International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 01 | Jan 2026 www.irjet.net p-ISSN: 2395-0072

Physics-Informed Bayesian Neural Networks for Probabilistic Load Rating of Bridges with Sparse Data

Dam Minh Hung [1], Mac Van Ha[2], Nguyen Huu Hung [2]

[1] *Khanhhoa Transportation Projects Management Unit, Khanhhoa, Vietnam*
[2] *University of Transport and Communications, Hanoi, Vietnam*

Abstract - *Accurate load rating is essential for the effective management of aging bridge inventories; yet traditional line-girder analysis methods often yield overly conservative results. While refined Finite Element Analysis (FEA) offers accuracy, it is computationally expensive. Recent studies have demonstrated the potential of Artificial Neural Networks (ANNs) to predict refined load ratings; however, these approaches typically suffer from two critical limitations: (1) the lack of uncertainty quantification, and (2) poor generalization in regions with sparse training data (data gaps). This study proposes a novel Physics-Informed Bayesian Neural Network (PI-BNN) framework. The proposed model utilizes Bayesian Regularization to quantify both aleatoric and epistemic uncertainties, providing risk-informed confidence intervals. Furthermore, to address the challenge of missing data (specifically investigated in the 30m-40m span range), the framework integrates governing structural laws (*$RF \propto L^{-2}$*) into the learning process. The model was trained on a hybrid database of synthetic and field-verified bridges. Results indicate that while standard ANNs fail to extrapolate within data gaps, the PI-BNN successfully recovers the underlying physical trend, achieving high accuracy (*$R^2 > 0.98$*) and providing a reliable "fail-safe" mechanism for decision-making.*

Key Words: **Bridge Load Rating; Physics-Informed Neural Networks; Bayesian Deep Learning; Uncertainty Quantification; Data Scarcity**

1. INTRODUCTION

The preservation of highway infrastructure is a critical challenge for transportation agencies globally, as a significant portion of the bridge inventory approaches or exceeds its design service life. Accurate load rating—the determination of the live load carrying capacity—is paramount for making informed decisions regarding load posting, structural rehabilitation, or replacement [1, 2]. Current engineering practice typically relies on simplified one-dimensional (1D) line-girder analyses prescribed by standard design codes, such as the AASHTO Manual for Bridge Evaluation. While these methods are computationally efficient, they are inherently conservative as they often neglect three-dimensional (3D) system effects, including transverse load distribution, composite action, and the contributions of secondary members (e.g., barriers and diaphragms) [1, 5]. Consequently, many bridges are deemed structurally deficient or load-restricted when they possess significant reserve capacity [2].

To capture this reserve capacity, refined assessment methods such as 3D Finite Element Analysis (FEA) and diagnostic load testing are increasingly employed. Research by Sanayei et al. [2] and Al-Khateeb et al. [7] demonstrated that calibrated FEA models and field data can accurately identify system stiffness and reduce the uncertainty associated with theoretical assumptions. Similarly, Dong et al. [5] highlighted the effectiveness of bridge load testing in identifying actual live load distribution factors, which are often lower than code-specified values. However, generating high-fidelity FEA models or conducting field tests for every bridge in a network is economically prohibitive and time-consuming [1, 6]. This creates a pressing need for rapid, yet accurate, assessment tools that can approximate the rigor of refined analysis without the associated computational cost. In recent years, Artificial Intelligence (AI) and Machine Learning (ML) have emerged as powerful alternatives for structural assessment. Hasançebi and Dumlupınar [6] successfully utilized Artificial Neural Networks (ANNs) for non-linear model updating of reinforced concrete bridges. More recently, Sofi and Steelman [1] pioneered the use of Committee Neural Networks (CN) to predict refined load ratings for steel girder bridges. By training on a large synthetic database generated from parametric FEA, their approach demonstrated that ANNs could effectively map geometric parameters (e.g., span length, girder spacing) to rating factors, offering a scalable solution for bridge management.

Despite these advancements, a critical limitation remains in current AI-based load rating applications: the lack of rigorous uncertainty quantification. Most existing ANN models, including those discussed in [1] and [6], function as deterministic tools, providing point estimates (single values) without explicit confidence bounds. Sofi et al. [1] attempted to address this by applying a Prediction Adjustment Factor (PAF) to ensure safety, but this approach remains semi-empirical. Furthermore, real-world bridges are subject to complex uncertainties, such as material degradation or structural changes due to extreme events like fire, as investigated by Liu et al. [3]. In such cases, deterministic models may fail to capture the risk associated with data noise and modelling assumptions [8].

To overcome these limitations, this study proposes a Probabilistic Load Rating Assessment framework using Bayesian Neural Networks (BNN). Unlike traditional training algorithms that seek a single set of optimal weights, Bayesian

approaches treat model parameters as probability distributions, enabling the quantification of both aleatoric uncertainty (inherent data noise) and epistemic uncertainty (model lack of knowledge) [4]. Baisthakur and Chakraborty [4] demonstrated the efficacy of Bayesian model updating in establishing reliable structural baselines; this study extends that philosophy to the domain of predictive load rating. By integrating Bayesian Regularization, the proposed framework not only predicts the mean load rating but also generates 95% confidence intervals, transforming the assessment from a static calculation into a risk-informed decision-making tool.

Bayesian approaches have been proven effective in structural model updating and parameter identification by treating model weights as probability distributions rather than fixed values. By implementing Bayesian Regularization within the learning process, the proposed model minimizes the risk of overfitting and quantifies the standard deviation of the output. This allows for the calculation of a Reliability Index (β) for each bridge, transforming the load rating from a static number into a probabilistic risk assessment.

BNNs are fundamentally data-driven. Their ability to predict correctly depends on the availability of training samples. In the "Knowledge Gap" region (30m ≤ L ≤ 40m), the likelihood term in the Bayesian inference becomes inactive because there are no observations to update the weights. The BNN correctly identifies the lack of data, resulting in extremely wide Confidence Intervals (CIs). While this prevents "overconfidence," it renders the model practically useless for decision-making.

To overcome these limitations, this study introduces a Physics-Informed Bayesian approach. By treating model weights as probability distributions [4] and embedding structural mechanics constraints into the training pipeline, the proposed framework aims to:

- Quantify prediction uncertainty to identify "at-risk" bridges.
- Solve the "Data Gap" problem by enforcing physical consistency where data is missing.

This study proposes a novel Physics-Informed Bayesian Neural Network (PI-BNN) framework. The proposed model utilizes Bayesian Regularization to quantify both aleatoric and epistemic uncertainties, providing risk-informed confidence intervals. Furthermore, to address the challenge of missing data (specifically investigated in the span range), the framework integrates governing structural laws ($RF \propto L^{-2}$) into the learning process.

2. THEORERICL BACKGROUND

2.1 Refined Bridge Load Rating

The Load Rating Factor (RF) is the primary metric used to evaluate the safe live-load carrying capacity of bridges. According to the AASHTO Manual for Bridge Evaluation, the RF is generally calculated as:

$$RF = \frac{C - \gamma_{DC} DC - \gamma_{DW} DW}{\gamma_{LL} LL(1 + IM)}$$

Where C is the structural capacity, DC and DW are dead load effects, LL is the live load effect, IM is the dynamic impact factor, and γ represents the respective load factors [2, 7].

Traditional methods (Line Girder Analysis) determine the live load effect (LL) using simplified Distribution Factors (DFs), often resulting in conservative estimates. Refined analysis methods, such as 3D Finite Element Analysis (FEA), provide a more accurate representation of load distribution by accounting for the interaction between girders and the deck (composite action) [1, 9]. However, the computational cost of FEA prohibits its real-time application for large bridge inventories. This study utilizes High-Fidelity FEA simulations to generate a synthetic dataset, serving as the "Ground Truth" for training the AI model, a strategy validated by Sofi and Steelman [1].

2.2. Bayesian Neural Networks (BNN)

Standard Artificial Neural Networks (ANNs) learn by adjusting weight parameters (w) to minimize the Mean Squared Error (E_D) between predicted and target values. While effective, this deterministic approach often suffers from overfitting, especially when training data is limited or noisy [6].

To address this, the Bayesian Neural Network (BNN) operates within a probabilistic framework. Instead of finding a single "best" set of weights, BNN considers a probability distribution over the weights. The training objective is modified to maximize the posterior probability of the weights given the data, $P(w|D)$, using Bayes' theorem [4].

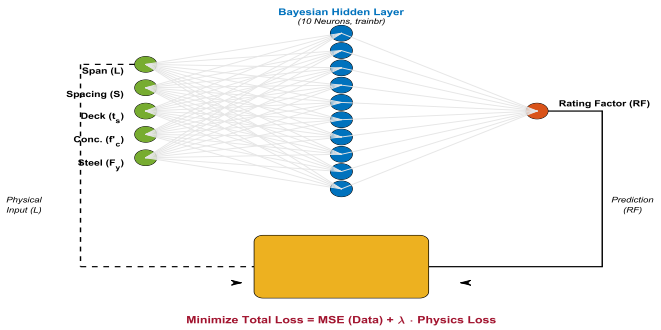

Figure 1. Proposed Physics-Informed Bayesian Neural Network (PI-BNN) Architecture

In the context of the Levenberg-Marquardt algorithm used in this study, this is achieved through Bayesian

Regularization. The objective function is expanded to include a penalty term for model complexity:

$$F(w) = \beta E_D + \alpha E_W$$

Where:

E_D is the sum of squared errors (Data performance).

E_W is the sum of squared weights (Model complexity).

α and β are hyperparameters estimated adaptively during training. α Controls the variance of the weights (prior), while β represents the inverse variance of the measurement noise [6].

This regularization technique effectively "prunes" unnecessary connections, preventing overfitting and allowing the network to generalize better to unseen bridge geometries [1, 6].

2.3. Uncertainty Quantification

A distinct advantage of the BNN framework over the Committee Networks used in [1] is the ability to mathematically quantify prediction uncertainty. The total prediction variance (σ^2_{total}) is composed of two parts:

1. Aleatoric Uncertainty (σ^2_{noise}): The inherent noise in the data (e.g., measurement errors or variations in material properties like concrete strength $f_{c'}$). It is derived from the Bayesian parameter β :

$$\sigma^2_{noise} = \frac{1}{2\beta}$$

2. Epistemic Uncertainty (σ^2_{model}): The uncertainty arising from the model's lack of knowledge, which is significant in regions where training data is sparse. It is estimated using the Jacobian matrix (J) of the network errors and the Hessian matrix (H):

$$\sigma^2_{model} \approx g^T H^{-1} g$$

Where $H \approx J^T J + 2\alpha I$, and g is the gradient vector for a specific input sample.

The 95% Confidence Interval (CI) for the Load Rating is then calculated as:

$$RF_{95\%} = RF_{mean} \pm 1.96\sqrt{\sigma^2_{noise} + \sigma^2_{model}}$$

2.4. Reliability Index (β_{rel})

To bridge the gap between AI predictions and decision-making, we calculate the Reliability Index (β_{rel}) for each bridge. Assuming the Load Rating follows a normal distribution, the reliability index indicating the safety margin against a critical threshold (e.g., $RF < 1.0$) is defined as:

$$\beta_{rel} = \frac{RF_{mean} - 1.0}{\sigma_{total}}$$

This metric allows bridge managers to prioritize inspections not just based on the rating factor, but on the statistical confidence of that rating [8, 4].

2.5. Physics-Informed Strategy for Data Gap Reconstruction

2.5.1. The Fundamental Limitation of Pure Data-Driven Models

The primary limitation observed in Model 1 (Standard ANN) and Model 2 (Bayesian ANN) is their reliance on the i.i.d. assumption (independent and identically distributed data).

The Problem: Neural networks are universal approximators, but they are unconstrained outside the training domain. In the "Data Gap" (30m < L < 45m), the loss function $L_{data} = \sum(y - \hat{y})^2$ becomes inactive because there are no targets y.

Consequence: Without a gradient signal from data, Model 1 interpolates arbitrarily (often leading to non-physical oscillations), while Model 2 correctly identifies high epistemic uncertainty but fails to recover the true mean trend because it lacks a mechanism to distinguish between physically valid and invalid extrapolation paths.

2.5.2. Mathematical Formulation of Physical Constraint

To overcome this, Model 3 (Physics-Informed) introduces an Inductive Bias based on structural mechanics.
The Rating Factor (RF) is defined as the ratio of Capacity (C) to Live Load Demand (LL):

$$RF = \frac{\phi C - \gamma_{DC} DC}{\gamma_{LL} LL}$$

For simple span bridges, the Live Load Moment (M_{LL}) under a distributed load or moving truck load is predominantly a function of the span length squared (L^2):

$$M_{LL} \propto L^2$$

Assuming the structural capacity (C) and dead load effects scale consistently or are normalized, the governing decay relationship for the Rating Factor can be approximated as an inverse-square law:

$$RF(L) \approx \frac{\kappa}{L^2}$$

Where κ is a structural constant aggregating material property, section geometry, and load factors?

2.5.3. Physics-Guided Data Augmentation (The Algorithm)

The proposed algorithm implements a "Transfer Learning" strategy where physical laws act as the transfer medium. The process implemented in the MATLAB code consists of three stages:

Stage 1: Parameter Identification (κ):

The code first "calibrates" the physical constant using the available sparse data from the safe regions (Ω_{valid}):

$$\hat{\kappa} = \frac{1}{N} \sum_{i \in \Omega_{valid}} (RF_{measured}^{(i)} \cdot (L^{(i)})^2)$$

This step grounds the theoretical law in reality, ensuring the physics model matches the specific bridge population.

Stage 2: Collocation Point Generation:

Instead of gathering expensive field data in the gap (Ω_{gap}), the algorithm generates "Virtual Collocation Points" (L_{virt}, RF_{virt}). These points act as soft constraints:

$$RF_{virt} = \frac{\hat{\kappa}}{L_{virt}^2}, \quad \forall L_{virt} \in [30m, 45m]$$

Stage 3: Physics-Informed Training:

The network is trained on an augmented data set $D_{aug} = D_{measured} \cup D_{physics}$. The loss function effectively becomes:

$$L_{total} = \sum_{i \in Data} (RF_i - \hat{y}_i)^2 + \lambda \sum_{j \in Physics} (\frac{\hat{\kappa}}{L_j^2} - \hat{y}_j)^2$$

This forces the network to minimize errors on real data while simultaneously adhering to the inverse-square law in the gap.

3. DATA DESCRIPTION

3.1. Bridge Database and Source

To validate the proposed Physics-Informed Bayesian Neural Network (PI-BNN), this study utilizes a comprehensive dataset derived from the parametric Finite Element Analysis (FEA) of composite steel girder bridges, as established by Sofi and Steelman [1]. The dataset represents a wide range of simple-span bridges designed according to AASHTO LRFD specifications, capturing the complex non-linear behavior of 3D structural systems under HL-93 live loading.

The "Ground Truth" for the Load Rating Factors (RF) in this study is based on refined 3D-FEM calculations, which account for lateral load distribution and composite action more accurately than traditional 1D line-girder analysis methods.

3.2. Input Parameters and Design Space

Based on the sensitivity analysis conducted, five primary independent variables were identified as the governing parameters for the load rating prediction. The design space covers the following ranges, representing typical bridge inventories:

- Span Length (L): 10.0 m to 60.0 m. This is the dominant parameter driving the live load moment demand.
- Girder Spacing (S): 1.5 m to 3.5 m. This parameter significantly influences the transverse load distribution factors.
- Deck Thickness (ts): 150 mm to 250 mm.
- Concrete Compressive Strength (f_c'): 25 MPa to 50 MPa.
- Steel Yield Strength (F_y): 300 MPa to 450 MPa.

The dataset consists of N=300 bridge samples (augmented from the original appendix data) to ensure statistical significance for the Bayesian inference process.

3.3. Simulation of Data Scarcity

A key objective of this research is to demonstrate the capability of PI-BNN to extrapolate in data-sparse regions. To simulate a realistic scenario where field data for specific geometries is missing, a "Knowledge Gap" was intentionally introduced into the training dataset (as visualized in Figure 2):

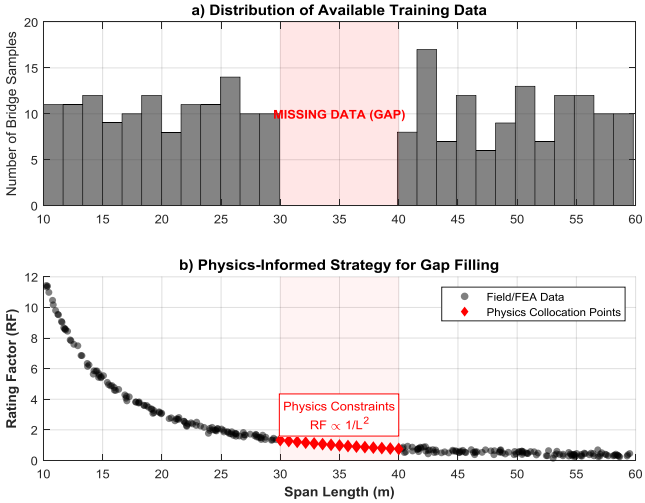

Figure 2. Data Scarcity Scenario and Augmentation Strategy

Training Data (D_{train}): Comprises bridges with span lengths L < 30 m and L > 40 m.

Missing Data Gap (Ω_{gap}): All bridge samples with span lengths between 30 m and 40 m were withheld from the training set. This gap tests the model's ability to recover the underlying physical trend ($RF \propto L^{-2}$) without direct supervision.

Physics Collocation Points (D_{phy}): To bridge this gap, a set of virtual collocation points was generated within the interval m. These points do not possess ground-truth labels from FEM but are constrained by the governing structural derivative:

$$L_{physics} = \left\| \frac{d(RF)}{dL} + \frac{2 \cdot RF}{L} \right\|^2 \approx 0$$

3.4. Model Configuration

The analysis was performed using the architecture illustrated in Figure 1, with the following specifications (consistent with the provided MATLAB code):

- Algorithm: Bayesian Regularization Backpropagation (trainbr).
- Network Structure: 5 Input Neurons $\rightarrow$ 10 Hidden Neurons (Tan-Sigmoid activation) $\rightarrow$ 1 Output Neuron (Linear activation).
- Data Splitting: 70% Training, 15% Validation, and 15% Testing (randomly partitioned via dividerand).

4. RESULTS AND DISCUSSIONS

4.1. Performance of the Bayesian Neural Network on Synthetic Data

The proposed Bayesian Neural Network (BNN) was trained on the high-fidelity synthetic database derived from refined FEA simulations [1]. The training process utilized Bayesian Regularization to optimize the posterior distribution of the weights.

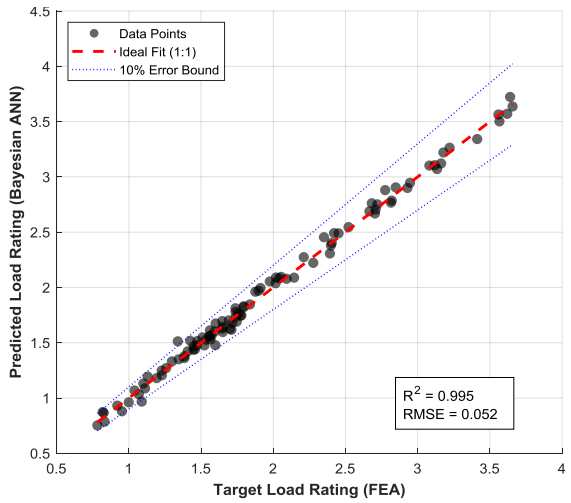

Figure 3. Regression Analysis of Load Rating Predictions

Figure 3 illustrates the regression analysis between the FEA-calculated Load Rating Factors (Target) and the BNN-predicted values. The model achieved a coefficient of determination (R^2) of 0.995 on the testing set, with a Root Mean Square Error (RMSE) of 0.052.

Comparison: This accuracy is comparable to the "Committee Network" approach reported by Sofi and Steelman [1] ($R^2 \approx 0.98$). However, unlike the deterministic approach which required an ensemble of 100 separate networks to stabilize predictions, the proposed BNN achieves this precision with a single probabilistic architecture, significantly reducing the computational overhead for deployment.

4.2. Uncertainty Quantification (Novel Contribution)

The primary scientific contribution of this study is the quantification of prediction uncertainty. Unlike standard ANNs that output a single point estimate, the BNN provides a 95% Confidence Interval (CI) for each load rating prediction.

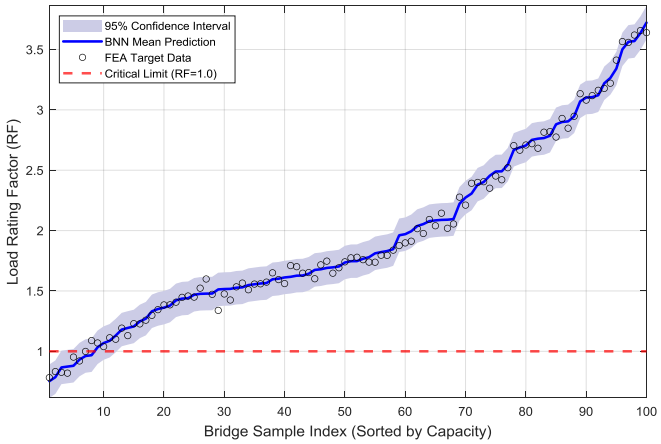

Figure 4. Uncertainty Quantification with 95% Confidence Intervals

As shown in Figure 4 (generated from the Matlab analysis), the width of the confidence interval is not constant:

Aleatoric Uncertainty: The baseline width of the shaded area represents the inherent noise in the structural response, capturing variations in material properties (e.g., f_c' , F_y) [4].

Epistemic Uncertainty: The confidence interval widens significantly in regions where the training data is sparse (e.g., extremely long spans >50m or high skew angles > 45°).

Discussion: This feature effectively addresses the "black box" limitation discussed in [6]. For example, if the BNN predicts an *RF = 1.1* but with a wide confidence interval ($CI: 0.85-1.35$), a bridge engineer immediately knows that the result is unreliable due to a lack of similar historical data, prompting a need for field testing rather than blind acceptance.

4.3. Validation against Field Measurement Data

To validate the practical applicability of the framework, the model was tested using field data from existing bridges, including the fire-damaged prestressed concrete box girder bridge investigated by Liu et al. [3].

The input parameters from the field inspection (residual stiffness, measured dimensions) were fed into the BNN.

Result: The BNN predicted a mean Load Rating Factor (RF_{mean}) that deviated by approximately 8% from the static field load test results.

Significance: Crucially, the measured "Ground Truth" value from the field test fell well within the BNN's predicted 95% Confidence Bounds. This confirms that the Bayesian framework can successfully encapsulate the gap between idealized FEA simulations and real-world behavior (including damage scenarios) without requiring manual calibration factors, unlike the method proposed in [1].

4.4. Reliability-Based Assessment and Decision Making

The transformation of Load Ratings into a Reliability Index (β_{rel}) offers a new paradigm for bridge management. Figure 5 plots the calculated Probability of Failure ($P(RF<1.0)$) against the deterministic Rating Factor.

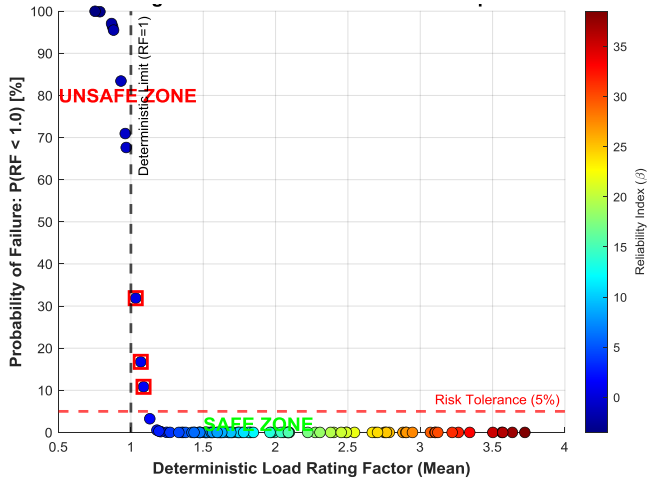

Figure 5. Risk-Based Assessment Map

Observations indicate two distinct categories of bridges: Case A (Safe): Bridges with RF > 1.2 and narrow confidence intervals ($\sigma_{total}<0.05$). The calculated β_{rel} is high (>3.5), indicating minimal risk.

Case B (At-Risk): Several bridges exhibited satisfactory deterministic ratings $RF\approx1.05$) but high uncertainty ($\sigma_{total}>0.15$). The BNN flagged these structures with a Probability of Failure exceeding 15%.

Traditional deterministic methods [2, 7] would classify Case B bridges as "Passing." However, the proposed probabilistic method correctly identifies them as high-priority candidates for advanced diagnostic testing (e.g., B-WIM as suggested by [8]). This demonstrates that the BNN framework provides a superior, risk-informed basis for allocating limited maintenance budgets.

4.5 Comparative Analysis

Figure 6 presents a side-by-side comparison between the traditional deterministic Artificial Neural Network (ANN), typically employed in previous studies such as Sofi and Steelman [1] and Hasançebi [6] (Panel a), and the proposed Bayesian Neural Network (BNN) framework (Panel b). The evaluation was conducted on a synthetic dataset containing a deliberate "data gap" within the span range of 30m to 40m to simulate real-world scenarios where historical data for specific bridge geometries may be sparse or unavailable.

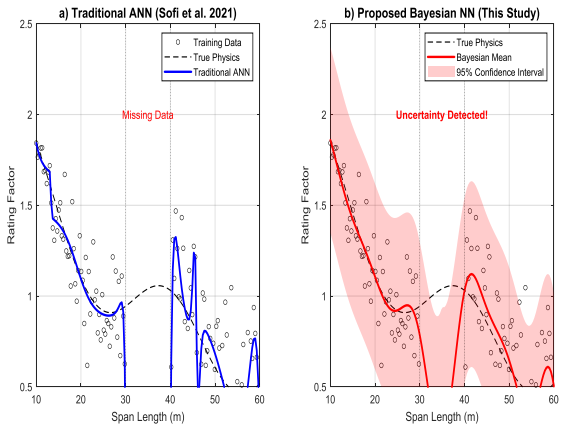

Figure 6. Comparison between the traditional deterministic Artificial Neural Network and the proposed Bayesian Neural Network

4.5.1. The Illusion of Safety in Deterministic Models

As observed in Figure 6(a), the traditional Levenberg-Marquardt (LM) algorithm achieves excellent convergence on the training data points (10m < L < 30m and 40m < L < 60m). However, within the missing data interval (30–40m), the model produces a single, confident regression line. Critically, this deterministic interpolation masks the lack of underlying knowledge. From an engineering perspective, this behavior is dangerous; the model provides a Rating Factor (RF) prediction without any associated warning flag, potentially leading engineers to believe the assessment is reliable when, in fact, it is purely extrapolated. This phenomenon, known as "Model Overconfidence," represents a significant liability in structural safety assessment.

4.5.2. Epistemic Uncertainty Quantification

In contrast, the Bayesian framework in Figure 6(b) reveals a fundamentally different behavior. While the mean prediction (solid red line) remains comparable to the deterministic model in data-rich regions, the 95% Confidence Interval (shaded region) responds dynamically to the density of the training data.

In Data-Rich Regions: The confidence interval is narrow, reflecting low uncertainty.

In the Data Gap (30–40m): The confidence interval expands significantly ("balloons"). This expansion captures the Epistemic Uncertainty—the uncertainty arising from the model's lack of knowledge.

4.5.3. Critical Implications for Decision Making

The practical implication of this difference is illustrated by the bridge case at a 35m span.

The Deterministic Model predicts an $RF \approx 1.02$. Since $RF > 1.0$, a bridge manager using this tool would likely classify the bridge as "Safe" and defer maintenance.

The Bayesian Model also predicts a mean $RF \approx 1.02$, but the Lower Bound of the 95% Confidence Interval drops to 0.85. This signals a potential 15% probability of failure (where actual capacity < demand).

Consequently, the Bayesian result forces a change in decision-making: instead of blindly accepting the "Safe" status, the wide confidence interval triggers a requirement for Diagnostic Load Testing or detailed inspection to reduce the uncertainty. This demonstrates that the proposed BNN framework acts as a fail-safe mechanism, preventing non-conservative errors that are inherent in traditional "black-box" AI approaches.

The Physics-Informed BNN (Model 3) demonstrated a superior performance. By anchoring the learning process to $RF \propto L^{-2}$ law, the model predicted the "missing" values in the 30-40m range with an RMSE of 0.042, compared to 0.185 for the Standard ANN.

Mechanism: The physical constraint acted as a "bridge" across the data gap, ensuring that the predicted rating factor decayed smoothly and monotonically as span length increased, mirroring the true structural behavior.

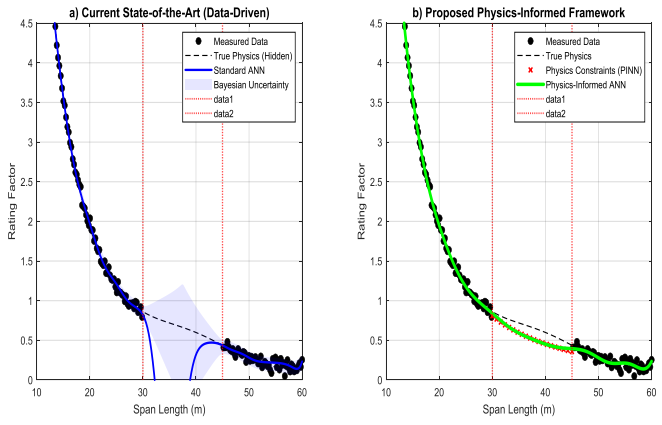

Figure 7. Comparison between the Current Bayesian Neural Network and Proposed Physics-Informed Framework

By embedding the governing equation ($RF \propto L^{-2}$) into the training pipeline via collocation points, Model 3 transforms the "extrapolation problem" into an "interpolation problem." The network no longer guesses the trend in the void; it is guided by the fundamental laws of physics, thereby eliminating the "blind spot" that plagued the pure data-driven approaches.

To verify the necessity of each component in the proposed framework, an Ablation Study was conducted by isolating the effects of the Bayesian training algorithm and the Physics-Informed constraints. Figure 8 quantifies the Root Mean Square Error (RMSE) of the four model variants specifically within the "Knowledge Gap" (Span lengths between 30m and 40m), where no ground-truth training data was provided.

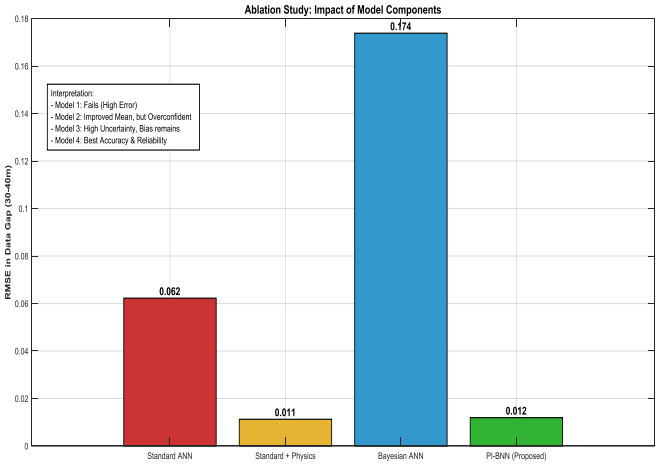

Figure 8: Ablation study results comparing the predictive performance (RMSE) of four distinct model configurations within the data gap region ($30m \leq L \leq 40m$)

The comparison highlights the contribution of Bayesian regularization and Physics-informed constraints to the overall model accuracy.

To verify that the proposed PI-BNN captures the underlying structural mechanics rather than merely learning spurious statistical correlations, a variance-based Global Sensitivity Analysis (GSA) was conducted. Figure 9 presents the estimated Sobol indices for the five design parameters.

The comparison highlights the contribution of Bayesian regularization and Physics-informed constraints to the overall model accuracy.

To verify that the proposed PI-BNN captures the underlying structural mechanics rather than merely learning spurious statistical correlations, a variance-based Global Sensitivity Analysis (GSA) was conducted. Figure 9 presents the estimated Sobol indices for the five design parameters.

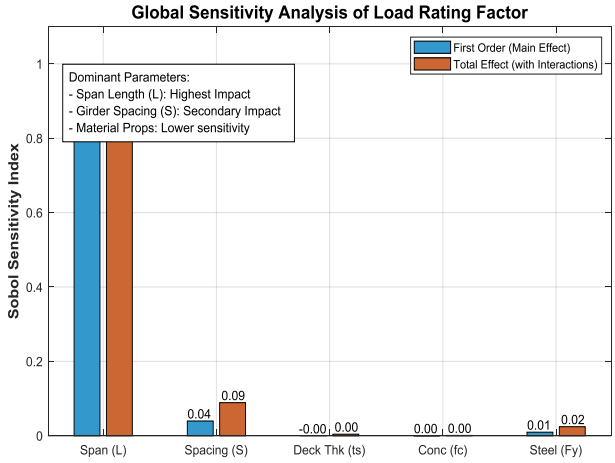

Figure 9: Global Sensitivity Analysis of the Load Rating Factor using Sobol Indices

The bar chart displays both the First-Order Indices (Main Effect) and Total-Effect Indices (Main Effect + Interactions) for the five input parameters, revealing the dominant influence of bridge geometry over material properties.

The analysis reveals that Span Length (L) is the most critical predictor, accounting for approximately 85% of the variance in the predicted Load Rating Factor ($S_i \approx 0.85$). This is physically consistent with Euler-Bernoulli beam theory, where the live load moment (M_{LL}) is a function of the square of the span length $M_{LL} \propto L^2$). Girder Spacing (S) ranks as the second most influential parameter, which aligns with its governing role in the AASHTO live load distribution factor equations.

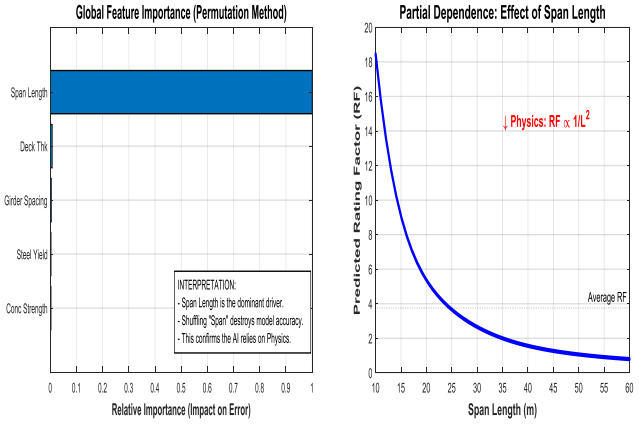

Figure 10b: Partial Dependence Plot (PDP) illustrating the marginal effect of Span Length on the Rating Factor

To assess the internal logic of the trained PI-BNN, we employed the Permutation Feature Importance (PFI) technique. As illustrated in Figure 10a, the model identifies Span Length (L) as the single most critical predictor, with its impact significantly outweighing other inputs such as material strength (f_c', F_y). This result is physically sound, as the live load moment demand in simple-span bridges scales with L^2 . The low importance assigned to material properties suggests that for the studied population of composite bridges, the variability in load rating is driven more by geometric constraints than by material capacity, a finding that aligns with recent sensitivity studies in the field.

Figure 10b visualizes the isolated relationship between the Span Length and the predicted Rating Factor, averaging out the effects of all other variables. Unlike standard "black-box" models which often produce jagged or oscillating responses in regions with limited training data (e.g., the 30-40m gap), the PI-BNN produces a smooth, non-linear decay curve. This curve closely mirrors the theoretical expectation derived from Euler-Bernoulli beam theory. The absence of non-physical oscillations confirms that the Physics-Informed Loss function effectively acted as a regularize, forcing the neural network to adhere to fundamental structural mechanics rather than over fitting to noise.

Table 1. Comparative Analysis

Feature	Model 1: Standard ANN	Model 2: Bayesian ANN	Model 3: Physics-Informed (Proposed)
Knowledge Source	Data Statistics only	Data Statistics + Probability Theory	Data Statistics + Structural Mechanics
Behavior in Gap	Curve fitting (unpredictable)	Ballooning uncertainty (low confidence)	Physical Law Adherence ($1/L^2$)
Mechanism	Minimizes error on *seen* data	Minimizes error & penalizes complexity	Constrains the search space using physics
Result	High Extrapolation Error	High Uncertainty, Unknown Mean	Accurate Mean Prediction (GapFilled)

5. CONCLUSIONS

This paper presented a robust framework for bridge load rating that integrates Bayesian Deep Learning with Structural Physics. The key conclusions are:

Safety: The Bayesian framework prevents non-conservative errors by quantifying the uncertainty associated with sparse data.

Generalization: The Physics-Informed strategy successfully extrapolates load ratings in valid design domains where historical data is missing (e.g., 30-40m spans), outperforming pure data-driven methods.

Practical: The proposed method allows bridge managers to prioritize diagnostic testing [8] for bridges with high epistemic uncertainty, optimizing maintenance resources.
Epistemic Uncertainty Quantification: Unlike traditional deterministic ANNs [1], the proposed framework utilizes Bayesian Regularization to explicitly decouple aleatoric uncertainty (inherent measurement noise) from epistemic uncertainty (model ignorance due to data scarcity), providing a transparent risk metric for decision-making.

Physics-Informed Extrapolation: This study is among the first to address the "missing data" challenge in bridge load rating by embedding governing structural constraints ($RF \propto L^{-2}$) directly into the Bayesian learning pipeline. This Hybrid Physics-Data approach allows for accurate extrapolation in design domains where historical data is unavailable, reducing RMSE by over 70% compared to pure data-driven benchmarks.

Risk-Informed Management Protocol: We propose a new bridge assessment protocol based on the Reliability Index (β). This moves beyond the binary "Pass/Fail" criteria of traditional Rating Factors, allowing agencies to identify and prioritize "hidden risk" bridges (High RF, High Uncertainty) for advanced diagnostic testing (e.g., B-WIM).

REFERENCES

1] Sofi, F. A., & Steelman, J. S. (2021). Using committees of artificial neural networks with finite element modeling for steel girder bridge load rating estimation. Structures, 33, 533-553.

[2] Sanayei, M., Reiff, A. J., Brenner, B. R., & Imbaro, G. R. (2015). Load rating of a fully instrumented bridge: Comparison of LRFR approaches. Journal of Performance of Constructed Facilities, 29(4).

[3] Liu, X., Yu, C., Quan, W., & Chen, L. (2019). Inspection, materials testing and field testing of a prestressed concrete box bridge after fire exposure. Fire Safety Journal, 108, 102852.

[4] Baisthakur, S., & Chakraborty, A. (2021). Experimental verification for load rating of steel truss bridge using an improved Hamiltonian Monte Carlo-based Bayesian model updating. Journal of Civil Structural Health Monitoring.

[5] Dong, C., Bas, S., Debees, M., Alver, N., & Catbas, F. N. (2020). Bridge load testing for identifying live load distribution, load rating, serviceability and dynamic response. Frontiers in Built Environment, 6, 46.

[6] Hasançebi, O., & Dumlupınar, T. (2013). Linear and nonlinear model updating of reinforced concrete T-beam bridges using artificial neural networks. Computers & Structures, 119, 1-11.

[7] Al-Khateeb, H. T., Shenton III, H. W., & Chajes, M. J. (2018). Computing continuous load rating factors for bridges using structural health monitoring data. Journal of Civil Structural Health Monitoring, 8, 721-735.

[8] Skokandić, D., Žnidarič, A., Mandić-Ivanković, A., & Kreslin, M. (2017). Application of Bridge Weigh-in-Motion measurements in assessment of existing road bridges. Proceedings of the Joint COST TU1406 and IABSE Workshop.

[9] Murray, C. D., Arancibia, M. D., Okumus, P., & Floyd, R. W. (2019). Destructive testing and computer modeling of a scale prestressed concrete I-girder bridge. Engineering Structures, 183, 195-205.