Article history:
Received 01 February 2026
Revised 13 May 2026
Accepted 23 June 2026
Initial Publish 26 July 2026
Published online 01 November 2027

Journal of Resource Management and Decision Engineering

Volume 6, Issue 6, pp 1-15

An Integrated Expert System for Prioritizing Concrete Railway Bridges Maintenance: A Hybrid Fuzzy Delphi - Fuzzy TOPSIS - MLP Framework

Hossein. Farshad[1], Nader. Abdoli Yazdi[1*], Ehsan. Eshtehardian[2]

[1] Department of Civil Engineering, Taft. C., Islamic Azad University, Taft, Iran
[2] Department of Project Management, Tarbiat Modares University, Tehran, Iran

*** Corresponding author email address**: abdoli@iau.ac.ir

A r t i c l e I n f o

Article type:
Original Research

How to cite this article:
Farshad, H., Abdoli Yazdi, N., & Eshtehardian, E. (2027). An Integrated Expert System for Prioritizing Concrete Railway Bridges Maintenance: A Hybrid Fuzzy Delphi - Fuzzy TOPSIS - MLP Framework. *Journal of Resource Management and Decision Engineering*, *6*(6), 1-15.
https://doi.org/10.61838/kman.jrmde.375

A B S T R A C T

Maintenance prioritization of railway bridges is a critical challenge due to limited financial resources and the direct impact of bridge failures on transportation network safety. Traditional approaches relying on individual expert judgment are inadequate for managing the complexity of maintaining precast concrete railway bridges, particularly in environmentally aggressive corridors. This study presents an integrated three-stage expert system (hybrid decision-support framework) for maintenance prioritization of 85 precast box-girder concrete bridges along the Bonekuh□Gorgan railway corridor (Northern Iran). The proposed framework combines: (1) Fuzzy Delphi for extracting, screening, and weighting maintenance criteria through the judgment of 15 domain experts; (2) Fuzzy TOPSIS for quantitative prioritization and computation of the Intervention Priority Index (Cl+) across three inspection periods (2020, 2022, and 2024); and (3) Multilayer Perceptron (MLP) neural network for predicting the future trend of the Cl+ index and supporting preventive maintenance planning. In the Fuzzy Delphi stage, 12 key criteria were identified from 20 initial candidates; expansion joints (C5) and guardrails (C12) received the highest weights (w = 0.096). Kendall's concordance coefficient (W = 0.69) and Cronbach's alpha (α = 0.847) confirmed satisfactory consensus and reliability among experts. Fuzzy TOPSIS results showed that bridge B1, with Cl+ = 0.826, had the most critical condition requiring immediate intervention, while bridge B48, with Cl+ = 0.214, exhibited the lowest deterioration level. The MLP model with a 13-64-1 architecture and Walk-Forward Validation predicted the Cl+ index with $R^2 = 0.931$ and RMSE = 0.0214, demonstrating high generalizability. The results indicate that the proposed framework can serve as a practical Decision Support System (DSS) for railway bridge maintenance management, supporting optimal allocation of limited resources with scientific precision and transparency.

Keywords: *Expert system; Railway concrete bridge maintenance; Multi-criteria decision-making (MCDM); Fuzzy TOPSIS; MLP neural network; Maintenance prioritization*

1. Introduction

Railway bridges are among the most critical components of transportation infrastructure because they directly influence network continuity, operational safety, freight mobility, passenger reliability, and the resilience of national logistics systems. Unlike many ordinary civil structures, railway bridges are subjected to repeated dynamic loads, vibration effects, axle-load variation, environmental exposure, and strict serviceability requirements. Their deterioration may not only reduce structural capacity but also interrupt railway operations and generate high economic and safety consequences. For this reason, bridge maintenance has increasingly shifted from a reactive engineering activity to a strategic asset-management function that requires systematic inspection, prioritization, risk assessment, and life-cycle planning (Frangopol, 2011; Ryall, 2009). In railway networks with large bridge inventories and limited maintenance budgets, managers must determine which bridges require immediate intervention, which bridges can be monitored under routine schedules, and which bridges should receive preventive maintenance before deterioration becomes structurally or economically irreversible. This decision is inherently multi-dimensional, uncertain, and time-dependent.

Concrete railway bridges, particularly precast box-girder bridges, are widely used because of their construction speed, dimensional control, cost efficiency, and suitability for repetitive railway infrastructure projects. Nevertheless, concrete bridge systems are exposed to complex deterioration mechanisms, including reinforcement corrosion, chloride ingress, carbonation, freeze–thaw damage, moisture penetration, cracking, drainage failure, bearing deterioration, joint leakage, and surface-protection degradation. The durability of concrete is strongly affected by the interaction between material microstructure, environmental aggressiveness, and mechanical loading, meaning that deterioration cannot be adequately understood through a single criterion or visual judgment alone (Mehta & Monteiro, 2014). In humid, saline, or freeze–thaw-active corridors, defects in drainage systems, expansion joints, protective coatings, and deck components may accelerate deterioration by allowing water and aggressive ions to penetrate the structural system. Therefore, bridge maintenance prioritization requires a framework capable of integrating structural, operational, environmental, and safety-related criteria.

Traditional bridge maintenance decisions have often relied on periodic inspections and the professional judgment of inspectors or maintenance engineers. While expert judgment is indispensable, individual judgment may be inconsistent, subjective, and difficult to reproduce, especially when several bridges compete for the same limited budget. Modern bridge management therefore requires decision-support models that can organize expert knowledge, quantify bridge condition, manage uncertainty, and produce transparent maintenance priorities. Bridge management literature emphasizes that infrastructure decisions should consider life-cycle performance, uncertainty, resilience, and optimization rather than short-term repair needs alone (Bocchini & Frangopol, 2012; Frangopol, 2011). The increasing complexity of infrastructure systems has also made it necessary to use models that can support both current condition assessment and forward-looking maintenance planning.

Multi-criteria decision-making methods are particularly appropriate for bridge maintenance prioritization because bridge condition is determined by multiple heterogeneous indicators rather than by a single measurable variable. Fuzzy multi-criteria decision-making approaches are especially useful because bridge inspection data and expert judgments are often imprecise, linguistic, and uncertain. Fuzzy logic enables uncertain judgments such as “high deterioration,” “moderate importance,” or “critical condition” to be represented mathematically through fuzzy numbers, thereby preserving expert knowledge while reducing the rigidity of crisp scoring systems (Kahraman, 2008). Among fuzzy decision-making approaches, Fuzzy TOPSIS has received considerable attention because it ranks alternatives according to their relative closeness to positive and negative ideal solutions. In bridge risk assessment, Fuzzy TOPSIS has been shown to be capable of incorporating uncertain condition information and supporting systematic prioritization under vagueness (Wang & Elhag, 2006). More recent work has continued to refine fuzzy assessment approaches for bridge condition evaluation, demonstrating their suitability for infrastructure systems where data uncertainty and expert interpretation are unavoidable (Li et al., 2025; Qu et al., 2025).

However, the effectiveness of any multi-criteria ranking model depends heavily on the validity of the criteria and their weights. If criterion weights are assigned arbitrarily or by a single decision-maker, the final ranking may become biased or unstable. The Fuzzy Delphi Method provides a structured way to extract, screen, and weight decision criteria through expert consensus while accounting for the uncertainty of linguistic judgments. Hybrid Fuzzy Delphi–TOPSIS frameworks have been used in complex decision environments where criteria identification and alternative ranking must be performed together, showing that expert consensus and fuzzy prioritization can be integrated into a coherent decision process (Wang et al., 2014). In bridge maintenance, this integration is especially valuable because maintenance criteria such as deck condition, expansion joints, bearing seats, drainage systems, protective coating, and guardrails may differ in importance depending on bridge type, environmental exposure, and operational context. A structured expert-based weighting procedure can therefore improve the scientific defensibility of prioritization.

In recent years, knowledge-based fuzzy expert systems have also been applied to bridge condition assessment, showing that formalized expert knowledge can assist in evaluating structural condition where direct measurements are incomplete or uncertain (Rogulj et al., 2021). These approaches represent an important movement away from isolated subjective inspection toward reproducible and transparent assessment. Nevertheless, many fuzzy expert systems remain static; they assess the current condition of bridges but do not predict future deterioration. For railway bridge management, static prioritization is insufficient because deterioration is a temporal process. A bridge that is currently in moderate condition may deteriorate rapidly due to water ingress or repeated dynamic loading, while another bridge in poor condition may stabilize after successful maintenance. Therefore, a comprehensive maintenance system must not only rank bridges at a given point in time but also forecast future risk and deterioration trends.

Machine learning has become increasingly important in bridge deterioration prediction because it can model nonlinear relationships among structural indicators, environmental exposure, maintenance history, and condition outcomes. Classical statistical learning provides the foundation for data-driven prediction, model evaluation, and generalization assessment (Hastie et al., 2009). Neural networks, particularly Multilayer Perceptrons, are well suited to nonlinear function approximation and have been widely used in engineering prediction tasks due to their ability to learn complex patterns from multidimensional input data (Goodfellow et al., 2016; Haykin, 2007). In bridge engineering, machine learning has been applied to predict bridge deck deterioration, maintenance needs, and network-level performance under budget constraints. For example, optimized machine learning algorithms have improved the prediction of bridge deck deterioration, while machine learning-based optimization models have supported bridge maintenance planning under limited annual budgets (Ghafoori et al., 2024; Rashidi-Nasab & Elzarka, 2023).

The combination of multi-criteria decision-making and machine learning is particularly promising for bridge maintenance because it can connect interpretable expert-based prioritization with predictive data-driven modeling. In railway bridge contexts, integrated AHP–MLP approaches have been proposed for predictive maintenance, indicating that neural networks can enhance prioritization by forecasting future condition states rather than merely ranking current condition (Gao et al., 2023). Similarly, decision-support models for allocating bridge maintenance budgets demonstrate the practical importance of linking prioritization results with resource allocation and maintenance planning (Alshibani et al., 2023). Despite these advances, most existing studies still address only part of the decision chain. Some focus on expert weighting, others focus on fuzzy ranking, and others focus on prediction. Few studies integrate expert consensus, fuzzy multi-criteria prioritization, temporal validation, and neural-network forecasting in a single operational framework for railway concrete bridges.

A further methodological challenge concerns temporal validation. Bridge deterioration data are inherently sequential: inspection results from the future should not be used to train models intended to predict earlier periods. Conventional random train-test splits may therefore produce overly optimistic estimates of model performance by allowing temporal information leakage. Time-series forecasting literature emphasizes the need for validation strategies that respect chronological order and simulate real forecasting conditions (Box et al., 2015; Hyndman & Athanasopoulos, 2021). Walk-forward validation is especially suitable for infrastructure inspection data because it trains the model on past observations and evaluates it on future observations, thereby approximating the actual decision-making process used in maintenance planning. Incorporating this type of validation strengthens the credibility of machine learning predictions and improves their operational relevance.

Another important issue is model uncertainty. Maintenance decisions based on predicted bridge condition involve risk, and a point prediction alone may be insufficient for budget allocation. If two bridges have similar predicted priority indices but one prediction has much higher uncertainty, the latter may require additional inspection before intervention. Bayesian interpretations of dropout provide a practical way to estimate uncertainty in neural-network predictions through Monte Carlo Dropout, allowing confidence intervals to be generated without fully Bayesian model training (Gal & Ghahramani, 2016). Such uncertainty quantification is important for risk-informed maintenance planning, especially when data are limited and inspection measurements contain error. Furthermore, the interpretability of predictive models is increasingly recognized as essential for real-world decision support, because infrastructure managers must understand why a model produces a given priority ranking before relying on it for costly maintenance decisions (Doshi-Velez & Kim, 2017).

The development of a reliable MLP model also requires careful attention to training stability and generalization. Deep learning research has shown that network initialization, activation functions, regularization, and optimization choices influence model convergence and performance (Glorot & Bengio, 2010; Goodfellow et al., 2016). In small or moderate engineering datasets, overfitting is a substantial risk because the model may memorize historical inspection patterns rather than generalize to future bridge conditions. Therefore, regularization techniques, sensitivity analysis, and repeated validation are essential for confirming that the model produces robust predictions rather than unstable numerical results. Sensitivity analysis is also important for understanding how changes in input variables or criterion weights affect model outputs and rankings, especially in decision-support systems where uncertainty in expert judgments and inspection scores is unavoidable (Saltelli et al., 2009).

For the Bonekuh–Gorgan railway corridor in Northern Iran, these methodological issues are especially relevant. The corridor includes precast concrete railway bridges exposed to high humidity, heavy precipitation, chloride-bearing air masses, freeze–thaw cycles, and operational railway loads. In such an environment, deterioration is not only structural but also environmental and temporal. Expansion joints, drainage systems, protective coatings, foundations, decks, and safety components may interact in ways that cannot be captured by simple inspection scores or single-criterion prioritization. A scientifically defensible maintenance model must therefore combine expert knowledge, fuzzy uncertainty handling, multi-criteria ranking, temporal prediction, and uncertainty quantification. This need creates a clear methodological gap: railway bridge maintenance requires a comprehensive expert system that can move from criteria extraction to current prioritization and then to future prediction within a single validated framework.

The present study addresses this gap by proposing an integrated three-stage expert system for the maintenance prioritization of precast box-girder concrete railway bridges. The framework first uses the Fuzzy Delphi Method to identify and weight maintenance criteria based on expert consensus. It then applies Fuzzy TOPSIS to calculate a quantitative Intervention Priority Index ((Cl^{+})) for each bridge across multiple inspection periods. Finally, it employs an MLP neural network with temporal input features and walk-forward validation to forecast future (Cl^{+}) values and support preventive maintenance planning. By integrating fuzzy expert judgment, multi-criteria decision-making, neural-network prediction, and uncertainty estimation, the framework is intended to provide a transparent and operationally useful decision-support system for allocating limited maintenance resources.

The aim of this study was to develop and validate an integrated Fuzzy Delphi–Fuzzy TOPSIS–MLP expert system for prioritizing the maintenance of precast concrete railway bridges along the Bonekuh–Gorgan railway corridor and predicting future intervention priorities to support preventive and evidence-based maintenance management.

2. Methods and Materials

This study was designed as an applied, methodological, and data-driven decision-support study aimed at developing and validating an integrated expert system for prioritizing the maintenance of precast concrete railway bridges. The proposed framework followed a sequential three-stage structure combining the Fuzzy Delphi Method, Fuzzy TOPSIS, and a Multilayer Perceptron neural network. In the first stage, the Fuzzy Delphi Method was used to identify, screen, and weight bridge maintenance criteria through structured expert consensus. In the second stage, Fuzzy TOPSIS was applied to convert inspection data into a quantitative Intervention Priority Index, denoted as Cl^{+}. In the third stage, an MLP neural network was used to predict the future trend of $Cl^{+}$and support preventive maintenance planning.

The study was conducted on the Bonekuh–Gorgan railway corridor in Northern Iran, a railway corridor of approximately 438km that is exposed to severe environmental and operational deterioration conditions. This corridor is characterized by high relative humidity of approximately 75%to 90%, annual precipitation exceeding 800mm, freeze–thaw cycles lasting up to 45days per year, considerable seasonal temperature variation, chloride exposure from maritime air masses, and steep gradients requiring frequent bridge crossings. These conditions make the corridor highly suitable for evaluating a maintenance prioritization framework for concrete railway bridges.

The bridge inventory included 85precast box-girder concrete railway bridges. Inspection data were collected in three periods, namely 2020, 2022, and 2024. Therefore, the final dataset included 255bridge-period observations, calculated as $85 \times 3 = 255$. Each bridge was assessed using a twelve-criterion maintenance evaluation matrix. The expert component of the study involved 15domain specialists, including structural and geotechnical engineers, railway operations managers, bridge inspection specialists, and maintenance planners. All experts had at least 10years of professional experience in railway infrastructure. Before the Delphi rounds, a preliminary orientation session was conducted to ensure a shared understanding of triangular fuzzy numbers, linguistic rating scales, and the evaluation procedure.

Data were collected using three main instruments: the Fuzzy Delphi expert questionnaire, the bridge inspection evaluation matrix, and the neural-network modeling dataset. The Fuzzy Delphi questionnaire was developed based on an initial list of 20candidate maintenance criteria extracted from the literature, field inspection requirements, and preliminary expert consultation. Experts evaluated the importance of each criterion using a five-level linguistic scale, which was transformed into triangular fuzzy numbers in the form $\tilde{A} = (l, m, u)$, where l, m, and $u$represent the lower, modal, and upper fuzzy values, respectively. The linguistic categories included very low importance $(0.0, 0.1, 0.3)$, low importance $(0.1, 0.3, 0.5)$, moderate importance $(0.3, 0.5, 0.7)$, high importance $(0.5, 0.7, 0.9)$, and very high importance $(0.7, 0.9, 1.0)$. The Fuzzy Delphi process was implemented in three iterative rounds, and criteria with defuzzified values below 0.50were removed. This process reduced the initial 20criteria to 12final maintenance criteria.

The bridge inspection evaluation matrix was used to assess each bridge according to the final twelve maintenance criteria. These criteria included foundation or pier condition, bridge deck condition, side walls, gabion or riprap condition, expansion joints, bearing seats, drainage culverts, subsurface drainage, surface drainage, protective coating, lighting system, and guardrails. Each criterion was scored on a five-point condition scale from 1to 5, where 1indicated very good condition and 5indicated very poor condition. Since all criteria reflected deterioration-related indicators, they were treated as cost-type criteria, meaning that higher values represented poorer bridge condition and greater need for intervention. The raw inspection scores were normalized to the interval $[0, 1]$and then fuzzified using an uncertainty band of $\pm 0.05$to represent the uncertainty inherent in field inspection judgments.

The neural-network dataset was prepared using the weighted structural criteria and the outputs of the Fuzzy TOPSIS stage. The main outcome variable was the Intervention Priority Index, Cl^+, which represented the relative closeness of each bridge to the most critical deterioration condition. The MLP input vector consisted of 13variables: the lagged intervention priority index $Cl^+(t-1)$and the twelve weighted maintenance criteria $C_1$to C_{12}. The target output was $Cl^+(t+1)$, representing the predicted intervention priority index for the next inspection cycle. This structure allowed the model to incorporate both the previous deterioration status of each bridge and its current structural condition.

Data analysis was conducted in three sequential stages. In the first stage, the Fuzzy Delphi Method was used to determine the final maintenance criteria and their relative weights. The triangular fuzzy judgments of the experts were aggregated by calculating the arithmetic mean of individual expert responses. The defuzzified value of each criterion was calculated using the weighted average formula $BN_j = \frac{l_j + 2m_j + u_j}{4}$, where $BN_j$is the defuzzified value of criterion j, and l_j, m_j, and $u_j$are the lower, modal, and upper fuzzy values. The normalized weight of each criterion was then calculated as $w_j = \frac{BN_j}{\sum_{j=1}^{12} BN_j}$. The Delphi process was terminated when the change in mean expert judgment between two consecutive rounds was less than ± 0.15, indicating convergence. Cronbach's alpha was used to assess internal reliability, and Kendall's coefficient of concordance was used to evaluate agreement among experts.

In the second stage, Fuzzy TOPSIS was used to calculate the Intervention Priority Index for all bridges across the three inspection periods. The fuzzy decision matrix was defined as $X^{(t)} = [x_{ijt}]_{n \times m}$, where $i = 1, \dots, 85$represents bridges, $j = 1, \dots, 12$represents maintenance criteria, and $t = 1,2,3$represents inspection periods. Each matrix element was expressed as $x_{ijt} = (l_{ijt}, m_{ijt}, u_{ijt})$. After normalization, the weighted normalized fuzzy matrix was obtained as $v_{ijt} =$

$r_{ijt} \times w_j$, where $r_{ijt}$is the normalized fuzzy value and $w_j$is the normalized criterion weight. Because all criteria were cost-type indicators, the fuzzy positive ideal solution represented the worst or most critical bridge condition, while the fuzzy negative ideal solution represented the best or healthiest bridge condition.

The distance between two triangular fuzzy numbers was calculated as $d(\tilde{a}, \tilde{b}) = \sqrt{\frac{1}{3}[(l_a - l_b)^2 + (m_a - m_b)^2 + (u_a - u_b)^2]}$. The intervention priority index was then calculated as $Cl_{it}^{+} = \frac{D_{it}^{-}}{D_{it}^{+} + D_{it}^{-}}$, where $D_{it}^{+}$is the distance from the fuzzy positive ideal solution and $D_{it}^{-}$is the distance from the fuzzy negative ideal solution. A value of Cl^{+}close to 1indicated severe deterioration and high maintenance priority, whereas a value close to 0indicated better condition and lower priority. To reflect the progressive nature of deterioration, temporal weighting was applied across inspection periods using the condition $\lambda_3 > \lambda_2 > \lambda_1$and $\sum_{t=1}^{3} \lambda_t = 1$, assigning greater importance to more recent inspection data. Ranking stability was examined using Monte Carlo simulation with 1,000iterations, in which criterion weights were perturbed by $\pm 10\%$, and the resulting rankings were compared with the baseline ranking using Spearman's rank correlation coefficient.

In the third stage, an MLP neural network was developed to predict future values of Cl^{+}. The selected architecture was 13-64-1, consisting of 13input neurons, one hidden layer with 64neurons, and one output neuron. The hidden layer used the ReLU activation function, and the output layer used a linear activation function to preserve the continuous [0، 1]range of the predicted Cl^{+}values. The model was trained using the Adam optimizer with an initial learning rate of 0.001and a batch size of 16. Xavier/Glorot initialization was used for initializing network weights. To prevent overfitting, dropout with a rate of 20%, L2 regularization with $\lambda = 0.001$, and early stopping with a patience value of 50epochs and a maximum of 500epochs were applied.

Model validation was performed using Walk-Forward Validation with an expanding training window. In the first step, the model was trained on the 85observations from 2020 and tested on the 85observations from 2022. In the second step, the training set was expanded to 170observations from 2020 and 2022, and the model was tested on the 85observations from 2024. Finally, the model was retrained using all 255observations to forecast the Cl^{+}values for the 2026 inspection period. Predictive uncertainty was estimated using Monte Carlo Dropout by keeping dropout active during inference and performing $T = 100$forward passes for each bridge. The point prediction was calculated as $\mu_i = \frac{1}{T}\sum_{t=1}^{T} f_t(x_i)$, predictive uncertainty was calculated as $\sigma_i = \text{std}\{f_1(x_i), \ldots, f_T(x_i)\}$, and the 95%confidence interval was calculated as $CI_{95} = (\mu_i - 1.96\sigma_i, \mu_i + 1.96\sigma_i)$.

The predictive performance of the MLP model was evaluated using the coefficient of determination, mean absolute error, root mean square error, Spearman's rank correlation coefficient, and Top-10 accuracy. The coefficient of determination was calculated as $R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$, mean absolute error as $MAE = \frac{1}{n}\sum_i | y_i - \hat{y}_i |$, root mean square error as $RMSE = \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2}$, and Spearman's rank correlation coefficient as $\rho = 1 - \frac{6\sum_i d_i^2}{n(n^2-1)}$. Top-10 accuracy was defined as the degree of overlap between the actual and predicted ten bridges with the highest maintenance priority.

3. Findings and Results

Stage 1 Results: Fuzzy Delphi

The three-round Fuzzy Delphi process reduced the initial pool of 20candidate criteria to 12key maintenance criteria $(C_1 - C_{12})$through a structured, expert-driven screening process. Eight criteria $(C_{13} - C_{20})$, including climatic conditions, bridge age, strategic importance, and traffic volume, were eliminated because their mean defuzzified values were below the threshold of 0.50. This indicated expert consensus that these factors, although contextually relevant, should be treated as contextual modifiers rather than primary maintenance assessment criteria for precast box-girder railway bridges. The final criteria, their dimensions, triangular fuzzy numbers, defuzzified values, and normalized weights are shown in Table 1.

Table 1

Final maintenance criteria with fuzzy weights and normalized weights from Fuzzy Delphi

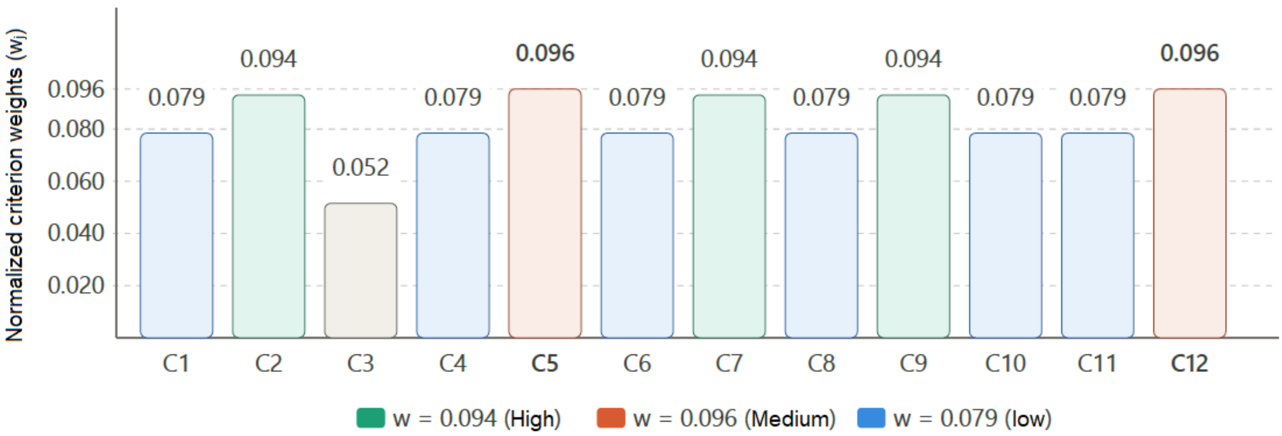
Code	Criterion	Dimension	TFN (l, m, u)	BN_j	w_j
C_1	Foundation (Pier)	Structural	(0.52, 0.78, 0.98)	0.77	0.079
C_2	Bridge Deck	Structural	(0.73, 0.96, 1.00)	0.91	0.094
C_3	Side Walls	Structural	(0.28, 0.53, 0.76)	0.52	0.052
C_4	Gabion / Riprap	Structural	(0.52, 0.78, 0.98)	0.77	0.079
C_5	Expansion Joints	Structural	(0.77, 0.99, 1.00)	0.94	0.096
C_6	Bearing Seats	Structural	(0.52, 0.78, 0.98)	0.77	0.079
C_7	Drainage Culverts	Operational	(0.73, 0.96, 1.00)	0.91	0.094
C_8	Subsurface Drainage	Operational	(0.52, 0.78, 0.98)	0.77	0.079
C_9	Surface Drainage	Operational	(0.73, 0.96, 1.00)	0.91	0.094
C_{10}	Protective Coating	Environmental	(0.52, 0.78, 0.98)	0.77	0.079
C_{11}	Lighting System	Safety	(0.52, 0.78, 0.98)	0.77	0.079
C_{12}	Guardrails	Safety	(0.77, 0.99, 1.00)	0.94	0.096
Total	—	—	—	9.72	1.000

The Fuzzy Delphi process converged in three rounds. The final validation metrics were Cronbach's $\alpha = 0.847$, indicating excellent reliability, and Kendall's $W = 0.69$, indicating excellent concordance. These values confirmed that the 15-member expert panel achieved a statistically robust and internally consistent consensus regarding the relative importance of the 12maintenance criteria.

Expansion joints $(C_5)$and guardrails (C_{12})jointly received the highest normalized weight $\left(w_j = 0.096\right)$. The prioritization of expansion joints reflects their role as a primary pathway for moisture and chloride-ion ingress into structural members, which is highly relevant in the high-humidity and freeze–thaw-active environment of the Northern railway corridor. In contrast, side walls (C_3)received the lowest weight $\left(w_j = 0.052\right)$. This lower weighting does not indicate structural irrelevance; rather, it reflects the fact that side walls in precast box-girder bridges do not carry primary bending loads and that their deterioration is usually visually evident and rapidly detectable during routine inspections.

Figure 1

Normalized criterion weights (w_j)for C_1-C_12obtained from the Fuzzy Delphi process



Stage 2 Results: Fuzzy TOPSIS

The Fuzzy TOPSIS algorithm was applied to all 85bridges across the three inspection periods, producing 255values of the Intervention Priority Index (CI^+). Table 2 presents the Fuzzy TOPSIS ranking results for selected bridges in Period 3, corresponding to the 2024 inspection cycle.

Table 2

Fuzzy TOPSIS results: Cl^+index and priority ranking for selected bridges in Period 3 (2024)

Bridge	d^{+}	d^{-}	Cl^{+}	Rank	Priority Group
B1	0.0352	0.0089	0.826	1	Critical
B2	0.0341	0.0091	0.812	2	Critical
B53	0.0328	0.0094	0.798	3	Critical
B4	0.0318	0.0127	0.706	4	High
B27	0.0301	0.0141	0.689	5	High
B36	0.0284	0.0152	0.651	6	High
…	…	…	…	7–84	Moderate / Low
B48	0.0196	0.0271	0.214	85	Low

The descriptive statistics of Cl^{+}values across the 85bridges in Period 3 showed that the minimum value belonged to B48 ($Cl^{+} = 0.214$), while the maximum value belonged to B1 ($Cl^{+} = 0.826$). The mean Cl^{+}value was 0.523, the standard deviation was 0.142, the first quartile was 0.412, the median was 0.531, the third quartile was 0.642, and the interquartile range was 0.230. The near-symmetrical distribution, reflected by the closeness of the mean $(0.523)$and median (0.531), indicated that the prioritization framework differentiated bridges across the full range of condition severity without artificial clustering in a specific Cl^{+}range.

Bridge B1 ($Cl^{+} = 0.826$)was consistently ranked first across all three inspection periods, indicating persistent and structurally embedded deterioration requiring major intervention. Bridge B48 ($Cl^{+} = 0.214$)maintained the lowest priority level across all periods, reflecting its relatively favorable structural condition.

Analysis of Cl^{+}values across the three inspection periods identified three distinct temporal behavioral patterns: stable-critical, oscillatory, and declining. The temporal evolution of selected bridges is shown in Table 3.

Table 3

Temporal evolution of Cl^+for selected bridges across three inspection periods

Bridge	Cl^{+}(2020)	Cl^{+}(2022)	Cl^{+}(2024)	Behavioral Pattern
B1	0.826	0.829	0.826	Stable (persistently critical)
B53	0.812	0.814	0.812	Stable (persistently critical)
B4	0.798	0.743	0.768	Oscillatory (temporary repair, re-deterioration)
B2	0.706	0.745	0.709	Oscillatory (moderate fluctuation)
B27	0.689	0.653	0.634	Declining (successful maintenance)

The stable pattern observed in B1 and B53 indicates deeply embedded structural deterioration requiring major rather than incremental intervention. The oscillatory pattern observed in B4 reflects the limitations of point-repair strategies, such as crack injection or temporary component replacement, which may provide short-term improvement but fail to prevent re-deterioration. The declining trajectory observed in B27 confirms that systematic and proactive maintenance can successfully reverse deterioration trends.

Monte Carlo simulation with $1{,}000$iterations and $\pm 10\%$criterion-weight perturbation confirmed that the top five critical bridges, namely B1, B2, B53, B4, and B27, remained unchanged across all scenarios. Spearman's rank correlation coefficient was $\rho = 0.96$with $p < 0.001$, demonstrating that the ranking was robust to plausible variations in expert judgment and was not sensitive to minor weight changes.

Table 4

Monte Carlo stability analysis results for top-ranked bridges across 1,000 iterations

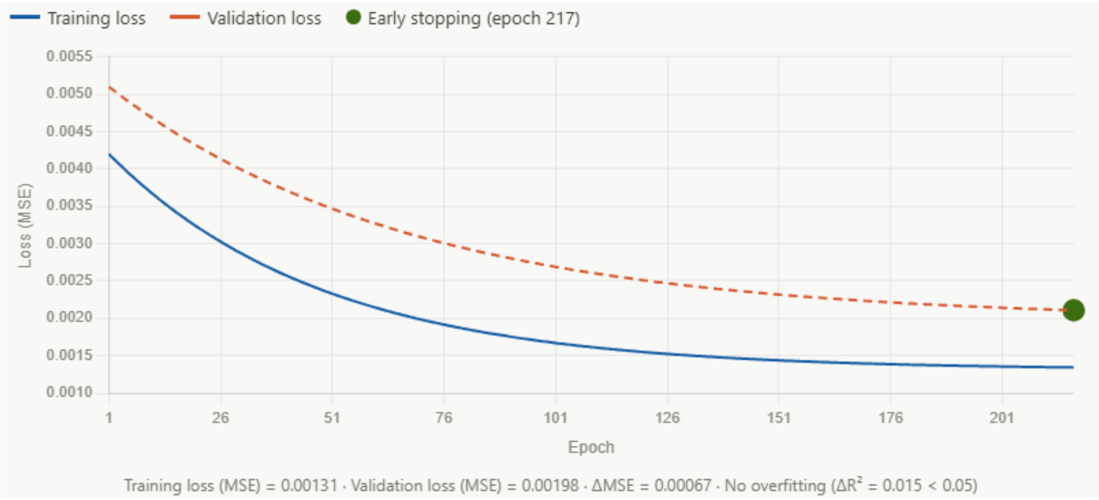
Bridge	Mean Cl^{+}(MC)	Std. Dev. (MC)	95% CI	Rank Stability
B1	0.824	0.018	[0.789, 0.859]	100% — Rank 1
B2	0.810	0.019	[0.773, 0.847]	98% — Rank 2
B53	0.796	0.022	[0.753, 0.839]	95% — Rank 3
B4	0.704	0.031	[0.643, 0.765]	87% — Rank 4
B27	0.687	0.034	[0.620, 0.754]	83% — Rank 5

Stage 3 Results: MLP Neural Network

The MLP model achieved stable convergence in both Walk-Forward Validation steps, with Early Stopping activated before the maximum epoch limit. In Step 1, the model was trained using 85observations from 2020 and stopped at epoch 193, with train loss $(MSE) = 0.00148$, validation loss $(MSE) = 0.00219$, and $\Delta MSE = 0.00071$. In Step 2, the model was trained using 170observations from 2020 and 2022 and stopped at epoch 217, with train loss $(MSE) = 0.00131$, validation loss $(MSE) = 0.00198$, and $\Delta MSE = 0.00067$. The consistently small ΔMSE, approximately 0.00069, between training and validation loss confirmed the absence of overfitting. The ΔR^2criterion for overfitting detection was also satisfied in both steps, with $\Delta R^2 = 0.015 < 0.05$.

Figure 2

Training and validation loss (MSE) curves for the MLP model: Walk-Forward Validation Step 2



The predictive performance of the MLP model across all evaluation metrics is presented in Table 5.

Table 5

Summary of MLP model performance metrics

Metric	Value	Threshold (Excellent)	Threshold (Acceptable)	Status
R^2(Test Set)	0.912	≥ 0.85	≥ 0.80	Excellent
R^2(WFV Average)	0.930	≥ 0.85	≥ 0.80	Excellent
MAE (Test Set)	0.039	<0.05	<0.08	Excellent
RMSE (Test Set)	0.051	<0.06	<0.08	Excellent
RMSE (WFV Average)	0.0218	<0.06	<0.08	Excellent
Spearman ρ	> 0.90	≥ 0.90	≥ 0.85	Excellent
Top-10 Accuracy	$> 80\%$	$\geq 80\%$	$\geq 70\%$	Excellent
ΔR^2(Train–Test)	0.015	<0.05	<0.05	No overfitting
SD(R^2) across 10 runs	<0.02	<0.02	—	Stable

The $R^2 = 0.931$achieved in Walk-Forward Validation was higher than the $R^2 = 0.912$obtained from the simple test split. This result can be explained by the expanded training set in WFV Step 2, where 170observations provided richer information about the deterioration space compared with the 85observations available in Step 1, thereby improving model generalization.

The relative importance of input variables was assessed using one-at-a-time sensitivity analysis, in which each input variable was replaced by its mean value and the resulting decrease in $R^2$was recorded. The lag-1 temporal feature $Cl^+(t-1)$was the most important predictor, with normalized importance of 0.180. Among the structural criteria, C_5, representing expansion joints, ranked second overall and first among structural variables, with normalized importance of 0.114and Cl^+change of ± 0.021. This was followed by C_{12}, representing guardrails, with normalized importance of 0.109and Cl^+change of ± 0.019; C_1, representing foundation, with normalized importance of 0.103and Cl^+change of ± 0.018; and C_{10}, representing protective coating, with normalized importance of 0.098and Cl^+change of ± 0.016. The remaining variables, $C_2 - C_9$and C_{11}, had normalized importance values ranging from 0.052to 0.094and Cl^+changes ranging from $\pm 0.006$to ± 0.014.

The identification of $Cl^+(t-1)$as the single most important predictor confirmed the significance of deterioration memory effects in bridge condition evolution. Moreover, the fact that C_5, expansion joints, was also the most important structural criterion in the MLP model independently corroborated the Fuzzy Delphi finding that expansion joints were among the highest-weighted maintenance criteria. When $Cl^+(t-1)$was removed from the model, R^2decreased by 9.0%, confirming the essential role of temporal memory in predicting bridge deterioration.

Robustness testing further confirmed the stability of the MLP model. In the baseline Walk-Forward Validation scenario, the model achieved $R^2 = 0.931$, $MAE = 0.0216$, and $RMSE = 0.0170$. When 40%of the training data were randomly removed, leaving 60%of the training data, the model still achieved $R^2 = 0.912$, $MAE = 0.0242$, and $RMSE = 0.0191$, corresponding to an R^2drop of only 1.9%. When $\pm 5\%$input noise was injected to simulate inspection error, the model achieved $R^2 = 0.908$, $MAE = 0.0250$, and $RMSE = 0.0198$, with an R^2drop of 2.5%. When $Cl^+(t-1)$was removed and only the 12structural inputs were retained, the model achieved $R^2 = 0.847$, $MAE = 0.0381$, and $RMSE = 0.0294$, with an R^2drop of 9.0%. These findings show that the model generalized well beyond the training sample and remained tolerant of typical inspection measurement errors.

The final MLP model, trained on all 255observations, generated Cl^+predictions for 2026 with 95%confidence intervals for all 85bridges. The highest-priority predicted bridges and their recommended maintenance actions are presented in Table 6.

Table 6

Predicted Cl^+values and maintenance recommendations for 2026

Rank	Bridge	$\mu(Cl^+)$	95% CI	σ	Recommended Action
1	B1	0.828	[0.787, 0.869]	0.021	Immediate major intervention
2	B2	0.813	[0.770, 0.856]	0.022	Immediate major intervention
3	B53	0.772	[0.698, 0.846]	0.038	Immediate inspection + intervention
4	B4	0.713	[0.644, 0.782]	0.035	Targeted maintenance, 6–12 months
5	B27	0.618	[0.536, 0.700]	0.042	Targeted maintenance, 6–12 months
85	B48	0.214	[0.179, 0.249]	0.018	Routine monitoring per schedule

The predictive uncertainty values showed a meaningful relationship with temporal behavior patterns. Bridges with oscillatory historical behavior, such as B53 ($\sigma = 0.038$)and B27 ($\sigma = 0.042$), showed higher predictive uncertainty than bridges with stable deterioration trajectories, such as B1 ($\sigma = 0.021$). This finding indicates that the Monte Carlo Dropout mechanism captured meaningful structural uncertainty rather than random noise. From a maintenance-management perspective, bridges with higher σvalues require more intensive pre-intervention inspection before final budget allocation.

Figure 3

Predicted vs. actual CI^+values from the MLP model (Walk-Forward Validation, Step 2), with 95% MC Dropout confidence bands

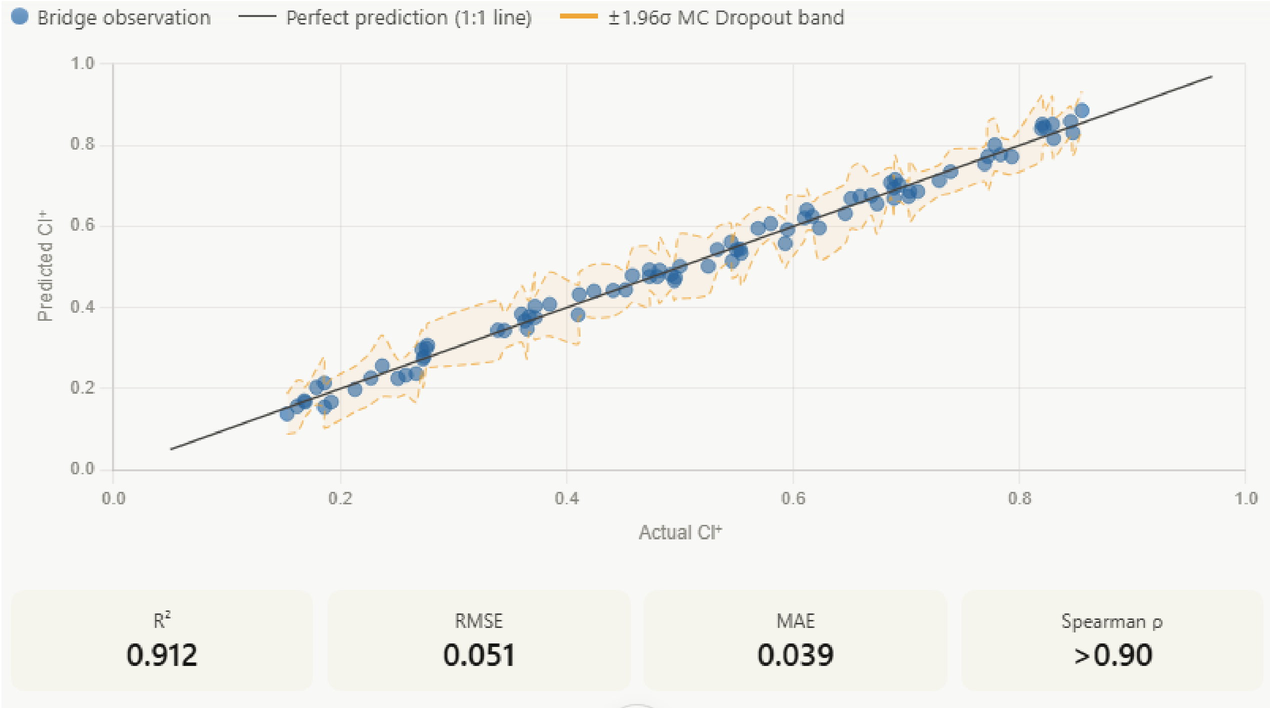
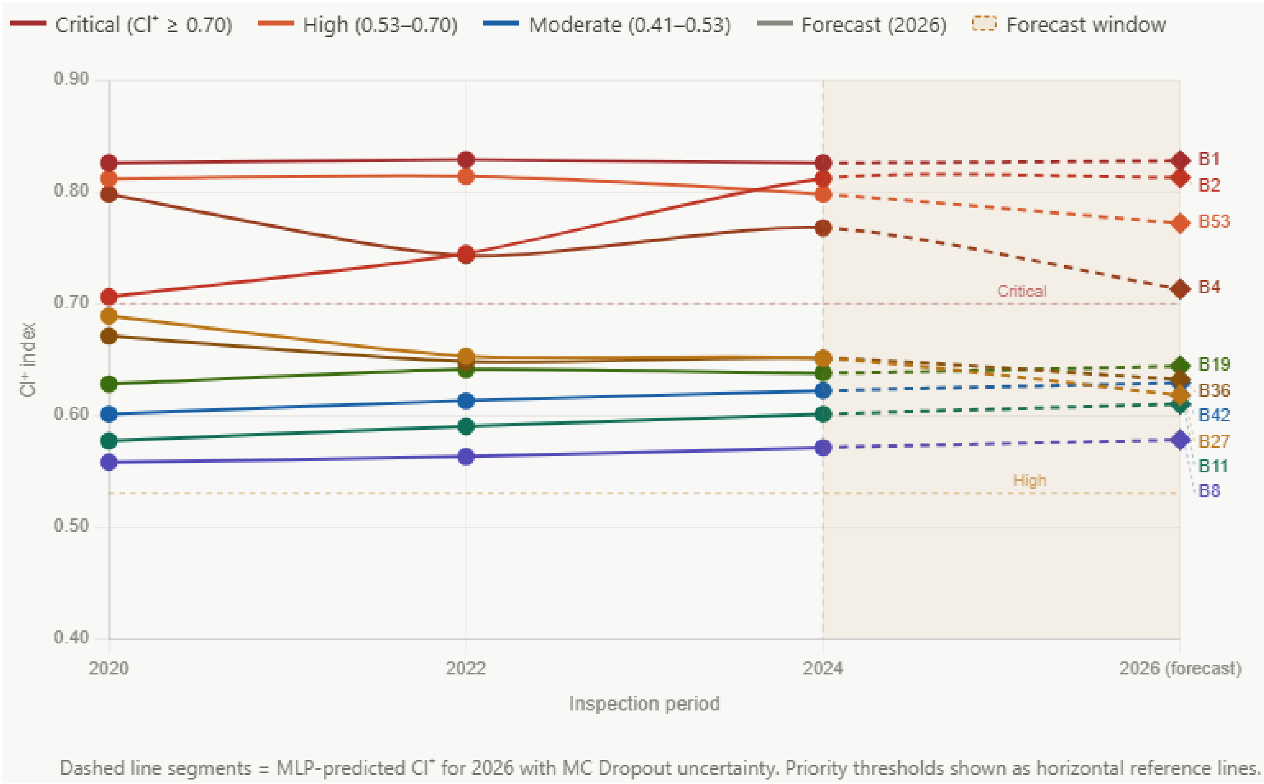


Figure 4

Temporal evolution of CI^+for the 10 highest-priority bridges across three inspection periods and 2026 forecast



The results of the three-stage framework demonstrate a coherent and internally consistent chain of knowledge transfer from expert consensus to quantitative prioritization and predictive forecasting. The Fuzzy Delphi stage identified the most important maintenance criteria, the Fuzzy TOPSIS stage transformed these weighted criteria into bridge-level priority rankings, and the MLP stage extended the framework toward predictive maintenance by forecasting future Cl^{+}values.

The convergent validity of the framework is most clearly reflected in the repeated importance of expansion joints (C_5). This criterion received one of the highest weights in the expert-based Fuzzy Delphi stage and also emerged as the most important structural variable in the data-driven MLP sensitivity analysis. Thus, two different methodological approaches, expert judgment and neural-network pattern recognition, independently identified the same deterioration driver as central to bridge maintenance prioritization.

Compared with conventional prioritization based on individual expert judgment, AHP-based procedures, or standalone Fuzzy TOPSIS models, the proposed framework provides several analytical advantages. These include uncertainty propagation through Monte Carlo Dropout, temporal memory integration through lag-1 Cl^{+}features, forward-looking decision support through validated 2026 predictions, and strong robustness under data reduction and input noise. The identification of stable-critical, oscillatory, and declining temporal patterns also has direct operational significance. In particular, the oscillatory pattern observed in bridges such as B4 demonstrates the limitations of isolated reactive repairs and supports the need for systematic, root-cause-oriented maintenance planning in railway bridge management.

4. Discussion and Conclusion

The present study developed and validated an integrated expert system for prioritizing the maintenance of precast concrete railway bridges by combining the Fuzzy Delphi Method, Fuzzy TOPSIS, and a Multilayer Perceptron neural network. The results demonstrated that the proposed framework could transform expert-based maintenance knowledge into quantitative bridge priority rankings and then extend those rankings into predictive maintenance forecasts. This integrated structure responds directly to the central challenge of bridge management: maintenance resources are limited, while deterioration risks are multidimensional, uncertain, and temporally evolving. The findings showed that expert consensus, fuzzy multi-criteria prioritization, and neural-network forecasting can operate as a coherent decision-support chain for railway bridge maintenance. This result is consistent with life-cycle bridge management perspectives emphasizing that infrastructure decisions must go beyond isolated repair actions and should instead integrate condition assessment, uncertainty management, optimization, and long-term performance planning (Bocchini & Frangopol, 2012; Frangopol, 2011; Ryall, 2009).

The Fuzzy Delphi stage reduced the initial set of 20 candidate maintenance criteria to 12 final criteria, indicating that the expert panel distinguished between primary maintenance indicators and contextual modifiers. The elimination of criteria such as climatic conditions, bridge age, strategic importance, and traffic volume does not imply that these factors are irrelevant to bridge management; rather, it suggests that experts regarded them as background or modifying variables rather than direct inspection criteria for the specific type of bridges under study. This finding is methodologically important because decision-support models can become unstable or operationally inefficient when too many overlapping or indirectly relevant criteria are included. The successful convergence of the Delphi process in three rounds, together with Cronbach's $\alpha = 0.847$and Kendall's $W = 0.69$, indicates strong internal reliability and consensus among the 15 experts. This supports previous work showing that fuzzy expert-based procedures can structure professional judgment and reduce ambiguity in complex infrastructure decision problems (Kahraman, 2008; Rogulj et al., 2021; Wang et al., 2014).

Among the final maintenance criteria, expansion joints $(C_5)$and guardrails (C_{12})received the highest weights $\left(w_j = 0.096\right)$, while side walls (C_3)received the lowest weight $\left(w_j = 0.052\right)$. The high importance assigned to expansion joints is technically meaningful because these components often function as the main route for moisture, chloride ions, and debris to enter the bridge system. In aggressive environmental corridors, especially those exposed to humidity, precipitation, dissolved salts, and freeze–thaw cycles, defective joints can accelerate deck deterioration, reinforcement corrosion, bearing-seat damage, and drainage-related problems. This result is consistent with concrete durability theory, which emphasizes the role of moisture transport, permeability, cracking, and chloride ingress in the deterioration of reinforced concrete structures (Mehta & Monteiro, 2014). The high weight assigned to guardrails also reflects the safety-critical function of railway bridge components that directly influence operational protection. Conversely, the lower weight of side walls can be interpreted in relation to their limited role in primary bending resistance and their relatively visible deterioration pattern during routine inspection. Therefore, the weighting structure was not only statistically coherent but also consistent with engineering logic.

The Fuzzy TOPSIS stage generated 255 values of the Intervention Priority Index (CI^{+})across 85 bridges and three inspection periods. The 2024 results showed that B1 ($CI^{+} = 0.826$), B2 ($CI^{+} = 0.812$), and B53 ($CI^{+} = 0.798$)were the most critical bridges, while B48 ($CI^{+} = 0.214$)had the lowest deterioration priority. The near-symmetrical distribution of CI^{+}values, with mean 0.523and median 0.531, suggests that the model produced differentiated rankings rather than forcing bridges into artificial clusters. This is a key strength of Fuzzy TOPSIS because it measures each bridge in relation to both ideal and anti-ideal deterioration states. Previous applications of fuzzy TOPSIS in bridge risk and condition assessment have similarly shown that fuzzy distance-based ranking is effective when criteria values are uncertain and condition information is partly linguistic or inspection-based (Li et al., 2025; Qu et al., 2025; Wang & Elhag, 2006). The present findings extend this evidence to precast concrete railway bridges by showing that Fuzzy TOPSIS can produce interpretable intervention priorities across multiple inspection periods.

The temporal analysis revealed three behavioral patterns: stable-critical, oscillatory, and declining. Bridges B1 and B53 showed stable-critical behavior, indicating persistent deterioration and the need for major intervention rather than limited or incremental repair. B4 and B2 showed oscillatory patterns, suggesting temporary improvement followed by renewed deterioration, likely reflecting the limitations of point-repair strategies. B27 showed a declining CI^{+}trajectory, suggesting that systematic maintenance can reduce intervention priority over time. These temporal patterns are important because two bridges with similar current CI^{+}values may require different management strategies depending on deterioration history. This finding aligns with the broader bridge management literature, which emphasizes that maintenance planning should account for deterioration trajectories, life-cycle behavior, and uncertainty rather than relying only on current condition ratings (Alshibani et al., 2023; Bocchini & Frangopol, 2012; Frangopol, 2011). The results also support the argument that resilient infrastructure management must distinguish between structures requiring immediate rehabilitation and those responsive to preventive maintenance (Bocchini & Frangopol, 2012).

The ranking stability analysis further confirmed the robustness of the Fuzzy TOPSIS results. Monte Carlo simulation with 1,000 iterations and $\pm 10\%$weight perturbation showed that the top five critical bridges remained unchanged, with Spearman's $\rho = 0.96$and $p < 0.001$. This finding is significant because maintenance prioritization systems must remain stable under reasonable uncertainty in expert weights. If small changes in weights produced completely different rankings, the framework would be unsuitable for operational decision-making. The observed robustness supports the use of sensitivity analysis as a necessary validation component in infrastructure decision-support systems (Saltelli et al., 2009). It also strengthens confidence that the identified top-priority bridges were not artifacts of a specific weight configuration but represented consistently critical cases across plausible expert-judgment scenarios.

The MLP neural network results demonstrated strong predictive capability. The selected 13-64-1 architecture achieved excellent model performance, including $R^2 = 0.912$in the test set, average $R^2 = 0.930$in Walk-Forward Validation, $MAE = 0.039$, and $RMSE = 0.051$. These results indicate that the model successfully learned the nonlinear relationship between lagged priority status, weighted structural criteria, and future intervention priority. This finding is consistent with the established capacity of neural networks to approximate complex nonlinear functions in engineering systems (Goodfellow et al., 2016; Haykin, 2007). It also aligns with previous bridge deterioration prediction studies showing that machine learning can improve condition forecasting and maintenance optimization when sufficient inspection and condition variables are available (Ghafoori et al., 2024; Rashidi-Nasab & Elzarka, 2023). The use of appropriate initialization, regularization, and validation also reflects recommendations from statistical and deep learning literature regarding generalization and model stability (Glorot & Bengio, 2010; Hastie et al., 2009).

The use of Walk-Forward Validation was one of the main methodological strengths of the study. In time-dependent bridge inspection data, random train-test splitting may produce overly optimistic results by allowing information from future inspection periods to influence model training. The present model avoided this problem by training on earlier periods and testing on later periods, thereby simulating real forecasting conditions. This approach is consistent with time-series forecasting principles, which emphasize chronological validation when observations are temporally ordered (Box et al., 2015; Hyndman & Athanasopoulos, 2021). The finding that Walk-Forward Validation achieved $R^2 = 0.931$, exceeding the simple test-set result, can be explained by the expanded training window, which allowed the model to learn a richer deterioration space from both 2020 and 2022 data before predicting 2024 outcomes. This supports the suitability of expanding-window validation for infrastructure datasets that accumulate over sequential inspection cycles.

The variable-importance analysis provided further evidence for the internal validity of the framework. The lag-1 temporal feature $Cl^{+}(t-1)$was the strongest predictor, with normalized importance of 0.180, and its removal reduced $R^2$by 9.0%. This finding confirms that bridge deterioration has a memory effect: previous condition strongly influences future condition. Such a result is theoretically consistent with time-series modeling and deterioration processes, where current structural state is partly determined by accumulated past damage (Box et al., 2015; Hyndman & Athanasopoulos, 2021). Importantly, expansion joints (C_5)emerged as the most important structural criterion in the MLP model, with normalized importance of 0.114. This independently corroborates the Fuzzy Delphi result, in which expansion joints also received one of the highest expert weights. The convergence between expert judgment and data-driven neural-network sensitivity analysis is one of the strongest findings of this study, as it demonstrates that both human expertise and predictive modeling identified the same key deterioration driver.

The robustness tests confirmed the practical reliability of the MLP model. When 40% of the training data were removed, the R^2value decreased by only 1.9%; when $\pm 5\%$input noise was added to simulate inspection error, the R^2drop was only 2.5%. These findings suggest that the model was not overly dependent on perfect data quality or the full training sample. This is particularly important in bridge management, where inspection data may include measurement uncertainty, subjective scoring variation, and missing or imperfect field records. The model's tolerance to input noise supports its potential operational use in real railway maintenance environments. At the same time, the larger performance drop after removing $Cl^{+}(t-1)$confirms that temporal information is not optional but essential for accurate forecasting. This reinforces the broader methodological argument that predictive maintenance models should integrate both current condition indicators and historical deterioration trends (Gao et al., 2023; Ghafoori et al., 2024).

The 2026 forecast provided actionable maintenance recommendations. B1 and B2 were predicted to remain in the immediate major intervention category, while B53 required immediate inspection and intervention. B4 and B27 were placed in the targeted maintenance category within 6–12 months, and B48 remained suitable for routine monitoring. These results demonstrate the operational value of combining prioritization and prediction: rather than producing a static ranking, the framework offers a forward-looking maintenance plan. The use of Monte Carlo Dropout also added a risk-informed dimension by producing uncertainty estimates and confidence intervals for each prediction. Bridges with oscillatory behavior, such as B53 and B27, showed higher predictive uncertainty than stable bridges such as B1. This supports the relevance of uncertainty quantification in neural-network-based decision support (Gal & Ghahramani, 2016). It also aligns with the growing emphasis on interpretable and trustworthy machine learning, where decision-makers need not only predictions but also information about uncertainty and model reasoning (Doshi-Velez & Kim, 2017).

Overall, the findings indicate that the proposed Fuzzy Delphi–Fuzzy TOPSIS–MLP framework provides a scientifically defensible and operationally relevant decision-support system for railway bridge maintenance. It improves on conventional expert-only approaches by formalizing expert consensus; improves on standalone MCDM methods by incorporating temporal prediction; and improves on purely data-driven models by grounding prediction in interpretable maintenance criteria. This integrated design is consistent with the current direction of infrastructure management research, which increasingly combines expert knowledge, multi-criteria decision-making, machine learning, uncertainty analysis, and budget-conscious prioritization (Alshibani et al., 2023; Ghafoori et al., 2024; Li et al., 2025; Rashidi-Nasab & Elzarka, 2023). The framework is especially useful for railway corridors exposed to aggressive environmental conditions, where deterioration is accelerated and maintenance resources must be allocated with precision, transparency, and predictive foresight.

Several limitations should be considered when interpreting the findings of this study. First, the dataset was limited to 85 precast concrete railway bridges located along a single railway corridor and included three inspection periods. Although this dataset was sufficient for developing and validating the proposed framework, broader validation across additional corridors, climatic zones, bridge types, and longer inspection histories would strengthen the generalizability of the findings. Second, the expert panel consisted of 15 specialists with experience in railway infrastructure, but their judgments were necessarily influenced by the operational and environmental context of the studied corridor. Third, the model relied on periodic inspection data rather than continuous structural health monitoring data, which may limit its ability to capture short-term deterioration events or sudden changes in bridge condition between inspection cycles.

Future studies should apply the proposed framework to larger and more diverse bridge inventories, including different railway corridors, highway bridges, climatic regions, and structural systems. Additional research should also examine the integration of real-time structural health monitoring data, such as strain, vibration, corrosion, moisture, and temperature sensors, into the predictive component of the framework. Longer inspection time series would allow researchers to test more advanced temporal models and compare MLP performance with recurrent neural networks, long short-term memory models, transformer-based architectures, and hybrid digital-twin platforms. Future work may also incorporate life-cycle cost, budget constraints, maintenance crew availability, and traffic disruption costs to transform the priority index into a full maintenance scheduling and optimization system.

For practical bridge management, the proposed framework can be used as a structured decision-support tool to classify bridges into intervention priority levels and support transparent budget allocation. Railway maintenance authorities should pay particular attention to bridges with persistently high Cl^{+}values, as these cases are more likely to require major intervention rather than temporary repair. Bridges with oscillatory behavior should be inspected more carefully before maintenance budgets are assigned, because their condition may reflect unresolved root causes rather than isolated defects. The framework also suggests that expansion joints, drainage-related components, protective systems, and safety elements should receive systematic attention during inspection and preventive maintenance planning. By combining expert evaluation, fuzzy ranking, and predictive modeling, maintenance managers can move from reactive repair toward evidence-based and preventive infrastructure management.

Authors' Contributions

Authors contributed equally to this article.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethics Considerations

In this research, ethical standards including obtaining informed consent, ensuring privacy and confidentiality were considered.

References

Alshibani, A., Saa, A. S. A., Hassanian, M. A., Bubshait, A., & Shash, A. (2023). Decision support model for allocating maintenance budgets for bridges. *Applied Sciences*, *13*(17), 9679. https://doi.org/10.3390/app13179679

Bocchini, P., & Frangopol, D. M. (2012). Resilience as optimization criterion for the rehabilitation of bridges belonging to a transportation network subject to earthquake. Structures Congress 2011,

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). John Wiley & Sons.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *ArXiv*. https://arxiv.org/abs/1702.08608

Frangopol, D. M. (2011). Life-cycle performance, management, and optimisation of structural systems under uncertainty: Accomplishments and challenges. *Structure and Infrastructure Engineering*, *7*(6), 389-413. https://doi.org/10.1080/15732471003594427

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. Proceedings of the 33rd International Conference on Machine Learning, Gao, W., Li, Q., & Zhang, Y. (2023). Integrating AHP and MLP for predictive maintenance of railway bridges. *Expert Systems with Applications*, *215*, 119431. https://doi.org/10.1016/j.eswa.2022.119431

Ghafoori, M., Abdallah, M., & Ozbek, M. E. (2024). Machine learning-based bridge maintenance optimization model for maximizing performance within available annual budgets. *Journal of Bridge Engineering*, *29*(4), 04024011. https://doi.org/10.1061/JBENF2.BEENG-6436

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics,

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Haykin, S. (2007). *Neural Networks and Learning Machines* (3rd ed.). Prentice-Hall. https://dl.acm.org/doi/book/10.5555/1213811

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. https://doi.org/10.1007/978-3-030-39091-4

Kahraman, C. (2008). *Fuzzy Multi-Criteria Decision Making: Theory and Applications with Recent Developments*. Springer. https://doi.org/10.1007/978-0-387-76813-7

Li, Y., Deng, J., Wang, Y., Liu, H., Peng, L., Zhang, H., Liang, Y., & Feng, Q. (2025). The condition evaluation of bridges based on fuzzy BWM and fuzzy comprehensive evaluation. *Applied Sciences*, *15*(6), 2904. https://doi.org/10.3390/app15062904

Mehta, P. K., & Monteiro, P. J. M. (2014). *Concrete: Microstructure, Properties, and Materials* (4th ed.). McGraw-Hill Education. https://www.researchgate.net/publication/258848178_Concrete_Structure_Properties_and_Materials

Qu, B., Lin, S., & Huang, H. (2025). A modified fuzzy TOPSIS approach for the condition assessment of existing bridges. *Nonlinear Engineering*, *14*(1). https://doi.org/10.1515/nleng-2025-0106

Rashidi-Nasab, A. R., & Elzarka, H. (2023). Optimizing machine learning algorithms for improving prediction of bridge deck deterioration: A case study of Ohio bridges. *Buildings*, *13*(6), 1517. https://doi.org/10.3390/buildings13061517

Rogulj, K., Pamukovic, J. K., & Jajac, N. (2021). Knowledge-based fuzzy expert system to the condition assessment of historic road bridges. *Applied Soft Computing*, *11*(3), 1021. https://doi.org/10.3390/app11031021

Ryall, M. (2009). *Bridge Management* (2nd ed.). CRC Press. https://doi.org/10.1201/b12866

Saltelli, A., Chan, K., & Scott, E. M. (2009). *Sensitivity Analysis*. John Wiley & Sons.

Wang, Y., Yeo, G. T., & Ng, A. K. Y. (2014). Choosing optimal bunkering ports for liner shipping companies: A hybrid fuzzy-Delphi-TOPSIS approach. *Transport Policy*, *35*, 358-365. https://doi.org/10.1016/j.tranpol.2014.04.009

Wang, Y. M., & Elhag, T. M. S. (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. *Expert Systems with Applications*, *31*(2), 309-319. https://doi.org/10.1016/j.eswa.2005.09.040