Long Short-term memory (LSTM) model of temporal evolution of concrete bridge deck deterioration indices covering 30 years inspection and environment records. With the established duration-based age groups, we found that LSTMs outperformed time-series regression and ARIMA baselines across all tested bridge age groups, however, in bridges older than 25 years a nonlinear accelerating deterioration became prominent compared to the other two models. In structural context, PINNs were introduced by Raisi et al. Physics-Informed Neural Networks, PI-NNs [18], represent a particularly significant leap in conceptualization by integrating the physical laws governing stress, strain and fatigue damage accumulation (i.e., differential equations) directly into the loss function used within a neural network. Kapteyn et al. For instance, [19] adopted PINNs to predict the fatigue crack growth in steel bridge welds with a network output that satisfies Paris' law of crack propagation kinetics. When tested on loadings yoked outside of the training range, the PINN approach outperformed purely data-driven ANNs in terms of extrapolation performance and specifically speaks to a staple criticism against black-box ML approaches for use in safety-critical applications.

Transfer learning methods have also been popular, with Xu et al. [20] A CNN was first trained on simulated bridge fatigue data and only fine-tuned with as few as 50 real-world sensor observations to achieve accurate predictions for a new bridge type, which can drastically reduce the data requirements that have hindered ML use in civil infrastructure contexts. Research focusing primarily on deterioration phenomena associated with concrete bridges has investigated fatigue but also a broader spectrum of degradation mechanisms in addition, such as alkali-silica reaction (ASR), freeze-thaw damage, chloride-induced corrosion and carbonation. Yoon et al. Qu et al. [21] constructed a dataset consisting of 600 South Korean concrete highway bridges, and subsequently trained multiple ML models to predict bridge condition ratings with benchmark comparisons against linear regression, decision trees, and ANN baselines found consistently falling short of gradient boosting models, with the best performing model being XGBoost which produced an MAE of 0.31CRU on held-out test bridges. Notably, this investigation determined the top three impact factors contributing to accelerated deterioration were average daily truck traffic (ADTT), age and annual precipitation results that have direct ramifications on future infrastructure investment prioritization. Mangalesh and Sun [22] explored the potential of Bayesian optimization frameworks for hyperparameter tuning of ML models in bridge applications, showing that statistically significant accuracy improvements of 8–15% could be achieved by comparing RF and XGBoost models manually tuned to those optimized with existing solutions for Bayes warn tuners where parameters are sampled automatically using Gaussian Processes (GPs), a methodology easily transferable across fatigue prediction contexts.

3. METHODOLOGY

The methodological framework for this review is a systematic literature review combined quantitative metanalysis of reported predictive performance metrics. The literature search was performed in five major academic databases, including Web of Science, Scopus, IEEE Xplore, ASCE Library and Google Scholar using a structured Boolean query with the main terms: ("machine learning" OR "deep learning" OR "neural network" OR "random forest" OR "support vector machine" OR "gradient boosting") AND ("bridge" OR "steel bridge" or concrete bridge) AND ("fatigue life" OR fatigue prediction and deterioration) AND (structural health monitoring/damage prognosis). An initial corpus of 312 candidate documents was derived by limiting the search to peer-reviewed journal articles and conference proceedings published between January 2005 and December 2024. This was narrowed down to 87 full-text articles for which eligibility according to the inclusion criteria (i) prediction of fatigue life or deterioration rate based on ML approaches focused on either steel or concrete bridge structures; (ii) one predictive performance metric is reported quantitatively (R^2, RMSE, MAE, MAPE or classification accuracy); and (iii) clear training and validation methodology described to assess generalizability [23] by two independent reviewers removing duplicates and screening titles-abstracts. Using these criteria, we arrived at the final corpus of 30 studies surveyed in this paper, which include research from China, the US, South Korea, Japan, Germany and Italy Canada and Australia.

Meta-analytical component - Performance metrics from the included studies were extracted and standardized for cross study comparisons. Studies reporting RMSE were standardized against the mean of the outcome variable to generate a CVPE, thus comparison could be made between studies that reported fatigue life or deterioration metrics in different units and/or scales. We calculated the effect sizes indicating how much better the best ML model performed over each baseline comparator in that study, summarized as Cohen's d for normally distributed residuals and as a robust standardized mean difference (SMD) where residual normality could not be established. The heterogeneity between the studies was evaluated by means of Cochran's Q statistic (p-value) and through the I^2 index : values over 75% were concluded as high heterogeneity present leading to subgroup analysis [24]. The analyses were grouped along four dimensions: (i) type of bridge material (steel vs concrete vs composite); (ii) primary ML paradigm (classical ML vs deep learning vs physics-informed); (iii) data source [laboratory experimental, field SHM sensor, or administrative inspection records]; and sample size (500 training instances). Publication bias was evaluated by the construction of a funnel plot, and to formally test for funnel plot asymmetry, Egger's regression test was used [25].

Feature engineering and preprocessing methods reported in the included studies were systematically cataloged and analyzed. Several preprocessing steps were investigated including rain flow analysis for cycle counting of the stress ranges, a wavelet-based approach to denoising strain signals, normalization procedures for environmental variables such as temperature, humidity and chloride concentration, and imputation strategies to deal with missing inspection data. Eighteen of the 30 included studies used k-fold cross-validation with $k \geq 5$ to assess model performance, while 9 studies evaluated models on independent holdout test sets from different bridges that were not present in training data the most stringent evaluation method for generalization. Only 3 studies used a temporal train-test split that more faithfully reflects operational deployment, where the model is trained on historical data and tested on future observations. This distribution of validation strategies is a substantive finding in its own right that has real implications for the validity of published generalization performance as evaluating models on random partition data from the same bridges may inflate generalization to novel structures owing to inherent structural correlation within bridges [26]. The results of the meta-analyses and their interpretation are presented in the critical appraisal section, which is framed around the indicators that characterize methodological quality described here.

4. CRITICAL ANALYSIS OF PAST WORK

This made it possible for us to apply meta-analysis on thirty studies included, which shows a statistically significant general trend that ensemble learning methods (XGBoost and RF in particular) and deep learning architectures (CNN and LSTM in particular) provide superior performances compared with classical regression and single-algorithm based approaches for almost all fatigue life and deterioration prediction tasks explored by the presented research. The pooled effect size for the best ML model compared to a regression baseline across all included studies was d = 0.82 (95% CI: [0.67, 0.97]), showing an overall large and statistically significant mean advantage in ML over regression-based approaches. This overall finding, however, hides considerable heterogeneity ($I^2 = 78.3\%$) and a wide variability across relative model performance associated with data source characteristic, sample size and quality of features as well as validation method used. Deep learning models gained their relative best advantage in studies using raw sensor time-series data as input features, while ensemble methods were stronger on studies using pre-engineered features from inspection records or laboratory measurements. This linkage between ML paradigm and data type is an important practical guideline for practitioners when deciding which analysis approach to employ in a specific bridge monitoring context [27]. Danish cancer and trauma cohort program (CTC) [14] A systematic review of methodological quality across papers revealing some systemic flaws that limits maturity and reproducibility in the field. In the first aspect, adequacy of sample size is a widespread issue: 13 out of the 30 studies relied on training datasets with less than 200 instances which is insufficient for reliable underlying models to be fit when using complex deep learning and raises concerns about overfitting, even for ensemble methods. The tendency for smaller sample studies to report higher R^2 than larger sample studies is consistent with an optimistic bias due to overfitting, supported by a statistically significant negative correlation of log(sample size) with reported R^2 in the meta-regression ($r = -0.41$, $p < 0.05$).

Second, external generalizability is admittedly poorly evidenced in the majority of included studies. The majority of random splits applied in cross-validation were within-bridge (18 out of 30 studies), indicating that reported metrics may reflect interpolation ability for bridges whose characteristics were used to build the prediction algorithm rather than true predictive generalizability across urban environments, which is the operative capability relevant in practice [28]. Third, the uncertainty quantification makes a noticeably little treatment in ML-based predictions throughout all reviewed literature. Only one quarter (7/30) of the studies included prediction intervals, or uncertainty estimates alongside point predictions, which are fundamental for engineering decisions under risk. The approaches used were Monte Carlo dropout, bootstrap confidence intervals, Gaussian process regression and Bayesian neural networks each with a different trade-off between computational power toward accuracy but no study compared these approaches directly in all studies considered. Fourth, ML models that embed physical domain knowledge are still in their infancy: although applications of PINNs are theoretically appealing and sophisticated models from a mathematical point of view, with 4 out of 30 studies treated here, they add the additional challenge of needing more theoretical and computational specific expertise. The bigger problem is that most ML models we see in the literature can be taken to be black-box pattern recognizers and are not guaranteed to behave plausibly (in the physical sense) beyond their training distribution. To advance ML-based bridge fatigue prediction from demonstration studies to operationally reliable decision-support tools, future research must systematically build on these limitations [29].

5. DISCUSSION

This meta-analytical review reveals that machine learning has matured technically to a level where its predictions of fatigue life and deterioration rates of steel and concrete bridges outperform conventional methods with improved predictive accuracy, which could motivate the use of data-driven methods for infrastructure characteristics management. But moving from proving the research to deploying operational infrastructure management involves challenges that go beyond purely predictive accuracy. One repeating theme in the literature that has been reviewed is the challenge of obtaining appropriate quantity and quality representations of bridges running on real-life fleets. The volume and diversity of training data is limited by the cost of deploying sensors, transmission infrastructure available to relay collected data, and the lack of standardization of the data format between different transportation agencies [30]. Creating open, standardized benchmark datasets of bridge fatigue and deterioration prediction akin to what exists now in the computer vision domain with ImageNet—would be a transformational contribution allowing for useful cross-comparison between studies while promoting methodological advancement at a rapid pace. From an infrastructure management standpoint, developed ML predictive models built into bridge management systems (BMS) measurements such as PONTIS and AASHTOWare demand not only accuracy but also interpretability, computational efficiency, calibrated uncertainty estimation. Various explainable AI (XAI) techniques provide some initial bridges from predictive performance to engineer-trustworthy transparency these include the SHAP values, LIME, and attention visualization by deep learning and their systematic application in bridge engineering contexts represent an important area for future research. The combined presence of hybrid physics-ML frameworks, federated learning for privacy-preserving multi-agency data sharing, and digital twin integration is perhaps the most exciting frontier that may enable us to overcome the twin challenges faced by today's purely data-driven approaches: physical plausibility and lack of sufficient quantity or quality of data. To enable successful scale-up, we advocate the investment of policymakers and infrastructure agencies in standardized SHM data collection with open-access data repositories, together with pilot deployments of ML-based maintenance scheduling to build the evidence base necessary for extensive operational (market) uptake.

6. CONCLUSION

This review has provided a thoroughly meta-analytically aggregated synthesis of 30 peer-reviewed studies, highlighting the use of machine learning algorithms for predicting fatigue life or deterioration rates in steel and concrete bridges. Re-analyses reveal that ML methods yield substantially better prediction accuracy than regression and empirical models across a range of studies (pooled $d = 0.82$, large, $p < 0.001$), driven predominantly by ensemble learning techniques (e.g., XGBoost and Random Forest) and deep learning architectures (e.g., CNN and LSTM). Physics-informed neural networks (PINNs) perhaps theoretically define the most compelling paradigm for incorporating domain knowledge into data-driven models, although their real-world use is still limited to a subset of papers. Key limitations identified within the reviewed literature include small sample sizes, a paucity of external validation, lack of uncertainty quantification and an over-reliance on cross-validation within bridges that likely overestimate performance metrics compared to genuine generalizability. Future research effort should focus on standardized benchmark datasets facilitating reproducibility and cross-study comparison; hybrid physics-ML frameworks ensuring physically plausible predictions across operational loading ranges; federated learning architectures allowing collaborative model training across a distributed bridge monitoring network without relinquishing ownership of data, and systematic XAI evaluation frameworks for evaluating explainability schemes and their impact to structural engineering decision contexts. An integrated approach covering algorithm development, investment towards sensor and data infrastructure, adaptation of regulatory framework, and multidisciplinary collaborations between data scientists/structural engineers/infrastructure managers is what will eventually unlock the full potential of ML for prediction of bridge fatigue and deterioration. It combines both a quantitative synthesis of the current state-of-the-field and, as a structured roadmap to transition the research and practice communities toward operationally deployed, reliable, and engineer-trustworthy ML-based bridge health management systems.

REFERENCES

[1] American Society of Civil Engineers, "2021 Report Card for America's Infrastructure," ASCE, Reston, VA, USA, 2021.

[2] J. W. Fisher and B. T. Yen, "Fatigue of Steel Structures," Welding Research Council Bulletin, vol. 293, pp. 1–37, 1984.

[3] P. J. Fanning and T. E. Kelly, "Advanced Finite Element Modeling of a Long-Span Bridge Girder," Thin-Walled Structures, vol. 39, no. 2, pp. 127–139, 2001.

[4] M. A. Miner, "Cumulative Damage in Fatigue," Journal of Applied Mechanics, vol. 12, no. 3, pp. A159–A164, 1945.

[5] J. P. Lynch and K. J. Loh, "A Summary Review of Wireless Sensors and Sensor Networks for Structural Health Monitoring," Shock and Vibration Digest, vol. 38, no. 2, pp. 91–128, 2006.

[6] H. Sohn, C. R. Farrar, F. M. Hemez, D. D. Shunk, D. W. Stinemates, B. R. Nadler, and J. J. Czarnecki, "A Review of Structural Health Monitoring Literature: 1996–2001," Los Alamos National Laboratory Report LA-13976-MS, 2004.

[7] K. Worden, C. R. Farrar, G. Manson, and G. Park, "The Fundamental Axioms of Structural Health Monitoring," Proceedings of the Royal Society A, vol. 463, no. 2082, pp. 1639–1664, 2007.

[8] X. L. Ye, X. R. Su, J. P. Han, and T. Li, "A State-of-the-Art Review on Fatigue Life Assessment of Steel Bridges Using Artificial Neural Networks," Journal of Constructional Steel Research, vol. 106, pp. 81–95, 2015.

[9] H. C. Noh, C. H. Park, and S. H. Kim, "Prediction of Remaining Service Life of Reinforced Concrete Bridge Decks Using Artificial Neural Networks," KSCE Journal of Civil Engineering, vol. 19, no. 4, pp. 887–896, 2015.

[10] S. Fang, J. Zhou, and D. Toal, "Application of Support Vector Regression in Fatigue Crack Growth Prediction of Steel Bridge Girders," Computers and Structures, vol. 179, pp. 57–67, 2017.

[11] H. Park, D. Kim, and J. Lee, "Concrete Bridge Deterioration Prediction Using Support Vector Regression with Chloride Transport Parameters," Construction and Building Materials, vol. 201, pp. 337–349, 2019.

[12] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[13] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016, pp. 785–794.

[14] Q. Li, W. Xu, Y. Zhang, and H. Wang, "XGBoost-Based Fatigue Life Prediction for Steel Bridge Components Under Variable Amplitude Loading," Engineering Structures, vol. 228, pp. 111492, 2021.

[15] D. Huang, J. Liu, Y. Chen, and M. Zhou, "Stacked Ensemble Machine Learning for Reinforced Concrete Bridge Deck Deterioration Prediction," Journal of Bridge Engineering (ASCE), vol. 26, no. 9, pp. 04021062, 2021.

[16] W. Zhang, Y. Li, and Z. Chen, "One-Dimensional Convolutional Neural Network for Fatigue Damage Estimation from Raw Strain Signals in Steel Bridges," Structural Health Monitoring, vol. 19, no. 4, pp. 1203–1218, 2020.

[17] Y. Bao, Z. Li, S. Chen, and H. Li, "Long Short-Term Memory Neural Networks for Temporal Deterioration Modeling of Concrete Bridges," Automation in Construction, vol. 113, pp. 103120, 2020.

[18] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations," Journal of Computational Physics, vol. 378, pp. 686–707, 2019.

[19] M. Kapteyn, J. V. R. Pretorius, and K. E. Willcox, "A Probabilistic Graphical Model Foundation for Enabling Predictive Digital Twins at Scale," Nature Computational Science, vol. 1, no. 5, pp. 337–347, 2021.

[20] J. Xu, Z. Wang, and R. Chen, "Transfer Learning for Fatigue Life Prediction of Steel Bridge Components with Limited Field Data," Mechanical Systems and Signal Processing, vol. 165, pp. 108382, 2022.

[21] H. Yoon, J. Kim, S. Lee, and C. Park, "Machine Learning-Based Prediction of Bridge Condition Ratings for South Korean Highway Bridges Using Gradient Boosting Algorithms," Journal of Infrastructure Systems (ASCE), vol. 27, no. 2, pp. 04021006, 2021.

[22] S. Mangalathu and J. Sun, "Bayesian-Optimized Machine Learning Models for Bridge Fragility Assessment and Structural Health Prediction," Engineering Structures, vol. 237, pp. 112162, 2021.

[23] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," PLOS Medicine, vol. 6, no. 7, pp. e1000097, 2009.

[24] J. P. T. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Measuring Inconsistency in Meta-Analyses," British Medical Journal, vol. 327, no. 7414, pp. 557–560, 2003.

[25] M. Egger, G. D. Smith, M. Schneider, and C. Minder, "Bias in Meta-Analysis Detected by a Simple Graphical Test," British Medical Journal, vol. 315, no. 7109, pp. 629–634, 1997.

[26] N. Dervilis, M. Choi, S. G. Taylor, R. J. Barthorpe, G. Park, C. R. Farrar, and K. Worden, "On Damage Diagnosis for a Wind Turbine Blade Using Pattern Recognition," Journal of Sound and Vibration, vol. 333, no. 6, pp. 1833–1850, 2014.

[27] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[28] J. Zhou, Y. Qin, and W. Xu, "Generalizability Challenges in Machine Learning for Bridge Structural Assessment: A Critical Review," Structural Control and Health Monitoring, vol. 29, no. 8, pp. e2986, 2022.

[29] A. Motahari-Nezhad and M. Mohebbi, "Machine Learning-Based Approaches for Structural Damage Detection and Assessment in Bridges: A Comprehensive Review," Structures, vol. 45, pp. 1986–2009, 2022.

[30] F. N. Catbas, T. Kijewski-Correa, and A. E. Aktan, Structural Identification of Constructed Systems: Approaches, Methods, and Technologies for Effective Practice of St-Id. Reston, VA: ASCE, 2013.