Towards Equitable Infrastructure Asset Management: Scour Maintenance Strategy for Aging Bridge Systems in Flood-prone Zones using Deep Reinforcement Learning

Amir Taherkhani[a], Weiwei Mo[a], Erin Bell[a], and Fei Han[a]*

[a]University of New Hampshire, Durham, NH, USA

*Corresponding author

Email: fei.han@unh.edu

Postal address: W171, Kingsbury Hall, 33 Academic way, Durham, New Hampshire, 03824

Towards Equitable Infrastructure Asset Management: Scour Maintenance Strategy for Aging Bridge Systems in Flood-prone Zones using Deep Reinforcement Learning

ABSTRACT

Bridges play a critical role in transportation networks; however, they are vulnerable to deterioration, aging, and degradation, especially in the face of climate change and extreme weather events such as floodings. Furthermore, bridges can significantly affect social vulnerability; their damage or destruction can isolate communities, inhibit emergency responses, and disrupt essential services. Maintaining critical bridges in a cost-effective and sustainable manner is crucial to ensure their longevity and protect vulnerable communities. To address the maintenance optimization problem of bridge systems considering the effects of time deterioration, flood degradation, and social vulnerability, this study proposes a deep reinforcement learning algorithm to optimally allocate resources to bridges that are at expected cost of failure due to scour. The algorithm considers the effects of flood degradation with different return periods and is trained using a Markov Decision Process as the environment. The study conducts four flood simulation scenarios using Geographic Information System data. The findings suggest that the deep reinforcement learning algorithm proposes a sequence of repair actions that outperforms the status quo, currently employed by bridge managers. The significance of this study lies in its valuable insights for cities worldwide on how to effectively optimize their limited resources for the maintenance and rehabilitation of critical infrastructure systems to decrease portfolio cost and increase social equity.

Keywords: *Flood Risk Management, Deep Reinforcement Learning, Geographic Information System, Social Vulnerability, Equitable Infrastructure Management*

Highlights:

1. Employed geographic information system to conduct flood simulation and identify vulnerable bridges.
2. Incorporated social vulnerability in maintenance planning and resource allocation to enhance the living experience of socially vulnerable people.
3. Developed a Markov decision process to emulate the bridge operational life in a defined horizon.
4. Created a Reinforcement Learning framework to obtain optimized maintenance strategy for system of vulnerable bridges.

1 INTRODUCTION

Bridges are vital components of the transportation network, acting as essential conduits for societal commuting needs. Maintaining the safety and functionality of these bridge systems is vital for community well-being and interconnectedness (Zhang, Agbelie, and Labi 2015). In the United States, the Federal Highway Administration (FHWA) has reported that over 46,000 bridges, approximately 7.5% of all US bridges, are structurally deficient necessitating urgent repair or replacement (FHWA 2020). The situation of aged infrastructure is exacerbated by increasing threats from climate change and environmental extremes, heightening bridges vulnerability to deterioration (Ilbeigi and Meimand 2020; Liu and El-Gohary 2022; Manafpour et al. 2018; Sinha, Labi, and Agbelie 2017) and potential failure (Mohammadi and El-Diraby 2021; Xiong et al. 2023; Yang and Frangopol 2018; Zhang et al. 2022).

Current bridge management practices typically focus on individual bridges rather than considering the entire bridge portfolio in a region with a systematic approach. This can lead to suboptimal resource allocation, increased maintenance costs, and neglect of system-wide needs and vital social objectives. While several research efforts (Tao, Lin, and Wang 2021; Zhang and Alipour 2020; Zhang and Wang 2017) have ventured into developing systematic and optimized bridge maintenance strategies, notable gaps remain, particularly in addressing flood-induced damage and integrating social equity considerations into maintenance prioritization.

Disruptions on bridge functionality and potential expected cost of failure place a huge impact on the surrounding communities, especially the disadvantaged communities. Facing limited mobility options, financial constraints, and geographic segregation, these communities often rely heavily on public transportation and local infrastructure, including bridges, to access essential services and are particularly vulnerable to disruptions (Koks et al. 2015; Tuccillo and Spielman 2022; Xu and Guo 2022). Thus, incorporating social equity into bridge maintenance prioritization is vital for reducing the disproportionate impact of bridge failures on disadvantaged communities (Anderson, Kiddle, and Logan 2022; Plough et al. 2013; Yang et al. 2018). Addressing these challenges requires a holistic and systematic maintenance strategy that accounts for the natural aging of infrastructure, the threats from natural hazards, cost optimization, and social equity over long-term planning horizons (Abdelkader et al. 2022; Bibri and Krogstie 2017; Jaafaru and Agbelie 2022).

Recent advancements in deep reinforcement learning (DRL) offer promising avenues for addressing these complex challenges. DRL algorithms excel at solving complex Markov Decision Processes (MDPs), effectively simulating real-world scenarios. In the context of bridge system asset management, MDPs model the dynamic progression of bridge conditions (States) influenced by decision-making (Actions, such as the timing of repairs or replacements)

and external factors (environmental impacts like aging and flood damages). This framework allows for a comprehensive understanding and optimization of maintenance strategies. Table 1 summarizes the latest progress in the development of optimized maintenance strategies using variations of DRL, both for individual bridges and bridge systems. These works provide stakeholders and decision makers powerful tools for improving bridge conditions while reducing the maintenance expenditures.

Table 1- Recent advancement on optimized bridge maintenance strategy development using DRL

	Problem	Algorithm	State-space	Action-space	Reward
Wei, Bao, and Li (2020)	Maintenance for a single bridge deck deterioration	DQN	7 structural components and age	Do-Nothing, minor repair, major repair, and replacement	Action cost and expected cost of failure
Cheng and Frangopol (2021)	Optimizing load rating planning of a single bridge girder	DQN	The number of load rating factor intervals	30 interval duration as for 30 years	Action cost and expected life-cycle cost
Zhou et al. (2022)	Maintenance for a single bridge deck deterioration	Multi-agent DQN	7 structural components and age	Do-Nothing, preventive scenario, corrective scenario, and rebuild	Action cost
Yang (2022)	Maintenance for a bridge network deterioration and expected cost of failure of deterioration	PPO	5 overall quality conditions for bridges	Do-Nothing, maintenance, repair, rehabilitation, and replacement	Action cost and expected cost of failure
Yang (2022a)	Maintenance for a single steel girder bridge	DCMA2C	1 reliability index	Do-Nothing, inspect, and maintenance	Action cost
Du and Ghavidel (2022)	Maintenance for a bridge system	DQN	3 structural elements with 7 indexes based on NBI	Do-Nothing, minor repair, major repair, and replacement	Action cost and closure cost due to actions
Asghari, Biglari, and Hsu (2023)	Maintenance for a single bridge	DQN and A2C	9 states including structural and managerial features	Do-Nothing, maintenance, rehabilitation, and reconstruction	A function of utility, agency cost, and user cost
Hamida and Goulet (2023)	Maintenance of a bridge system	A2C and DQN	Variable structural components	Do-Nothing, routine maintenance, preventive maintenance, repair, and replace	Action cost

Note: DQN = Deep Q-Network; PPO= Proximal Policy Optimization; DCMA2C= Deep Centralized Multi-Agent Actor-Critic; A2C= Actor-Critic

Despite the significant advancement in these pioneering works, there remain two notable gaps to be resolved: 1) the spatial consideration of flood-induced damages to a bridge system in maintenance strategy development; and 2) the integration of social equity

considerations to ensure the development of infrastructure solutions that are just and equitable for all communities. This paper proposes a DRL-based optimization framework tailored to bridge systems in flood-prone areas, integrating considerations of aging, flood damage, economic costs, and social equity. Through this research, we aim to provide stakeholders with a robust tool for enhancing bridge resilience while ensuring fair and equitable infrastructure development for all communities.

Novelty of this work: 1) Flood Damage Consideration: It specifically addresses the gap in current bridge management strategies by considering spatial analysis of flood-induced damages. This includes modeling and anticipating the effects of flooding on bridge integrity and determining proactive maintenance strategies to mitigate these risks; 2) Social Equity Integration: Another novel aspect is the integration of social equity considerations into the maintenance prioritization process. This ensures that the developed strategies not only enhance the resilience of bridge infrastructure but also do so in a manner that is just and equitable, minimizing the disproportionate impact of bridge load limits, bridge closure, and bridge failures on vulnerable communities; 3) Comprehensive Framework: The proposed DRL-based framework systematically considers the dynamic interactions between bridge management decisions and decision outcomes, including economic costs, bridge conditions due to aging and flood damages, and social impact. The framework is adaptable to future management scenarios and climate conditions.

2 TESTBED DESCRIPTION

In this study, we use the Suffolk County (Boston area), MA, as the area testbed to develop the DRL-based decision support framework for bridge maintenance. Suffolk County is located in Massachusetts with a population of 800,000. Demography of Suffolk County is as follows: 38.8% of the people are White, 25% are Black, 22.9% are Hispanic, 9.1% are Asian,

and the rest have two or more races. 24.4% of the households have children under 18 years old, 10.2% are older than 65 in Suffolk County, and 20.6% of the people are below the poverty line (CDC/ATSDR 2022). The county has experienced an increase in both stormwater and coastal flooding, largely attributed to climate change. This increase is driven by more frequent and intense wet weather events, as well as rising sea levels affecting the coastal community (Dahl, Fitzpatrick, and Spanger-Siegfried 2017; Kirshen, Knee, and Ruth 2008; Ray and Foster 2016). The National Bridge Inventory (NBI) database (NBI 2020) provides detailed physical and functional attributes of the bridges in the study area, from which we identified 451 highway bridges in the initial analysis. From there, we further identified bridges that are the most vulnerable to flooding using the HAZUS tool to be the focus of our study. HAZUS is a program developed by the Federal Emergency Management Agency (FEMA) to assist in the assessment of potential losses from natural disasters such as hurricane, earthquake, and floods (Scawthorn et al. 2006). Four riverine flooding scenarios were modeled with return periods of 10, 50, 100, and 500 years. From these simulations through HAZUS, we can estimate the water level near each bridge during flooding and further calculate the flood depth. Figure 1 shows the results of the simulations for the four riverine flooding scenarios, revealing that stronger floods lead to wider inundation that impact more bridges.

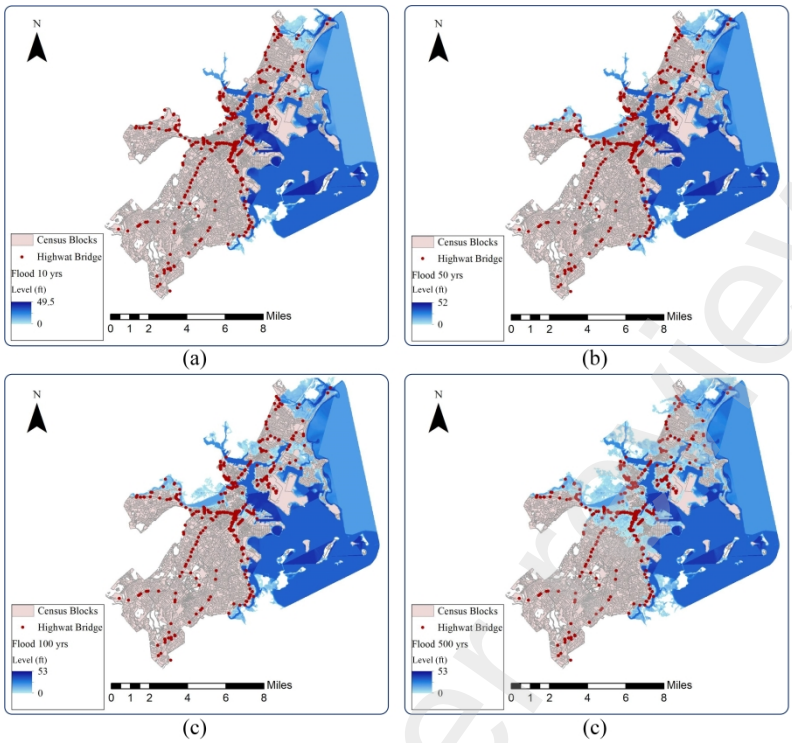

Figure 1- Flood level estimated from HAZUS flooding simulation for Sulfolk County, MA: (a) water level (measured from the ground level) during flood with 10-year return period; (b) water level during flood with 50-year return period; (c) water level during flood with 100-year return period. (d) water level during flood with 500-year return period.

Based on the flooding simulation results, there are 19 bridges that experience inundation in all four flooding scenarios, and these bridges are selected for the subsequent comparison of maintenance analyses. Table 2 summarizes the basic information of the 19 selected bridges of study.

Table 2- Basic information and scour vulnerabilities for critical bridges in Suffolk County, MA

ID	Bridge name	Average daily traffic (ADT)	Number of spans	Bridge type	Current scour rating	Bridge scour ratings after flooding with different return periods				Normalized social impact index (SII)
						10 years	50 years	100 years	500 years	
1	Saratoga Street	34100	5	Prestressed concrete continuous	8	5	5	5	5	0.63
2	Longfellow Bridge	29100	11	Steel	5	8	5	5	5	0.98
3	Harvard Bridge	32273	25	Steel continuous	8	8	5	5	5	0.99
4	Charles River Dam	43250	1	Steel	8	8	5	5	5	0.93
5	North Washington Street Bridge	43300	2	Steel continuous	3	5	5	5	5	0.93
6	Chelsea Street Bridge	13400	1	Steel	8	5	5	5	5	0.87

7	Maurice J. Tobin Memorial Bridge	33500	3	Steel continuous	9	5	5	5	5	0.82
8	Andrew McArdle Bridge	22435	1	Steel	5	5	5	5	5	0.84
9	Granite Avenue Bridge	47600	1	Steel	4	5	5	5	5	0.74
10	Alford Street Bridge	55000	1	Steel	4	5	5	5	5	0.78
11	Summer Street Bridge	27400	6	Steel continuous	5	5	5	5	5	1.00
12	Congress Street Bridge	17100	2	Steel	5	5	5	5	5	0.99
13	Morrissey Boulevard Bridge	77500	2	Steel	9	5	5	5	5	0.91
14	Seaport Boulevard	28448	3	Prestressed concrete	8	3	3	3	3	0.98
15	Nepsonset Bridge	76500	9	Steel	8	3	3	3	3	0.72
16	Parkway Plaza	18200	1	Steel	9	8	5	5	5	0.67
17	Woods Memorial Bridge	10000	1	Cable-stayed	9	5	5	5	5	0.66
18	Fellsway Bridge	31293	3	Steel continuous	9	3	3	3	3	0.66
19	Salem Turnpike	35302	3	Prestressed concrete	7	5	5	5	5	0.42

Note: Data used in this table are from CDC/ATSDR 2022; NBI 2020; Scawthorn et al. (2006)

3 METHODS

3.1. Bridge Condition Modeling

We use scour depth to measure a bridge's structural condition, which further informs its need of maintenance actions. Scour occurs when soil around the bridge pier and the foundation elements is eroded by flowing water. This erosion undermines the bridge foundations and could lead to bridge failure if not properly addressed (Hung and Yau 2017; Pizarro, Manfreda, and Tubaldi 2020). The NBI uses a scour rating system to characterize the susceptibility of a bridge to scour, which is measured on a scale from 0 to 9. A rating of 9 indicates that the bridge is considered safe from scour, while a rating of 0 suggests critical vulnerability. Figure 1 illustrates how the NBI bridge scour ratings are determined based on the relative elevation of the ground surface with respect to the foundation elements of the bridge. The scouring effect causes the lowering of the ground level and consequently a reduction in the bridge scour rating. As prescribed by the NBI, a bridge's scour rating is adjusted to 7 whenever a protective measure, such as the application of riprap, is applied.

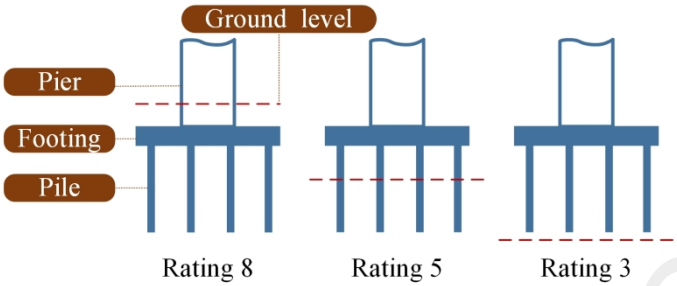

Figure 2- NBI bridge scour rating criterion based on the relative elevation of the ground surface (red dashed lines) with respect to the foundation elements.

Scouring can be either flood-induced or naturally occurring. The flood-induced scour depth, which is defined as the thickness of soil removed by the flood-induced scouring process, can be estimated using the HEC-18 equation developed by the National Cooperative Highway Research Program (NCHRP) (Richardson 1987). The equation has been adopted by AASHTO and has been revised over time to accommodate complex pier shapes and flow conditions (Arneson et al. 2012; Liang, Wang, and Yu 2019). The equation estimates the maximum scour depth h_b for a bridge pier given a flooding event:

$$\frac{h_b}{h_p} = 2.0K_1K_2K_3\left(\frac{D}{h_p}\right)^{0.65}\mathrm{F}^{0.43} \qquad \text{Eq. 1}$$

where h_p = flow depth from water surface to ground surface near the pier, which is obtained from the simulation results of HAZUS; D = width of the pier, K_1, K_2, and K_3 = bridge pier's correction factors regarding the shape of the pier, the pier slenderness and the angle of attack of the flow, and the river bed condition, obtained according to Liang et al. (2017), and F = Froude number that characterizes the type of flow, which is further determined by the Manning Equation (Chanson 2004) using Eq. 2.

$$F = \frac{V}{\sqrt{gh_p}} \qquad \text{Eq. 2}$$

where V = flow velocity at the pier, g = acceleration of gravity, and h_p is the flow depth. The flow velocity can be estimated using the Manning equation (Sturm 2001):

$$V = \frac{1.49}{n} R^{2/3} S^{1/2} \quad \text{Eq. 3}$$

where R is the hydraulic radius (ft), S is the slope of the channel, and n is the Manning roughness coefficient, determined by the channel material. The hydraulic radius R and the slope are calculated based on the Digital Elevation Model (DEM) of the area and flood simulation results obtained from HAZUS. The Manning coefficient n in this study was assumed to be 0.05, which is typical for a riverbed with light brush and weed vegetation (Arcement and Schneider 1989). The scour vulnerability for each bridge corresponding to different flood scenarios is summarized in Table 2.

In addition to flood-induced degradation, bridges naturally deteriorate over time. We used the deterioration function proposed by Agrawal, Kawaguchi, and Chen (2010) to capture the time deterioration of scour ratings r over time (T):

$$r = 9 - 0.0464T - 0.0002T^2 \quad \text{Eq. 4}$$

3.2. Social Vulnerability Impact Analysis

In this study, social vulnerability map developed by Centers for Disease Control and Prevention (CDC/ATSDR 2022) is used for social vulnerability analysis. CDC identifies 16 social vulnerability factors and groups them into four categories including 1) socioeconomic status, 2) household characteristics, 3) racial and ethnic minority status, and 4) housing type and transportation. The granularity of the data corresponds to the 2020 census tract level. To develop a Social Impact Index (SII) for each bridge, the following steps are taken:

1. We calculate the normalized social vulnerability score $SV_{ni,j}$ of social vulnerability factor i for census tract j:

$$SV_{n_{i,j}} = \frac{SV_{i,j}}{\max_j (SV_{i,j})} \quad \text{Eq. 5}$$

where $SV_{i,j}$ is the number of people under the social vulnerability factor i (e.g., people with disabilities) in census tract j, and $\max_j(SV_{i,j})$ is the maximum number of vulnerable people under the same factor i among all census tracts in the testbed. The normalized social vulnerability index $SV_{ni,j}$ ranges from 0 to 1, 0 being the least vulnerable and 1 being the most vulnerable.

2. We then calculate the overall social vulnerability index SV_{nj} for the census tract j. SV_{nj} integrates different categories of social vulnerability and represents the overall social vulnerability for each census tract j:

$$SV_{n_j} = \frac{1}{|\mathtt{I}|} \sum_{i \epsilon_{\mathtt{I}}} SV_{n_{i,j}} \qquad \text{Eq. 6}$$

where $\mathtt{I}$ is the set of social vulnerability factors ($\mathtt{I}$ = {aged 65 or older, below 150% poverty, unemployed, …}), and $|\mathtt{I}|$ is the size of set $\mathtt{I}$.

3. Finally, we calculate the Social Impact Index SII_b for each bridge b. SII_b is defined as the weighted average of the normalized social vulnerabilities SV_{nj} for each census tract j, with the weights being determined by the distance $d_{b,j}$ between bridge b and the centroid of census tract j to account for the vicinity of the bridge to each influenced census tract:

$$SII_b = \frac{1}{|\mathtt{J}|} \sum_{j \epsilon_{\mathtt{J}}} \frac{SV_{n_j}}{e^{(d_{b,j}/r_{eq})}} \qquad \text{Eq. 7}$$

where r_{eq} is the equivalent radius of the study area ($r_{eq} = \sqrt{\frac{Total\ Area}{\pi}}$ = 6.39 km), $\mathtt{J}$ is the set of all census tracts, and $|\mathtt{J}|$ is the size of set $\mathtt{J}$. The Social Impact Index SII_b captures the impact of bridge b on its surrounding socially vulnerable communities. For numerical stability, we normalize SII_b to the 0-1 range:

$$SII_{n_b} = \frac{SII_b}{\max_{b \epsilon \mathsf{B}}(SII_b)} \quad \text{Eq. 8}$$

where SII_{n_b} is the normalized *SII* for bridge b, and B is the set of all bridges. Figure 3 illustrates the spatial locations of each bridge along with their associated SII and cost of failure.

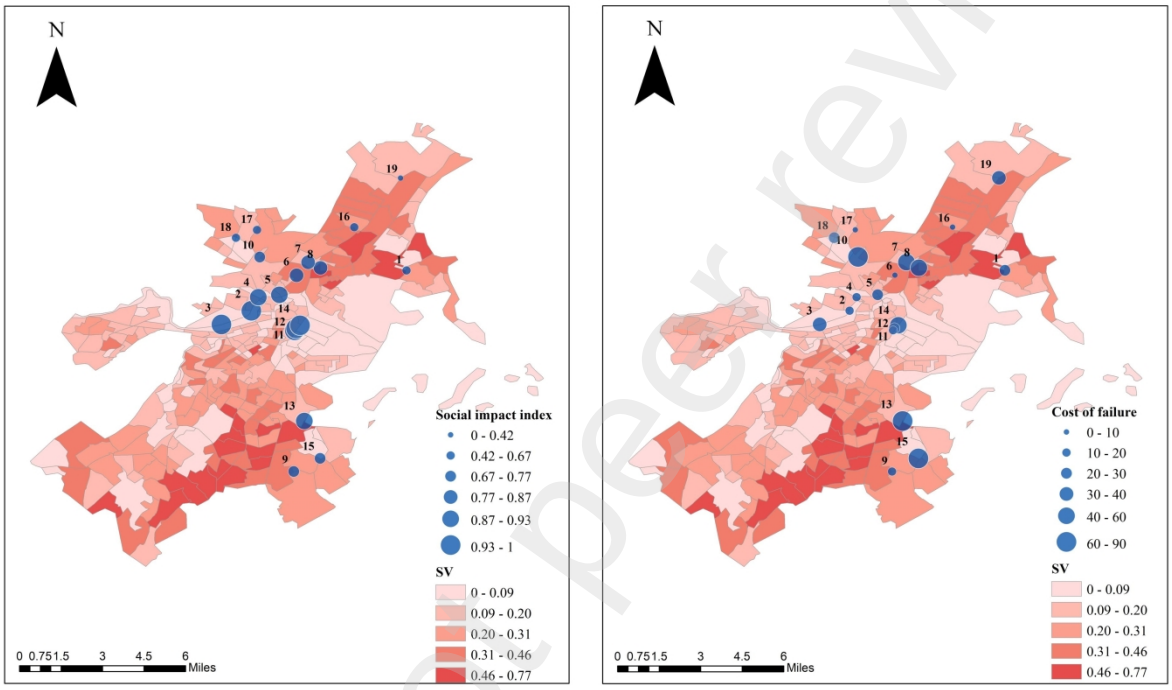

Figure 3- Bridges locations, SII, and cost of failure, social vulnerability of census tracts in Suffolk County, MA; a) bridges' SII; b) physical cost of failure of bridges

3.3. Economic Cost Assessment

In 2010, the Federal Highway Administration and Florida Department of Transportation Mclemore et al. (2010) proposed a framework to assess the cost of failure for bridges. As part of this framework, they developed a formula to estimate the cost of failure for bridges based on the bridge ratings according to the NBI database. The cost of failure consists of three key components: 1) the cost of bridge replacement, 2) the cost of detours, and 3) the cost of fatalities resulting from bridge failure:

$$CF_b = C_1 eWL + \left[C_2\left(1 - \frac{ADTT}{100}\right) + C_3\frac{ADTT}{100}\right]DAd + \left[C_4 O\left(1 - \frac{ADTT}{100}\right) + C_5\frac{ADTT}{100}\right]\frac{DAd}{S} + C_6 X \quad \text{Eq. 9}$$

where CF_b = cost of failure ($) for bridge b; C_1 = unit rebuilding cost per bridge deck area ($/ft^2), obtained from Table 2.2 of Mclemore et al. (2010); e = cost multiplier coefficient, obtained from Table 10 of Stein and Sedmera (2006); W = bridge width (ft), obtained from NBI item 52; L = bridge length (ft), obtained from NBI item 49; C_2 = cost of running cars ($/mile0), obtained from Table 2.3 of Mclemore et al. (2010); C_3 = cost of running truck ($/mile) , obtained from Table 2.6 of (NCHRP report 2006); D = detour length (mi), obtained from NBI item 19; A = average daily traffic, obtained from NBI item 29; d = detour duration (days), obtained from Table 2.8 of Mclemore et al. (2010); C_4 = value of time per adult ($/hr), obtained from Table 2.4 of Mclemore et al. (2010); O = average occupancy rate, obtained from Table 2.5 of Mclemore et al. (2010); $ADTT$= average daily truck traffic (% of average daily traffic (ADT)), obtained from NBI item 109; C_5 = value of time for truck ($/hr), obtained from Table 2.6 of Mclemore et al. (2010); S = average detour speed (40 mph); C_6 = cost of life lost ($); and X = number of deaths resulting from failure, obtained from Table 11 of Stein and Sedmera (2006).

To account for the social impact of each bridge on surrounding socially vulnerable communities, we modify the economic cost of failure CF_b for bridge b using the normalized Social Impact Index SII_{nb} for bridge b obtained in Section 3.2:

$$CF_{b,mod} = {SII_{n_b}}^{\alpha} {CF_b}^{1-\alpha} \quad \text{Eq. 10}$$

where $CF_{b,mod}$ is the modified cost of failure for bridge b, and α is the vulnerability influence factor that varies between 0 and 1. When α is set to 0, $CF_{b,mod}$ reverts to the original cost of failure CF_b, disregarding the social impact of potential bridge failure. On the other hand, setting

α = 1 represents full consideration of bridge's social impact. We will consider five values of α (α = 0, 0.25, 0.50, 0.75, and 1) in this study.

In any given year, different flood scenarios occur with corresponding probabilities (e.g., 50-year return period flood has a probability of occurrence = 1/50 = 2%), leading to potential bridge failure. The expected cost of failure C_f for the entire system of bridges ($) for that year is:

$$C_f\,(\$) = \sum_{b \epsilon \mathsf{B}} \left(CF_{b,mod} \sum_{fld \epsilon \mathsf{F}} P(failure \,|\, fld, r_b) \times P(fld) \right) \qquad \text{Eq. 11}$$

where B is the set of all bridges being considered; $CF_{b,mod}$ is the modified cost of failure for bridge b, obtained from Eq. 10; $P(failure|\, fld, r_b)$ is the conditional probability of failure for bridge b given a specific flood scenario *fld* and the scour rating r_b of bridge b, obtained from Table 3 (adapted from Pearson, Stein, and Jones (2002)); F is the set of flood scenarios, and $P(fld)$ is the probability of occurrence of flood with a specific return period. For example, $P(fld)$=0.02 for flood with return period of 50 years. It is important to note that as the bridges' ratings are affected by both time deterioration and flood-induced degradation, the expected cost of failure C_f changes over time.

Table 3- Annual probability of failure for bridges under different flood scenarios depending on the NBI scour rating (Mclemore et al. 2010).

Scour Rating	Flood with 500 years return period	Flood with 100 years return period	Flood with 50 years return period	Flood with 10 years return period
0	1	1	1	1
1	0.01	0.01	0.01	0.01
2	0.005	0.006	0.008	0.009
3	0.0011	0.0013	0.0016	0.002
4	0.0004	0.0005	0.0006	0.0007
5	0.000007	0.000008	0.00004	0.00007
6	0.00018	0.00025	0.0004	0.0005
7	0.00018	0.00025	0.0004	0.0005
8	0.000004	0.000005	0.00002	0.00004
9	0.0000025	0.000003	0.000004	0.000007

In addition to the expected cost of failure, we also considered the maintenance cost associated with the bridge repair. Precise data regarding the scour repair costs for bridges are notably absent in existing literature. To estimate the repair costs in developing the optimization framework, we assume the repair cost C_{rb} for bridge b to be a fraction of the cost of failure CF_b (Eq. 9), where the fraction η depends on the current bridge scour ratings, as summarized in Table 4.

$$Cr_b = \eta \times CF_b \quad \text{Eq. 12}$$

Table 4- Fraction of cost of failure (η) for repairing actions bridges depending on the bridge scour ratings

Rating 0	0.000444
Rating 1	0.000250
Rating 2	0.000160
Rating 3	0.000111
Rating 4	0.000082
Rating 5	0.000062
Rating 6	0.000049
Rating 7	0
Rating 8	0
Rating 9	0

3.4. Social Vulnerability Evaluation

In this section, we introduce a metric to evaluate how maintenance decisions improve social equity. This metric, the SII-weighted average probability of failure, measures a bridge portfolio's social impact on local vulnerable communities, and it is calculated in three steps:

1. Determine the probability of failure $P_{b,i}$ for bridge b in year i:

$$P_{b,i} = \sum_{fld \epsilon \mathbb{F}} P(failure \,|\, fld, r_{b,i}) \times P(fld) \quad \text{Eq. 13}$$

2. Compute the average probability of failure $\overline{P_b}$ for bridge b over 20 years:

$$\overline{P}_b = \frac{1}{20} \sum_{i=1}^{20} P_{b,i} \quad \text{Eq. 14}$$

3. Calculate weighted-average probability of failure $P_{SII\text{-}weighted}$ for the entire bridge portfolio, using the social impact index (SII) as the weight:

$$P_{SII-weighted} = \frac{\sum_{b \epsilon \mathbf{B}} \overline{P}_b \times SII_b}{\sum_{b \epsilon \mathbf{B}} SII_b} \quad \text{Eq. 15}$$

A larger value for $P_{SII\text{-}weighted}$ suggests a management outcome in which bridges with greater *SII* (with greater social impact) end up with greater probability of failure. On the other hand, lower $P_{SII\text{-}weighted}$ suggest that priority has been given to bridges with larger social impact (i.e., larger SII). By associating probability of failure with the social impact index (SII), this metric provides insights into the potential societal impacts of bridge portfolio management.

3.5. Deep Reinforcement Learning (DRL) Modeling

Framework: We use Deep Reinforcement Learning (DRL) to develop a maintenance strategy aimed at minimizing the economic costs associated with potential bridge failure and bridge maintenance while promoting social equity within neighboring communities over a 20-year span. Figure 4 shows the structure and workflow of the DRL model framework, in which the artificial intelligent (AI) agent interacts with its Environment, which is modeled a Markov Decision Process (MDP), as illustrated in Figure 5. At each time step (we consider one-year time increment), the artificial intelligent (AI) agent can choose to repair one bridge, two bridges, or Do-Nothing (i.e., not repairing any bridge). Based on the maintenance decision, the state of the Environment (i.e., bridge ratings) will be updated (e.g., improvement by maintenance, flood-induced degradation, time deterioration), and the AI agent will receive a feedback reward from the Environment according to the objective function defined by Eq. 16. Based on the updated Environment state and the reward received from the Environment, the AI makes a new maintenance decision for the next time step. This feedback loop iterates until the end of the 20-year time horizon, concluding one learning cycle. This learning cycle will be repeated as the

AI agent continuously refines its decision strategy with the objective of maximizing the cumulative rewards over 20 years, until the model converges. Specifically, the decision strategy learning is done using the Proximal Policy Optimization (PPO) algorithm, a type of policy gradient based DRL, which is thoroughly explained in Appendix 2.

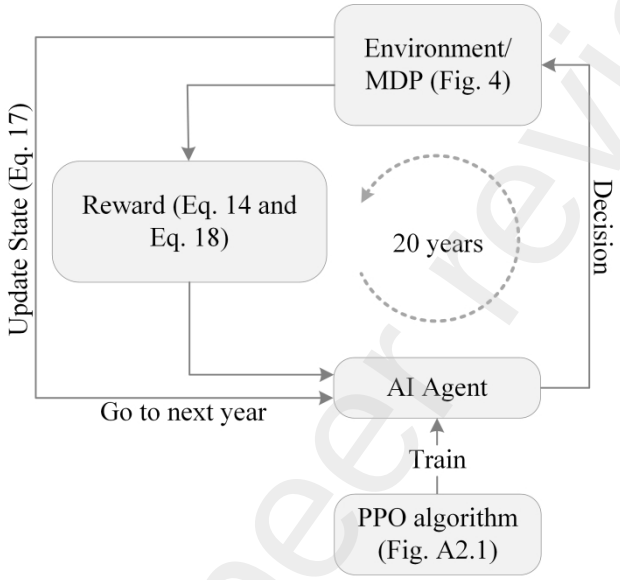

Figure 4- DRL framework to optimize the bridge portfolio maintenance.

Objective function: The objective function of this optimization problem can be written as:

$$\boldsymbol{Min.} \sum_{b\in \mathsf{B}} \sum_{t\in \mathsf{T}} \gamma^t \left[C_{r_b}(r_{b,t}, CF_{b,\,mod}) x_{b,t} + \sum_{f\in \mathsf{F}} [C_f(fld, r_{b,t}) x_{b,t}] \right] \qquad \text{Eq. 16}$$

Subject to:

$$r_{(b,t)} = \begin{cases} r_{(b,t-1)} - det(r_{(b,t-1)}) - \frac{deg(r_{(b,t-1)}, fld)}{7} & if\ \ x_{b,t-1} = 0 \\ & if\ \ x_{b,t-1} = 1 \end{cases} \qquad \text{Eq. 17}$$

$$\sum_{b\in \mathsf{B}} x_{b,t} \leq 2 \qquad \text{Eq. 18}$$

$$\forall b \in \mathsf{B} = \{1,\ldots,19\},\ t \in \mathsf{T} = \{1,\ldots,20\},\ \forall f \in \mathsf{F} = \{10, 50, 100, 500\}, x_{b,t} \in \{0,1\},\ r_{b,t} \in \{0,1,\ldots,9\}$$

Where γ is the discount factor, 0.99; $CF_{b,mod}$ is the modified cost of failure for bridge b (Eq. 10); C_{r_b} is the cost associated with repairing bridge b with condition rating r and $CF_{b,mod}$; $r_{b,t}$ is an integer variable representing the scour rating of bridge b in year t; $x_{b,t}$ is a binary decision variable that takes value 1 if bridge b is repaired in year t, and 0 otherwise; $C_f(fld, r_{b,t})$ is the expected cost of flooding for bridge b with the scour rating r in flood scenario *fld* at time t; $det(r_{(b,t)})$ is the deterioration rate of bridge b with scour rating r at time t; and $deg(r_{(b,t)}, fld)$ is the expected flood-related degradation for bridge b at time t in flood scenario *fld.*

In a reinforcement learning (RL) problem, the environment is the system in which an agent operates and learns to make optimized decisions. The emulator of the environment is typically modeled as a Markov Decision Process (MDP). MDP is a framework used to model decision-making problems in situations where outcomes depend on both the actions taken by a decision-maker and the state of the agent in its environment. The outcome of a decision at each time step is determined by the current state and the chosen action. MDP is numerically defined by a tuple M = ($\mathbf{S}$, $\mathbf{A}$, T, R), consisting of four main components:

States: $\mathbf{S} = \{\boldsymbol{s_1}, \boldsymbol{s_2}, \ldots, \boldsymbol{s_t}\}$, where $\boldsymbol{s_t}$ denotes the state of the system at time t. In this study, $\boldsymbol{s_t}$ is represented by a vector of 19 elements, each corresponding to the rating of one of the 19 bridges.

Actions: $\mathbf{A} = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_t}\}$, where $\boldsymbol{a_t}$ denotes the action taken at time t, which is represented as a vector of 2 elements, each corresponding to one action. Under each state $\boldsymbol{s}$, there are 20 options for each action: pick one of the 19 bridges to maintain or Do-Nothing .

Transition function: Transition function T defines the dynamics of the system. When the agent takes action $\boldsymbol{a_t}$ under state $\boldsymbol{s_t}$ at time t, the new state $\boldsymbol{s_{t+1}}$ of the system at time t+1 can be obtained using the transition function T:

$$\boldsymbol{s}_{t+1} = T(\boldsymbol{a_t}, \boldsymbol{s}_t) \quad \text{Eq. 19}$$

In this study, the transition function is obtained from the flood-induced degradation, time deterioration, and the maintenance decision.

Reward function: After the agent takes an action at time t, it will receive a reward R_{t+1} from the environment as feedback depending on the current system state $\boldsymbol{s_t}$ and the action $\boldsymbol{a_t}$. The reward in an RL framework must be designed to reflect the objective function for the underlying optimization problem. The reward mechanism guides the agent towards learning an optimal policy by incentivizing actions that lead to desirable outcomes. In this study, R_{t+1} is defined as the negative sum of the expected cost of failure $C_f(\boldsymbol{s_t})$ evaluated under state $\boldsymbol{s_t}$ (Eq. 11) and the repairing cost $C_r(\boldsymbol{a_t}, \boldsymbol{s_t})$ (Eq. 12):

$$R_{t+1} = \begin{cases} -C_f(s_t) - C_r(a_t, s_t) & if\ choose\ bridge\ with\ rating < 7 \\ -10^6 & Otherwise \end{cases} \quad \text{Eq. 20}$$

The second line of the equation is introduced to discourage the agent from repairing a bridge with a rating ≥ 7 by penalizing the agent with a large negative reward (-10^6). This is done because the bridge scour rating will be reset to 7 after maintenance according NBI standard, and thus there is no benefit of maintaining a bridge with ratings ≥ 7. Figure 5 shows the MDP developed for this study.

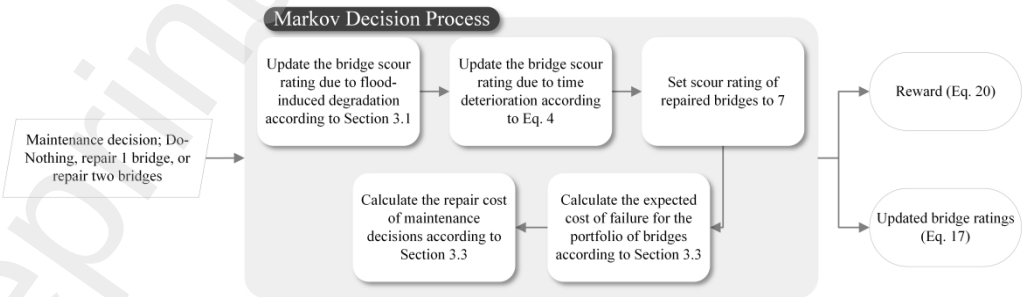

Figure 5- Markov Decision Process (MDP) for sequential decision-making problem of bridge system maintenance

4 RESULTS AND DISCUSSION

4.1. Optimization considering only economic cost

In RL, the learning curve graphically illustrates the evolution of the RL agent's performance, showcasing how its decision-making strategy progressively improves with respect to the cumulative rewards (or costs) through iterations. Figure 6 shows the learning curve obtained from the Proximal Policy Optimization (PPO). We only consider the economic cost for the objective function in this section ($\alpha = 0$ in Eq. 10), and the results with incorporation of social impact will be presented in Section 0 and Section 4.4. Initially, there is a fast decrease in cost, suggesting that the agent has swiftly learned to avoid repairing bridges with a rating of 7 or higher to avoid the substantial penalties (Eq. 20). After that, the cost continues decreasing at a slower rate as the agent continues to explore and refine more effective decision-making strategies. We selected the strategy learned by the RL agent at the end of 80,000 episodes.

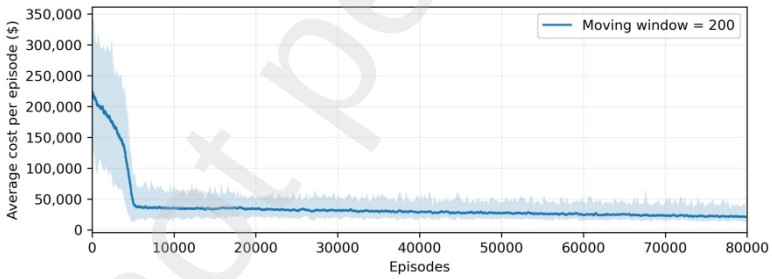

Figure 6- PPO learning curve for learning with no consideration of social equity (α=0 in Eq. 10). The dark blue curve shows a moving average with a window size of 200.

In the optimized scenario, thirteen bridges are never chosen to be repaired during the entire period potentially due to: 1) their initial ratings are high (e.g., above 7), 2) their cost of failure is relatively low, 3) the flood-induced degradation and time deterioration is slow. For example, bridge 4 has an initial scour rating of 9, (thus the time-related deterioration rate is low according to Eq. 4), and its flood-related degradation is not significant (Table 2). Thus, it is never selected by the RL agent over the 20-year period. On the other hand, bridge 15 is chosen twice for maintenance due to the large rate of degradation (Figure 7) and the great potential

economic cost if failure ever occurs, thus receiving greater priority from the RL agent in the optimization.

Figure 7 illustrates the history of the scour rating for four representative bridges as the output of the RL optimization considering only economic cost in the objective function. The RL agent's attitude towards bridges with lower rate of degradation (i.e., bridge 2 and bridge 4) is significantly different than the more vulnerable bridges that are located in highly flood-prone areas (i.e., bridge 15). The high flood degradation rate necessitates repairs to avert potential failures. This observation suggests that the agent's primary objective is to sustain acceptable scour ratings for all bridges throughout the entire horizon, motivated by the understanding that even a single severely deteriorated bridge can significantly elevate the risk profile of the entire portfolio. Additionally, it is important to acknowledge that the agent's decision-making process is influenced not only by the flood degradation rate but also by other critical factors, including initial scour ratings (Table 2). These factors collectively shape the agent's behavior in prioritizing and executing repair actions across the bridge portfolio.

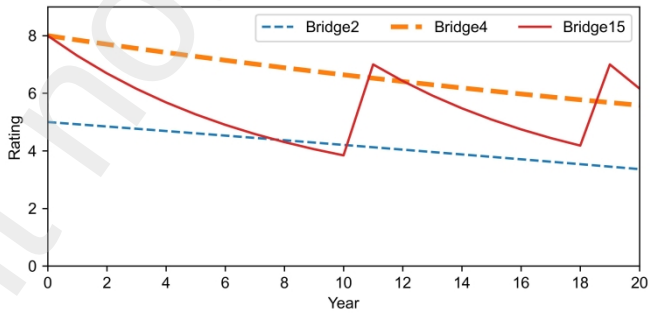

Figure 7- Scour rating for bridges in a horizon of 20 years as a result of the optimized maintenance strategy considering only economic cost in the objective function with no consideration of social equity (α=0 in Eq. 10).

4.2. Model validation with baselines

To assess the performance of reinforcement learning algorithms, it is essential to establish baseline benchmarks. For this purpose, we compare the optimized sequence of

maintenance actions generated by PPO with four baseline scenarios: "Do-Nothing," random policy, expert policy, and policy without consideration of social vulnerability analysis (i.e., no cost adjustment applied).

Baseline 1- Do-Nothing policy: The Do-Nothing baseline is often used as a fundamental reference for RL model validation. This baseline corresponds to a policy where the agent takes no action regardless of its current state. If the learned policy fails to outperform the Do-Nothing policy, it indicates that the RL algorithm is not effectively learning from the environment. Although in certain real-world scenarios, choosing the Do-Nothing strategy may not necessarily be a poor strategy, comparing the RL algorithm against this baseline helps the modeler understand how much value is being gained through active decision-making.

Baseline 2- random policy: Another commonly used baseline is the random action policy, in which actions are chosen randomly from the available action space. It represents a simple and naïve strategy that lacks any deliberate decision-making or learning mechanism. Despite its simplicity, the random action policy serves as a valuable baseline for evaluating the performance of RL algorithms.

Baseline 3- expert policy: The expert policy, also known as the status quo, represents the existing policy traditionally employed by the bridge decision-makers. Under this policy, two bridges with the minimum ratings at each timestep are selected for maintenance. This approach tends to minimize the number of non-functional bridges. The ultimate objective of the expert policy is to maintain the system's overall functionality rather than minimizing the total cost over the planning horizon.

Table 5 shows costs associated with the proposed PPO and three baselines described above. In baseline 1, 197% improvement in the proposed PPO policy over Do-Nothing Policy shows a significant effectiveness of the proposed algorithm. Furthermore, a difference of 386%

for the proposed PPO over random policy shows that the algorithm has successfully learned from the rewards and experiences, whose outcomes is not random. In baseline 3, a 52% improvement highlights that the proposed PPO has discovered more efficient actions than the status quo, due to achieving an optimal equilibrium between action cost and the expected cost of failure. Given these findings, it is clear that the proposed algorithm's performance is significant, as it outperforms the established baselines.

Table 5- Model evaluation results on cumulative costs

Policy	Action cost ($)	Expected cost of failure ($)	Total cost ($)	Cost differences (%)
PPO (without social vulnerability, i.e., α = 0 in Eq. 10)	3,009	3,594	6,603	-
Baseline 1: Do-Nothing	0	1,9631	19,631	197%
Baseline 2: random policy	20,919	4,568	25,487	386%
Baseline 3: expert policy	9,397	645	10,042	52%

Our analysis further demonstrates the superior performance of the proposed algorithm, when it incorporates SV considerations (α = 1 in Eq. 10), in comparison to the established baselines. As illustrated in Table 6, the SII-weighted average probability of failure (Eq. 15) reveals a significant improvement with the application of our proposed approach over do-nothing policy, random policy, and expert policy. According to Table 5 and Table 6, the model's practicality is not merely in terms of portfolio cost optimization but extends crucially to the enhancement of social vulnerability outcomes. Such results underscore the significance of integrating SV considerations into the algorithmic framework, thereby yielding not only economic but also societal benefits.

Table 6- Model evaluation results on social vulnerability

Policy	$P_{SII\text{-}weighted}$ (Eq. 15)	$P_{SII\text{-}weighted}$ differences (%)
PPO (with social vulnerability, i.e., α = 1 in Eq. 10)	0.0054%	-
Baseline 1: Do-Nothing	0.0800%	1391.72%
Baseline 2: random policy	0.0505%	841.01%
Baseline 3: expert policy	0.0167%	211.42%

4.3. The effect of SV on the algorithm's behavior

We utilize two factors of the normalized modified cost of failure and the frequency of repairs for each bridge to show the significance of SV on the algorithm's behavior. For normalization, we divide the modified cost of failure for each bridge by the average modified cost of failure across all bridges, yielding a mean value of 1. Thus, a bridge with a normalized modified cost of failure above 1 is deemed more critical in terms of failure costs than the average bridge. In Figure 8, we explore the variation in normalized modified costs of failure for Bridges 5, 12, and 15 across different α values (Figure 8a), alongside the corresponding number of repairs dictated by the algorithm for each bridge concerning different α value (Figure 8b). The first column of the heatmaps represents scenarios devoid of SV consideration, highlighting Bridge 15 with a normalized modified cost of failure of 2.23—a notably high value within the portfolio. Bridges 5 and 12 exhibit lower normalized modified cost of failure, leading to no repair actions for these bridges, while Bridge 15 is repaired twice (as shown in Figure 8b). Conversely, when examining the scenario with α equals 1, Bridge 15, despite its higher cost of failure, registers a lower SII, diminishing its priority when solely considering social vulnerability. Under this scenario, Bridges 5 and 12, characterized by higher SII values, display increased modified costs of failure, prompting the algorithm to schedule Bridge 5 for two repairs and Bridge 12 for one repair, leaving Bridge 15 without any repair due to its reduced SII. The stark contrast between the extreme scenarios of $\alpha = 0$ and $\alpha = 1$, as depicted in Eq. 10, underscores the algorithm's sensitivity to the incorporation of SV in the analysis. The variations in repair frequency and prioritization between these scenarios conclusively demonstrate that SV plays a crucial role in shaping maintenance strategies for bridges. This analysis also highlights the algorithm's adaptability to different prioritization metrics, concerning the contribution of SV to bridge maintenance planning problems.

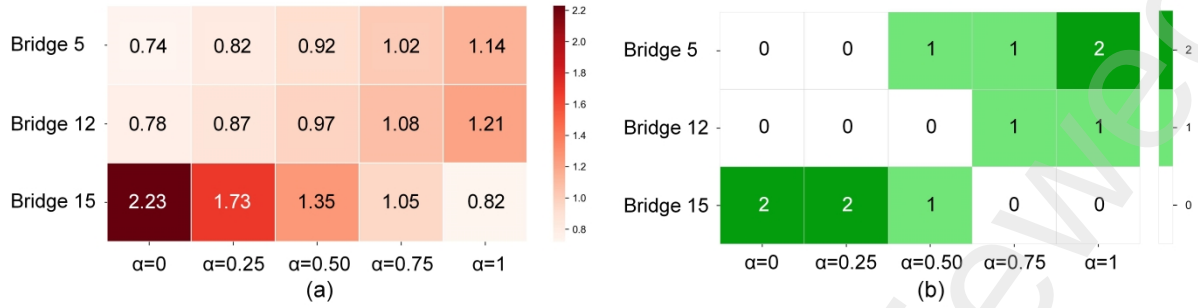

Figure 8- (a) Normalized modified cost of failure for bridge 5, 12, and 15 across various α values in Eq. 10; (b) Number of repairs for bridge 5, 12, and 15 in the optimized episode concerning various α values in Eq. 10.

4.4. Understanding the tradeoff between the economic cost and social equity

Figure 9 illustrates the tradeoff between the total physical cost (including the repair costs C_r and the expected cost of failure C_F) and $P_{SSI\text{-}weighted}$ (Eq. 15) associated with different α values. As expected, $P_{SSI\text{-}weighted}$ decreases as the contribution of SV increases in the calculation of the modified cost of failure (Eq. 10). This suggests that by factoring in SV, bridges that are more critical to society are identified and prioritized for maintenance. Accordingly, the incorporation of SV in maintenance planning results in an increased total cost for the bridge portfolio because accounting for SV leads to the selection of more expensive maintenance strategies, which are necessary to mitigate risks on socially critical bridges. Thus, the inclusion of social vulnerability in bridge maintenance planning does not come without added costs, but it can lead to a socially optimized portfolio of bridges with a lower risk of failure. This cost is justifiable from a societal welfare perspective to balance the financial investment and societal equity. Figure 9 provides bridge owners with valuable managerial insights regarding the implications of SV consideration, by showing the financial impact of integrating SV into maintenance planning.

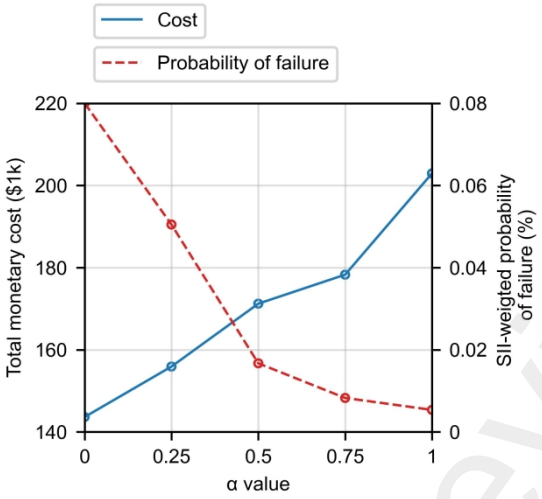

Figure 9- Physical cost and SII-weighted average probability of failure across different values for α

5 CONCLUSIONS

This study has introduced a sequential decision-making framework designed to reduce the overall costs associated with bridge maintenance and failure in flood-prone areas, while simultaneously improving social equity among affected communities. Through a detailed case study in Suffolk County, MA, involving various flooding scenarios, we have demonstrated the framework's capacity to identify bridges at risk and evaluate their repair and failure costs comprehensively. Key innovations include the use of the Social Impact Index (SII) to integrate social equity considerations into bridge management decisions and the application of the Proximal Policy Optimization (PPO) algorithm for strategic planning over a 20-year horizon.

Our findings reveal that the proposed framework not only outperforms traditional management strategies but also highlights the significance of including social vulnerability (SV) in infrastructure decision-making processes. The use of deep reinforcement learning has proven effective in navigating the complexities of bridge maintenance under flood risk, offering a promising tool for policymakers and infrastructure managers. Furthermore, we have quantified the financial implications of integrating SV into portfolio management, revealing a 41% increase in portfolio costs over a 20-year horizon when SV consideration is the only

determining factor, as opposed to physical cost factors. One the other hand, neglecting SV in the decision-making processes leads to a significant 91% exacerbation of SV. To facilitate a balanced consideration of SV and physical costs in decision-making, we have also provided an approach that assists decision-makers in determining the relative weight of each factor within their decision-making processes

Looking ahead, the integration of budget constraints into the decision-making model represents a logical next step, ensuring that the framework reflects the financial realities of bridge maintenance projects. Additionally, a broader examination of the bridges within the context of urban transportation networks could yield further insights into optimizing infrastructure resilience in the face of climate change and urbanization pressures. This study paves the way for more sophisticated, equitable, and effective approaches to managing critical infrastructure in vulnerable regions.

APPENDIX 1. Corrections factor for Eq. 1

Table A1.7, Eq. A1.21, and Where θ is the angle of attack of water, L is the length of the pier, and D is the diameter of the pier.

Table A1.8 determine the correction factors for K_1, K_2, and K_3 in Eq. 1, respectively. These values are obtained from Liang et al. (2017).

Table A1.7- Correction factor K_1 in Eq. 1

Shape of pier nose	$\boldsymbol{K_1}$
Square	1.1
Round	1
Cylinder	1
Group of cylinders	1
Sharp	0.9

$$\boldsymbol{K_2 = (\cos\theta + \sin\theta \times \frac{L}{D})^{0.65}} \qquad \text{Eq. A1.21}$$

Where θ is the angle of attack of water, L is the length of the pier, and D is the diameter of the pier.

Table A1.8- Correction factor K_3 in Eq. 1

Bed condition	K_3
Clear-water scour	1.1
Plane bed and antidune flow	1.1
Small dunes	1.1
Medium dunes	1.1 to 1.2
Large dunes	1.3

APPENDIX 2. Proximal Policy Algorithm

RL encompasses two primary types of approaches: model-based and model-free methods. In model-based RL, the agent learns a model of the environment and utilizes it to predict future states and rewards. Model-free RL directly learns optimal policies from the interaction with the environment without explicitly learning a model of the environment. Within model-free RL, there are two main subcategories: value-based and policy-based methods. Value-based methods, such as Deep Q-Network (Mnih et al. 2015), aim to estimate the value of each state or state-action pair and select actions that maximize these value estimates. They often employ iterative algorithms to update value estimates based on observed rewards. In contrast, policy-based methods, like Proximal Policy Optimization (PPO) (Schulman et al. 2017), optimize the policy (in PPO, the deep neural network) parameters directly to maximize expected rewards. These methods typically rely on gradient-based optimization (e.g., gradient ascent) techniques and learn from the agent's interactions with the environment.

In policy-based RL, policy refers to the decision-making strategy that an agent employs to maximize the cumulative reward. A policy function is often parameterized using an artificial neural network. It determines the actions to be taken by the agent based on the state of the

agent-environment system. For discrete action spaces, the policy is represented using a softmax function over the action probabilities:

$$\pi(\boldsymbol{a} \mid \boldsymbol{s}; \boldsymbol{\theta}) = softmax(f(\boldsymbol{s}; \boldsymbol{\theta})) \quad \text{Eq. A2.22}$$

where $\boldsymbol{a}$ represents action, $\boldsymbol{s}$ represents the state, $\boldsymbol{\theta}$ represents the policy parameters (in this case, the weights and biases of the deep neural network), and $\boldsymbol{f}(\boldsymbol{s}; \boldsymbol{\theta})$ is the output of the neural network.

PPO is a popular and effective policy-based algorithm in RL. PPO addresses the challenge of finding a balance between exploration and exploitation by iteratively updating the policy in a way that ensures stable and efficient learning (Schulman et al. 2017). Unlike traditional policy-based methods, PPO employs a proximal (surrogate) objective function (

Eq. A2.23) that incorporates a constraint on the size of the policy update (gradient clipping).

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t;\theta)}{\pi_{\theta_{\text{old}}}(a_t|s_t;\theta_{old})}$$

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t\left[min\left[r_t(\theta)\hat{A}_t, clip\left(r_t(\theta), 1-\varepsilon, 1+\varepsilon\right)\hat{A}_t\right]\right] \quad \text{Eq. A2.23}$$

$$\hat{A}_t(s,a) = Q(s,a) - V(s) \quad \text{Eq. A2.24}$$

$$Q(s,a) = \mathbb{E}[G_t|(s_t = s, a_t = a)] \quad \text{Eq. A2.25}$$

$$V(s) = \mathbb{E}[G_t|(s_t = s)] \quad \text{Eq. A2.26}$$

where $L^{CLIP}(\theta)$ is the objective function of the PPO algorithm. $\boldsymbol{\pi_\theta(a_t|s_t; \theta)}$ represents the current policy for taking action $\boldsymbol{a}$ given state $\boldsymbol{s}$ and parameter $\boldsymbol{\theta}$, $\boldsymbol{\pi_{\theta\_old}(a_t|s_t; \theta_{\text{old}})}$ represents the old policy for taking action $\boldsymbol{a}$ given state $\boldsymbol{s}$ and old parameter $\boldsymbol{\theta_{old}}$, ε is the hyperparameter that controls the clipping extent. The state-action value function $Q(\boldsymbol{s}, \boldsymbol{a})$ estimates the expected cumulative reward when taking action $\boldsymbol{a}$ in state $\boldsymbol{s}$ and following a specific policy thereafter (Eq. A2.25). It represents the quality (value) of taking the specific action $\boldsymbol{a}$ in the state $\boldsymbol{s}$. The state value function $V(\boldsymbol{s})$ estimates the expected cumulative reward when starting in state s and following a specific policy thereafter, representing the overall quality (value) of being in that state (Eq. A2.26). $\hat{A}(\boldsymbol{s},\boldsymbol{a})$ is the advantage function (Eq. A2.24) for state $\boldsymbol{s}$ and action $\boldsymbol{a}$. $\boldsymbol{\hat{A}(s,a)}$

is called "advantage function" because by subtracting the $V(\boldsymbol{s})$ from the $Q(\boldsymbol{s}, \boldsymbol{a})$, the difference represents the advantage or disadvantage of taking the specific action $\boldsymbol{a}$ in state $\boldsymbol{s}$ compared to the average performance of all possible actions in state $\boldsymbol{s}$.

$\boldsymbol{L^{CLIP}(\theta)}$ is the empirical average of the minimum of two terms. These terms are two competing functions. The first term aims at encouraging the policy update, which increases the selection probability of actions with higher advantage. On the other hand, the second term introduces a constraint on the policy update by clipping the ratio of the new and old policy to prevent significant deviations from the old policy, ensuring more stable training. The feature of clipping contributes to the success of PPO in achieving robust, stable, and reliable policy optimization, making it a popular choice in various reinforcement learning applications. Figure A2.10 illustrates the algorithm of PPO.

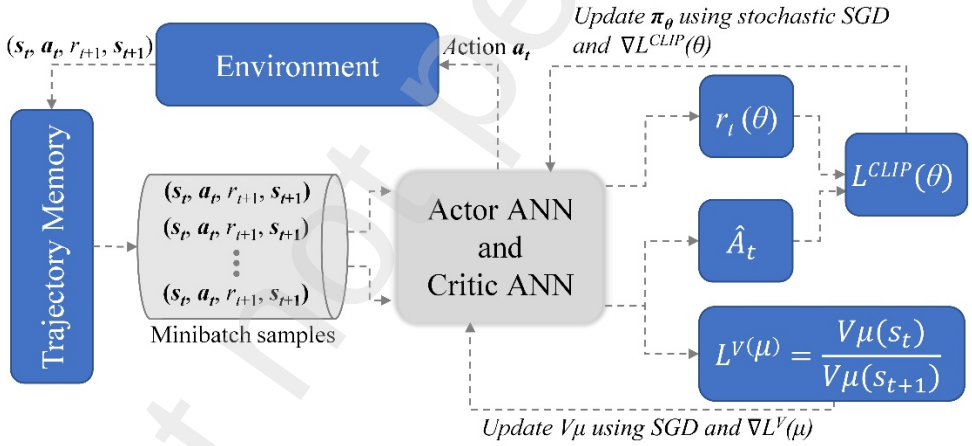

Figure A2.10- Overall scheme of PPO

Entropy Coefficient is a hyperparameter of PPO algorithm that controls the contribution of the actor's entropy (randomness) to the overall objective function during training. In the context of PPO, entropy coefficient encourages exploration by discouraging the policy from becoming overly deterministic and favoring actions with high probabilities. The entropy of the actor can be defined in terms of the PPO objective function (Eq. A2.27) and (Eq. A2.28).

$$L^{PPO}(\theta) = L^{CLIP}(\theta) - cL^{S}(\pi_\theta) \quad \text{Eq. A2.27}$$

$$L^{S}(\pi_\theta) = -\sum_{a} \pi_\theta(a \mid s) \log(\pi_\theta(a \mid s)) \quad \text{Eq. A2.28}$$

where $L^{PPO}(\boldsymbol{\theta})$ is PPO objective function to be maximized, $L^{CLIP}(\boldsymbol{\theta})$ is the clipped surrogate objective (Eq. A2.23), $\boldsymbol{c}$ is the entropy coefficient, $L^{S}(\boldsymbol{\pi_\theta})$ is the actor's entropy term, and $\boldsymbol{\pi_\theta(a|s)}$ is the probability of selecting action $\boldsymbol{a}$ given state $\boldsymbol{s}$ according to the policy network $\boldsymbol{\pi}$ with parameters $\boldsymbol{\theta}$. By including the entropy term in the objective function and adjusting the entropy coefficient $\boldsymbol{c}$, better control of balance between exploration and exploitation during training is possible. A higher entropy coefficient encourages more exploration, while a lower coefficient promotes exploitation and a more deterministic policy.

APPENDIX 3. Model development and training

The model is developed using Python 3.8 and PyTorch 2, and it runs on a machine equipped with an Intel Core-i9-10920X CPU, 128 GB of memory, and an NVidia RTX A5000 GPU. The PPO algorithm utilizes the Adam optimizer with a step size of 0.0005 for gradient ascent. The neural network architecture comprises two hidden layers with 64 nodes each, same for both the actor and critic network. The training process requires approximately 36 hours to complete. After conducting a grid search to identify the optimal hyperparameters, a minibatch size of 64 and a discount factor ($\boldsymbol{\gamma}$) of 0.99 are determined. To encourage more exploration in the initial episodes and more exploitative behavior as learning progresses, the entropy coefficient is designed to gradually decay from 0.05 to 0.0005.

6 REFERENCES

Abdelkader, Eslam Mohammed, Osama Moselhi, Mohamed Marzouk, and Tarek Zayed. 2022. "An Exponential Chaotic Differential Evolution Algorithm for Optimizing Bridge Maintenance Plans." *Automation in Construction* 134:104107.

Agrawal, Anil K., Akira Kawaguchi, and Zheng Chen. 2010. "Deterioration Rates of Typical Bridge Elements in New York." *Journal of Bridge Engineering* 15(4):419–29.

Anderson, MJ, DAF Kiddle, and TM Logan. 2022. "The Underestimated Role of the Transportation Network: Improving Disaster & Community Resilience." *Transportation Research Part D: Transport and Environment* 106:103218.

Arcement, George J., and Verne R. Schneider. 1989. *Guide for Selecting Manning's Roughness Coefficients for Natural Channels and Flood Plains*. Vol. 2339. US Government Printing Office Washington, DC.

Arneson, L. A., L. W. Zevenbergen, P. F. Lagasse, and P. E. Clopper. 2012. *Evaluating Scour at Bridges*. National Highway Institute (US).

Asghari, Vahid, Ava Jahan Biglari, and Shu-Chien Hsu. 2023. "Multiagent Reinforcement Learning for Project-Level Intervention Planning under Multiple Uncertainties." *Journal of Management in Engineering* 39(2):04022075.

Bibri, Simon Elias, and John Krogstie. 2017. "On the Social Shaping Dimensions of Smart Sustainable Cities: A Study in Science, Technology, and Society." *Sustainable Cities and Society* 29:219–46.

CDC/ATSDR. 2022. "Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research, Analysis, and Services Program. CDC/ATSDR Social Vulnerability Index 2020 Database MA."

Chanson, Hubert. 2004. *Hydraulics of Open Channel Flow*. Elsevier.

Cheng, Minghui, and Dan M. Frangopol. 2021. "A Decision-Making Framework for Load Rating Planning of Aging Bridges Using Deep Reinforcement Learning." *Journal of Computing in Civil Engineering* 35(6):04021024.

Dahl, Kristina A., Melanie F. Fitzpatrick, and Erika Spanger-Siegfried. 2017. "Sea Level Rise Drives Increased Tidal Flooding Frequency at Tide Gauges along the US East and Gulf Coasts: Projections for 2030 and 2045." *PloS One* 12(2):e0170949.

Du, Ao, and Alireza Ghavidel. 2022. "Parameterized Deep Reinforcement Learning-Enabled Maintenance Decision-Support and Life-Cycle Risk Assessment for Highway Bridge Portfolios." *Structural Safety* 97(March). doi: 10.1016/j.strusafe.2022.102221.

Federal Highway Administration. n.d. "Bridges & Structures. Bridge Condition by Highway System 2020." Retrieved June 2, 2023 (https://www.fhwa.dot.gov/bridge/nbi/no10/condition19.cfm).

Hamida, Zachary, and James-A. Goulet. 2023. "Hierarchical Reinforcement Learning for Transportation Infrastructure Maintenance Planning." *Reliability Engineering & System Safety* 235:109214.

Hung, Chung-Chan, and Wen-Gi Yau. 2017. "Vulnerability Evaluation of Scoured Bridges under Floods." *Engineering Structures* 132:288–99.

Ilbeigi, M., and M. Ebrahimi Meimand. 2020. "Statistical Forecasting of Bridge Deterioration Conditions." *Journal of Performance of Constructed Facilities* 34(1).

Jaafaru, Hussaini, and Bismark Agbelie. 2022. "Bridge Maintenance Planning Framework Using Machine Learning, Multi-Attribute Utility Theory and Evolutionary Optimization Models." *Automation in Construction* 141:104460.

Kirshen, Paul, Kelly Knee, and Matthias Ruth. 2008. "Climate Change and Coastal Flooding in Metro Boston: Impacts and Adaptation Strategies." *Climatic Change* 90(4):453–73.

Koks, Elco E., Brenden Jongman, Trond G. Husby, and Wouter JW Botzen. 2015. "Combining Hazard, Exposure and Social Vulnerability to Provide Lessons for Flood Risk Management." *Environmental Science & Policy* 47:42–52.

Liang, Fayun, Chen Wang, Maosong Huang, and Yu Wang. 2017. "Experimental Observations and Evaluations of Formulae for Local Scour at Pile Groups in Steady Currents." *Marine Georesources & Geotechnology* 35(2):245–55.

Liang, Fayun, Chen Wang, and Xiong Yu. 2019. "Performance of Existing Methods for Estimation and Mitigation of Local Scour around Bridges: Case Studies." *Journal of Performance of Constructed Facilities* 33(6):4019060.

Liu, Kaijian, and Nora El-Gohary. 2022. "Bridge Deterioration Knowledge Ontology for Supporting Bridge Document Analytics." *Journal of Construction Engineering and Management* 148(6).

Manafpour, Amir, Ilgin Guler, Aleksandra Radlinska, Farshad Rajabipour, and Gordon Warn. 2018. "Stochastic Analysis and Time-Based Modeling of Concrete Bridge Deck Deterioration." *Journal of Bridge Engineering* 23(9).

Mclemore, S., S. Zendegui, J. Whiteside, M. Sheppard, M. Gosselin, H. Demir, P. Passe, and M. Hayden. 2010. "Unknown Foundation Bridges Pilot Study." *Federal Highway Administration & Florida Department of Transportation*.

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, and Georg Ostrovski. 2015. "Human-Level Control through Deep Reinforcement Learning." *Nature* 518(7540):529–33.

Mohammadi, Alireza, and Tamer El-Diraby. 2021. "Toward User-Oriented Asset Management for Urban Railway Systems." *Sustainable Cities and Society* 70:102903.

National Bridge Inventory Database. n.d. "National Bridge Inventory 2020." Retrieved June 2, 2023 (https://www.fhwa.dot.gov/bridge/nbi.cfm).

NCHRP report. 2006. *Risk-Based Management Guidelines for Scour at Bridges with Unknown Foundations*. Transportation Research Board of the National Academies Washington, DC.

Pearson, Dave, Stuart Stein, and J. Sterling Jones. 2002. *HYRISK Methodology and User Guide*. US Department of Transportation, Federal Highway Administration.

Pizarro, Alonso, Salvatore Manfreda, and Enrico Tubaldi. 2020. "The Science behind Scour at Bridge Foundations: A Review." *Water* 12(2):374.

Plough, Alonzo, Jonathan E. Fielding, Anita Chandra, Malcolm Williams, David Eisenman, Kenneth B. Wells, Grace Y. Law, Stella Fogleman, and Aizita Magaña. 2013. "Building Community Disaster Resilience: Perspectives from a Large Urban County Department of Public Health." *American Journal of Public Health* 103(7):1190–97.

Ray, Richard D., and Grant Foster. 2016. "Future Nuisance Flooding at Boston Caused by Astronomical Tides Alone." *Earth's Future* 4(12):578–87.

Richardson, E. V. 1987. "FHWA Technical Advisory-Scour at Bridges (Draft Report)." *US Department of Transportation, Federal Highway Administration*.

Scawthorn, Charles, Paul Flores, Neil Blais, Hope Seligson, Eric Tate, Stephanie Chang, Edward Mifflin, Will Thomas, James Murphy, and Christopher Jones. 2006. "HAZUS-MH Flood Loss Estimation Methodology. II. Damage and Loss Assessment." *Natural Hazards Review* 7(2):72–81.

Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. "Proximal Policy Optimization Algorithms." *arXiv Preprint arXiv:1707.06347*.

Sinha, Kumares C., Samuel Labi, and Bismark RDK Agbelie. 2017. "Transportation Infrastructure Asset Management in the New Millennium: Continuing Issues, and Emerging Challenges and Opportunities." *Transportmetrica A: Transport Science* 13(7):591–606.

Stein, Stuart, and Karsten Sedmera. 2006. *Risk-Based Management Guidelines for Scour at Bridges with Unknown Foundations*. Transportation Research Board of the National Academies Washington, DC.

Sturm, Terry W. 2001. *Open Channel Hydraulics*. Vol. 1. McGraw-Hill New York.

Tao, Weifeng, Peihui Lin, and Naiyu Wang. 2021. "Optimum Life-Cycle Maintenance Strategies of Deteriorating Highway Bridges Subject to Seismic Hazard by a Hybrid Markov Decision Process Model." *Structural Safety* 89:102042.

Tuccillo, Joseph V., and Seth E. Spielman. 2022. "A Method for Measuring Coupled Individual and Social Vulnerability to Environmental Hazards." *Annals of the American Association of Geographers* 112(6):1702–25.

Wei, Shiyin, Yuequan Bao, and Hui Li. 2020. "Optimal Policy for Structure Maintenance: A Deep Reinforcement Learning Framework." *Structural Safety* 83:101906.

Xiong, Wen, C. S. Cai, Rongzhao Zhang, Huiduo Shi, and Chang Xu. 2023. "Review of Hydraulic Bridge Failures: Historical Statistic Analysis, Failure Modes, and Prediction Methods." *Journal of Bridge Engineering* 28(4):03123001.

Xu, Gaowei, and Fengdi Guo. 2022. "Sustainability-Oriented Maintenance Management of Highway Bridge Networks Based on Q-Learning." *Sustainable Cities and Society* 81:103855.

Yang, David Y. 2022a. "Adaptive Risk-Based Life-Cycle Management for Large-Scale Structures Using Deep Reinforcement Learning and Surrogate Modeling." *Journal of Engineering Mechanics* 148(1):04021126.

Yang, David Y. 2022b. "Deep Reinforcement Learning–Enabled Bridge Management Considering Asset and Network Risks." *Journal of Infrastructure Systems* 28(3):04022023.

Yang, David Y., and Dan M. Frangopol. 2018. "Risk-Informed Bridge Ranking at Project and Network Levels." *Journal of Infrastructure Systems* 24(3):04018018.

Yang, Yifan, S. Thomas Ng, Frank J. Xu, and Martin Skitmore. 2018. "Towards Sustainable and Resilient High Density Cities through Better Integration of Infrastructure Networks." *Sustainable Cities and Society* 42:407–22.

Zhang, Guojing, Yongjian Liu, Jiang Liu, Shiyong Lan, and Jian Yang. 2022. "Causes and Statistical Characteristics of Bridge Failures: A Review." *Journal of Traffic and Transportation Engineering (English Edition)* 9(3):388–406. doi: 10.1016/j.jtte.2021.12.003.

Zhang, Ning, and Alice Alipour. 2020. "Two-Stage Model for Optimized Mitigation and Recovery of Bridge Network with Final Goal of Resilience." *Transportation Research Record* 2674(10):114–23. doi: 10.1177/0361198120935450.

Zhang, Weili, and Naiyu Wang. 2017. "Bridge Network Maintenance Prioritization under Budget Constraint." *Structural Safety* 67:96–104.

Zhang, Zhibo, Bismark R. Agbelie, and Samuel Labi. 2015. "Efficiency Measurement of Bridge Management with Data Envelopment Analysis." *Transportation Research Record* 2481(1):1–9.

Zhou, Qi-Neng, Ye Yuan, Dong Yang, and Jing Zhang. 2022. "An Advanced Multi-Agent Reinforcement Learning Framework of Bridge Maintenance Policy Formulation." *Sustainability* 14(16):10050.