

Semi-Markov process-driven maintenance scheduling for Tainter gate system considering multiple limit states

Structural Health Monitoring
2025, Vol. 24(4) 2031–2051
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14759217241277969
journals.sagepub.com/home/shm
 Sage

John Thedy¹, Kuo-Wei Liao² and Yi-Ting Hung³

Abstract

In extended periods of operation, Taiwan's reservoir electromechanical systems increasingly require substantial maintenance. This research adopts the semi-Markov process, which accommodates non-exponential distribution of state durations, to formulate optimal maintenance strategies for Tainter gate systems that are noted for their prolonged dormancy and significant operational uncertainties. The methodology initiates with the estimation of failure probabilities across four condition states, analyzing deterioration through the Weibull distribution for both general and latent limit states. The general limit state accounts for time-induced deterioration and effects of dormancy using a Bayesian–Weibull first-order reliability method, while the latent limit state addresses activation failures. Employing the semi-Markov process, an annual transition matrix is computed and combined with failure probabilities to assess the Tainter gates' system reliability. To identify the most efficient maintenance schedule, a genetic algorithm is applied, targeting the minimization of failure probabilities for both limit states below predefined thresholds and cost reduction. Numerical simulations validate the framework's efficacy, demonstrating its potential to enhance maintenance planning objectivity and decrease dependence on subjective assessments. The findings highlight the predominance of the general limit state in dictating system failure and underscore the risk Tainter gates face during transition from dormancy to activation, emphasizing the need for thorough monitoring.

Keywords

Maintenance, semi-Markov, Bayesian Weibull, FORM, optimization

Introduction

In Taiwan, the Water Resources Agency, as the governing body for water management, has digitized historical maintenance records for numerous reservoirs and established an information platform that enables the search and statistical analysis of the distribution of repair frequencies for various components. Hydraulic gates and their electromechanical control equipment are crucial and long-standing facilities within reservoirs. Based on the concept of the lifecycle, it is essential to ensure the safety and functionality of the gates and their control equipment. Due to prolonged or frequent usage, the hydraulic-gate-related facilities inevitably experience wear and aging, making appropriate maintenance management strategies vital for extending the service life of these facilities. Currently, routine inspection operations are predominantly conducted through manual visual inspections as a form of

preventative maintenance. To advance to more sophisticated predictive maintenance, concrete, scientifically based, and quantifiable data are necessary. Manual inspections, limited by the engineers' experience, slightly lack objectivity and do not offer real-time capabilities. To mitigate human errors, the installation of monitoring instruments has become a common solution, enhancing the quality of the existing database with quantifiable data. Moreover, contemporary

¹National Taiwan University, Taipei, Taiwan

²Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan

³Purdue University, West Lafayette, IN, USA

Corresponding author:

Kuo-Wei Liao, Department of Bioenvironmental Systems Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan.

Email: kliao@ntu.edu.tw

monitoring instruments, often integrated with various communication technologies such as the Internet of Things, can transmit onsite data in real time, reducing the inconvenience of onsite inspections while maintaining immediacy. Consequently, this study aims to transform monitoring data into scientifically based information for reliability analysis to identify failure modes, serving as a foundation for developing maintenance strategies to prevent failure incidents. This provides real-time information on hydraulic gates and their electromechanical control equipment, serving as a reference for maintenance management by agency personnel.

The Tainter gate, a critical element within water management frameworks, fulfills a crucial function in mitigating floods, facilitating irrigation, and generating hydroelectric power. These radial gates, crafted for high efficiency and robustness, are adept at handling significant hydraulic pressures, rendering them exceptionally suitable for modulating water flow within dams and waterways. Through precise regulation of water discharge, Tainter gates play a vital role in averting flood situations, securing a reliable water source for agricultural activities, and sustaining optimal conditions for the generation of energy. Within the context of Taiwan's water infrastructure, the Tainter gate stands out as a pivotal component. Distinguished from other manufactured equipment, Tainter gates exhibit a unique operational characteristic: they spend a greater portion of time in a dormant state relative to their "active periods." In this study, the "active period" refers to the time when the gate swings up (typically) to allow water to discharge downstream or when the gate swings down to dam the upstream pool again. Nonetheless, the imperative for these systems to remain operational at all times is underscored by their integral role in water management systems. Kirubakaran et al.¹ delineate four distinct maintenance strategies, as follows: corrective maintenance (CM) addresses failures as they occur, optimizing profit margins by reacting to issues, representing the initial approach in industrial maintenance. Time-based preventive maintenance (TM) schedules interventions to avoid unexpected failures, with its costs planned within the system's design life, and has been compared with CM in studies like those by Stenström et al.² Condition-based maintenance (CBM) continuously monitors system health to perform maintenance when specific degradation levels are reached, often using probabilistic methods for assessment.^{3,4} Predictive maintenance anticipates future system conditions through numerical models to guide preemptive maintenance,⁵ integrating a forward-looking approach to system upkeep.

To ascertain the most appropriate maintenance strategy for a given system, numerous scholars have

employed the analytical hierarchy process (AHP).^{6,7} Bevilacqua et al.⁷ incorporated multiple factors into their AHP analysis, including safety, the criticality of machinery, maintenance costs, frequency of failure, duration of downtime, and operational conditions. The assignment of maintenance strategies to individual machines is predicated on their respective AHP scores. Furthermore, various studies have advocated for the adoption of hybrid maintenance policies, such as integrating CBM with opportunistic maintenance (OM),^{8–10} or combining TM with OM.^{11,12} Regarding the determination of optimal maintenance intervals or schedules, a plethora of methodologies has been proposed. The predominant approach involves scheduling based on the optimization of a cost function,^{13–18} with a general cost function typically formulated as depicted in Equation (1).

$$\text{Min: } C(T) = \sum_{i=1}^{n_M} \sum_{j=1}^{n_C} \frac{C_j}{T_i} \quad \text{s.t.: } P_f \leq P_{\text{limit}} \quad (1)$$

where $C(T)$ is the total cost function, n_M is the total number of maintenance events, T_i is the time interval, n_C is the total number of actions per maintenance event, and C_j is the measure of cost per action. Note that each maintenance event may include more than one action. The optimization of the cost function $C(T)$ is pursued with the stipulation that the system's failure probability (P_f) is maintained beneath a predetermined threshold. In the literature, the costs associated with replacement and preventive maintenance are commonly addressed.^{13,14} It is posited that, at a certain juncture in the design life, the aggregate cost of replacement may outweigh the costs incurred from ongoing, imperfect maintenance at regular intervals. Consequently, an optimization of the maintenance interval duration and its corresponding costs becomes imperative. Diverging from conventional analyses that predominantly focus on replacement and maintenance expenses, some scholars, such as Vaurio,^{15,16} have devised a cost function that delineates the ratio of cost components relative to the duration of maintenance intervals, encompassing accidental and additional repair costs, alongside preventive maintenance and testing expenses. Furthermore, research by Ahmadi et al.¹⁷ extends the consideration of cost components to include inspection, repair, restoration, accidental, and downtime costs. Uniquely, Nguyen et al.¹⁸ evaluate the implications of adopting novel technologies. They contemplate the scenarios of inaction, preventive maintenance, or the substitution of deteriorated machinery with advanced technology, asserting that such technological advancements are presumed to enhance production efficacy. The postponement of

technological upgrades is thus inferred to adversely affect machine productivity in comparison to the deployment of newer technologies. Hu and Du¹⁹ introduce an evaluation of a lifetime cost optimization model that integrates physics-based reliability analysis. In their framework, total costs are categorized into initial design costs and maintenance costs. Maintenance costs are further divided into preventive and CM. The possibility of the system entering preventive or CM is also considered. Two types of design variables are examined: the first type includes the dimensions of components and system configurations, and the second type pertains to product operations, such as the number of preventive maintenances. They also consider warranty and post-warranty periods to better reflect the real conditions of system design life.

In addition to cost considerations, the optimization function denoted as Equation (1) also emphasizes the significance of a reliability constraint function. A broad range of formulations has been introduced to accurately encapsulate the uncertainties intrinsic to the system. The determination of the deterioration state and its associated probability of failure is critical to ensure the high quality of the optimization outcomes. With the increasing adoption of CBM, probabilistic approaches are frequently employed as a risk threshold to ascertain the necessity for maintenance interventions. A prevalent method for risk evaluation involves the application of the Weibull probability density function (PDF) to compute the likelihood of failure. An illustrative example of this approach is the methodology adopted by the U.S. Army Corps of Engineers (USACE), which utilizes the hazard rate $h(t)$ to evaluate the degree of hazard risk associated with a component, as specified in Equation (2).²⁰ This method facilitates a comprehensive risk assessment, enabling the identification and prioritization of maintenance actions based on the probabilistic risk of component failure, as shown below.

$$h(t) = K_1 K_2 K_3 \frac{f(t)}{1 - F(t)} = K_1 K_2 K_3 \times \beta \frac{(td)^{\beta-1}}{\alpha^\beta} \quad (2)$$

where $f(t)$ and $F(t)$ is the PDF and the cumulative density function (CDF) of Weibull distribution, respectively. The Weibull distribution, characterized by its scale parameters, is employed for the PDF, in which α and β are the scale and shape parameters, respectively. K_1 , K_2 , and K_3 are factors related to the environment, demand of components, and temperature effect, respectively.

Additionally, the variable d , representing the ratio of operating time to total time, is used to account for the dormancy effect on the system's reliability. Maintenance optimization studies frequently leverage

the Weibull failure probability and its variants as a probabilistic constraint to model the deterioration process and failure likelihood of systems.^{21,22} Such studies often segment the design life of a machine system into discrete states, each representing a specific level of system deterioration and associated failure probability. Transition probability methods, including Markov chains and their variants, are commonly utilized to simulate these deterioration states, offering a systematic approach to model the progression of wear and tear in systems. This probabilistic modeling extends beyond machinery, with applications in structural and bridge systems where deterioration rates are similarly segmented into various states by researchers.^{23,24} For instance, Sobanjo²³ defines four different condition scores (CSs) for bridges based on deterioration levels, employing the Weibull PDF to estimate the sojourn time probabilities across these states. This methodology enables the determination of transition probabilities between states, informing the optimization of maintenance intervals and risk assessments. Moreover, studies have incorporated semi-Markov processes as a probability constraint for maintenance optimization, demonstrating its utility in applications such as the determination of optimal maintenance intervals, where different states of deterioration are identified and analyzed.^{25,26} In the work by Kalantarnia et al.,²⁷ the system is divided into six states of deterioration, with the failure probability of each state calculated and used as a constraint in the optimization process. To enhance the accuracy and relevance of these models, some studies also apply Bayesian methods to update the Weibull parameters, allowing for the predictive model to be refined as new assessment information becomes available. This approach ensures that the optimization of maintenance schedules and risk assessments remains aligned with the most current understanding of the system's condition and performance.^{28,29} Vega et al.³⁰ conducted two types (static and dynamic) of optimal maintenance decisions for miter gates, in which semi-Markov transition probability was employed to define six distinct condition states. Notably, a strain measurement device is affixed to the miter gate to gather strain data for developing a numerical simulation model. This model is then used to estimate the damage state, and Bayesian methods are applied to update the condition rating.

Although numerous studies utilize semi-Markov processes to compute transition probability matrices, their application within the context of Tainter gate systems has not been documented at the time of writing this publication. For instance, a search in the Web of Science database for literature combining "hydraulic gate" and "semi-Markov" yields no results. Expanding the search criteria to include "hydraulic gate" and "Markov" results in only four publications.^{31–34} These

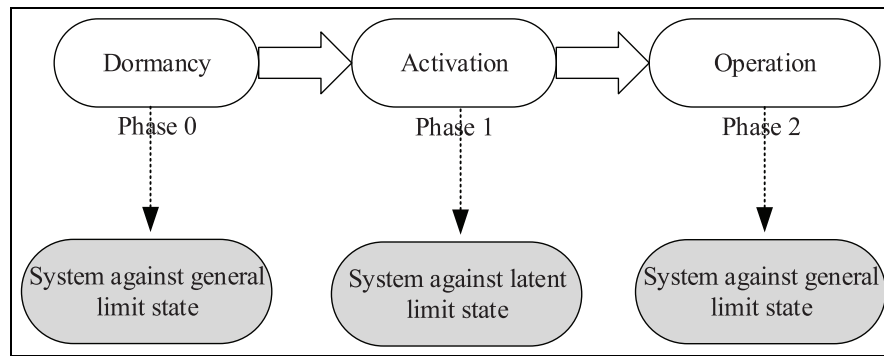


Figure 1. Limit states considered in this study.

studies incorporate the concept of Markov chains, yet do not employ semi-Markov processes to calculate time-varying transition probability matrices. The innovation of this study lies not only in the introduction of semi-Markov processes but also in the integration of these processes with the Weibull distribution to calculate failure probabilities, where the Weibull distribution is derived from Bayesian updating. The updating process utilizes first-order reliability analysis (FORM) to compute component probabilities. While these analytical tools are frequently used by researchers, their integration into a comprehensive framework specifically applied to Tainter gates is non-existent.

Many researchers, as previously mentioned, have proposed maintenance strategies that optimize costs while ensuring reliability. These studies each focus on distinct aspects, presenting varying analytical frameworks. For instance, some research primarily focuses on systems where operational periods far exceed dormancy periods, which are not applicable to Tainter gate systems. For systems that are in a prolonged dormant state, many studies integrate optimization with reliability calculations, considering the impact of system degradation on system performance. For example, Hu and Du¹⁹ uses FORM to analyze component and system reliability, employing their own sampling approach to convert time-dependent reliability analysis into a time-invariant format. Meanwhile, Vega et al.³⁰ combine Bayesian theory with finite-element models and operational condition assessment to derive Markov transition probabilities, thus calculating time-dependent reliability under different states.

This study closely aligns with Vega's approach³⁰ but has several distinctions. Both our study and Vega's employ Weibull distributions for time-dependent reliability analysis, differing from Hu's approach.¹⁹ Although our study also uses FORM, it is merely a process to update Weibull parameters. While both our study and Vega use Bayesian theory, Vega et al.³⁰ update the error ratio to obtain Markov transition

probabilities, whereas our application of Bayesian theory updates Weibull parameters. A novel aspect of our study is the introduction of semi-Markov processes to compute time-varying Markov transition probabilities, providing a more systematic calculation of state changes over time. Additionally, for the Tainter gate system, this study proposes a new optimized maintenance schedule that considers long dormancy periods. It meticulously distinguishes between two types of failures and outlines distinct solutions for each, emphasizing the importance of differentiating maintenance types.

The structure of this article is outlined as follows: section "Problem definition and research objective" elaborates on the problem definition and objectives of this study. Section "Methodology" describes the methods used to calculate the CDF for the two limit state functions discussed. Section "Numerical simulation of Tainter gate" presents a numerical example to demonstrate the effectiveness of the proposed framework. Finally, Section "Conclusion" offers a brief conclusion of the study.

Problem definition and research objective

Every system has unique characteristics, which make certain maintenance policies more suitable. For instance, implementing a CM policy on a Tainter gate system may not be an effective strategy, as the system cannot afford to fail even once. The failure at Folsom Dam illustrates this point well. Additionally, many Tainter gates are designed to endure long periods of dormancy, typically becoming operational only during the rainy season. This study puts focus on a gate system with three states: dormancy (Phase 0), activation (Phase 1), and operation (Phase 2), as illustrated in Figure 1. Authors observed that a long dormancy system commonly will have two governing types of failure. As mentioned earlier, two types of failure limit states considered here are general and latent failure

limit states. The general limit state defined as the failure relates to durability and operational limit states. The former focuses on the deterioration of materials or components over time due to environmental exposure, wear and tear, or other factors that reduce the lifespan of the gate system. The latter concerns the operational performance and reliability of gate system, where the limit is reached when the system can no longer perform its intended function efficiently or safely. The latent limit state considers the failure probability during activation.

This study assumes that systems initiate from Phase 0 (dormancy), where a prolonged period of inactivity introduces a failure probability preventing entry into Phase 1 (activation). In the context of Tainter gates or spillways, such failures might arise from electrical circuit issues, lubrication deficiencies, or other latent defects hindering activation. This type of failure is defined as latent failure limit state. Upon successful machine activation, it is then subjected to operational demand loads. This research consolidates Phases 0 and 2, essentially categorizing limit state functions into two main periods: activation and non-activation. The rationale for merging Phases 0 and 2 stems from previous studies, such as those by the USACE,²⁰ which have successfully integrated these limit states. Practical observations have shown that field engineers are particularly concerned with the failure probabilities during the activation phase, hence the introduction of a specific latent limit state to address this concern.

This study aims to develop a framework that prevents excessive or insufficient maintenance for systems with prolonged dormancy periods. It differentiates maintenance into two categories: Type I, which includes cleaning, repairing, updating, and replacing parts, and Type II, which focuses on inspection. According to practical operations, reservoir management units operate the Tainter gates several times each year, even in the absence of flooding (when there is no actual operational need), to ensure smooth functioning. This process of operation is referred to as inspection maintenance, or Type II maintenance, in this article. Mathematically, Type I maintenance is posited to extend the expected life of the machinery, whereas Type II maintenance serves to assess the health condition of the system through methods like sensor monitoring or visual inspections, indicating the machine's deterioration state. Given the brief nature of the activation period, it is impractical to conduct Type I maintenance during activation. This study simplifies, without losing representativeness, by assuming that Type I maintenance enhances reliability against failures associated the general limit state, while Type II maintenance has impact on the probability of latent failures during activation. Moreover, Type I maintenance

Table 1. Input for BW-FORM.

CS from inspection	BW-FORM input			
	Mean capacity at year 0 (μ_{c0})	Demand lifespan (μ_d)	Capacity cov	Demand cov
CS 4	100	40.0	1.0	1.0
CS 3	75			
CS 2	50			
CS 1	25			

BW-FORM: Bayesian–Weibull first-order reliability method;
CS: condition score; cov: coefficient of variation.

influences reliability during dormant and operational periods, a factor incorporated through the proposed Bayesian–Weibull first-order reliability method (BW-FORM), with details provided in the Methodology section. Another assumption is that a Tainter gate system experiences four different CS (for details, refer to Table 1) throughout its design life. Semi-Markov processes are utilized to model the transition probabilities from a healthy CS to lower states. For each CS, general (e.g., weariness) and latent failure probabilities are defined using the Weibull PDF, with general failure limit state parameters updated via Bayesian methods based on inspection data. The failure probabilities derived from general and latent limit states are then integrated with semi-Markov to assess the system's overall/total failure probability.

To optimize the scheduling of Type I and II maintenance, a genetic algorithm (GA) is employed, with the assumption that Type I maintenance rejuvenates the machine's mean life and Type II (inspection) reduces latent failure risk. The optimization aims to minimize maintenance costs while ensuring that the probabilities of general and latent limit states do not exceed a predefined threshold. The outcome is an optimal maintenance schedule within an estimated budget. The cost of failure is not considered here. Section “Methodology” will detail the proposed methodology, including the definition and process of generating failure probabilities for general and latent failure limit states.

Methodology

The proposed framework initiates by assessing the CS number, which symbolizes its health status. Initially, it constructs the failure probabilities for both general and latent limit states across CS levels. The general failure limit state employs the BW-FORM method^{28,29,35,36} for its probability calculation per CS, whereas the latent state's failure probability assumes a Weibull PDF. Sections “General limit state using BW-FORM”

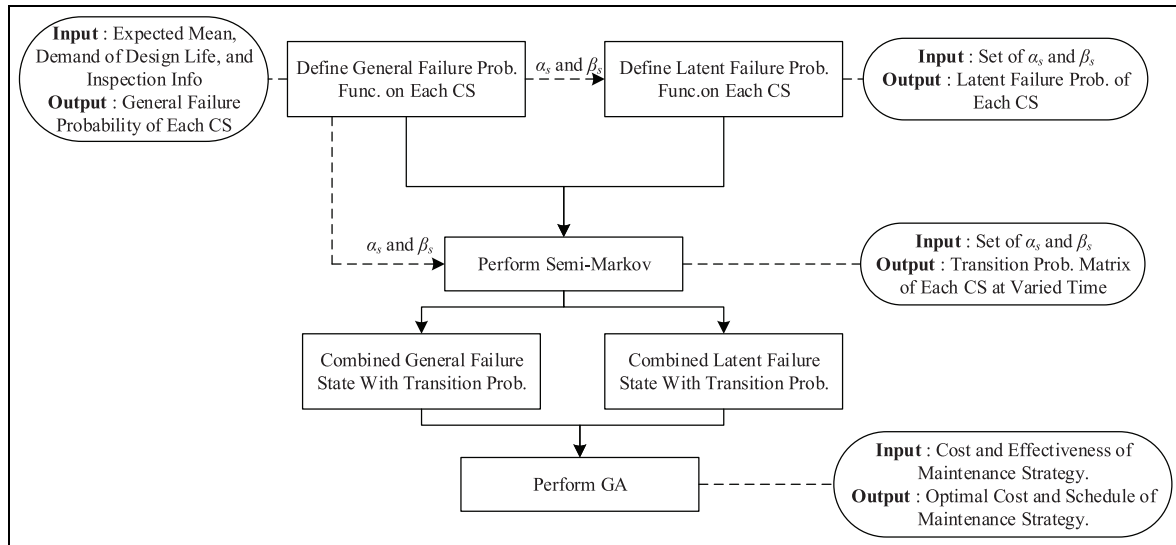


Figure 2. General scheme of the proposed method.

and “Latent failure limit state” elaborate on the methodologies for computing these probabilities, respectively. Subsequently, sojourn time data for each CS inform the second phase, utilizing a semi-Markov process to translate these data into CS transition probabilities. As depicted in Figure 2, latent failure CDF requires inputs of Weibull parameters (α_s and β_s), which are derived from an updated Weibull distribution via BW-FORM. This creates an interaction between the general and latent limit states. Similarly, fitting sojourn times to a Weibull PDF necessitates selecting fitting Weibull parameters (α_s and β_s), derived from an updated Weibull CDF via BW-FORM (section “General limit state using BW-FORM”). Section “Semi-Markov process” will detail the semi-Markov formulation. The subsequent stage integrates the failure probabilities with occurrence probabilities from semi-Markov, serving as input for GA optimization to schedule maintenance. This optimization aims to prevent limit state failure probabilities from exceeding thresholds while minimizing overall costs, detailed in section “Optimization of maintenance schedule.” Additionally, this section will discuss maintenance’s impact on reducing failure probabilities.

General limit state using BW-FORM

This study applies the Weibull distribution as the failure PDF for hydraulic electromechanical components, commonly used due to its flexibility in modeling life data. To incorporate the dormant effect, Bayesian updating is utilized to refine the health status

assessment based on inspection data. This approach allows for the integration of new information to adjust the failure probability estimates, thereby providing a more accurate reflection of the component’s current condition. The methodology for Bayesian updating in conjunction with Weibull distribution parameters is detailed further in this section (Figure 3), highlighting how inspection data can effectively inform and refine health status predictions. Bayesian updating has been extensively utilized in various engineering applications.^{28,29} In this study, the results obtained from the FORM³⁵ are treated as observations to update the parameters of the Weibull distribution. A similar analytical framework is presented in the work of Tian et al.,³⁶ although they employed a response surface method-based reliability analysis instead of the FORM used in this study.

Figure 3 shows the general scheme of the proposed BW-FORM. The step-by-step procedure is described as follows:

Step 1: Initially, it is posited that the probability of failure for a specified lifespan (t_{life}) adheres to a Weibull PDF, characterized by parameters α_w and β_w , as delineated in Equation (3). The parameters α_w and β_w are determined through calculations based on the residual life, which is estimated from historical maintenance records or expert assessments. This estimation process is considered to reflect the expert judgment within the framework of Bayesian analysis. The methodology involves initially setting the shape parameter (β_w) to a predefined value, followed by the application of the maximum likelihood estimation technique to

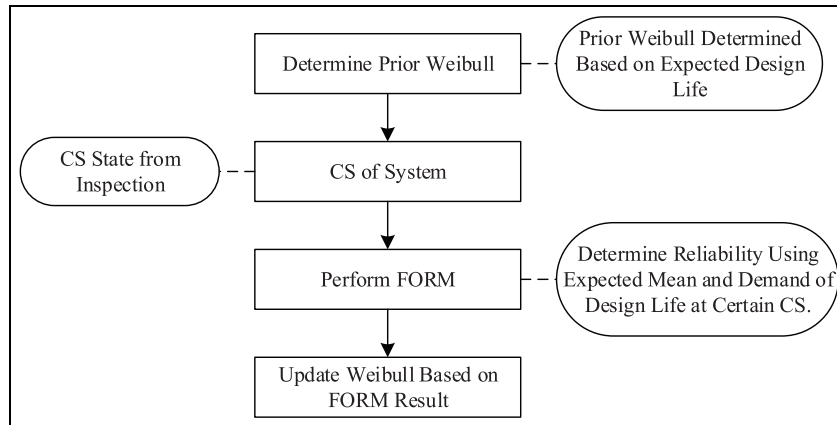


Figure 3. General scheme of the proposed BW-FORM.
BW-FORM: Bayesian–Weibull first-order reliability method.

ascertain the scale parameter (α_w). β_w is assumed to follow uniform distribution with lower and upper bound of 0.05 and 2, respectively.

$$f_w(t_{\text{life}}) = \frac{\beta_w}{\alpha_w} \left(\frac{t_{\text{life}}}{\alpha_w} \right)^{\beta_w - 1} e^{-\left(\frac{t_{\text{life}}}{\alpha_w} \right)^{\beta_w}}. \quad (3)$$

Step 2: The value of α_w in Equation (3) is assumed to follow inverse gamma distribution as shown in Equation (4) with mean value equal to obtained α_w at Step 1 and coefficient of variation (cov) = 1.

$$f_{\alpha_w}(\alpha_w) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{\alpha_w} \right)^{a+1} e^{-\left(\frac{b}{\alpha_w} \right)}. \quad (4)$$

Step 3: The failure probability is derived from the CS value using the FORM, as depicted in Equation (5). This equation incorporates μ_c and μ_d as the mean values, and σ_c and σ_d as the standard deviations, of the system's capacity and demand over its lifespan, respectively. β_{FORM} value of 0.5 is postulated to delineate the threshold of failure. Furthermore, it is hypothesized that the expected lifespan deteriorates over time, with the level of degradation articulated in Equation (6). The expected design life at year 0, denoted as μ_{c0} , is established based on the CS level. It is critical to acknowledge that FORM analysis extends from year $t = 0$ up to the year when β_{FORM} reaches 0.5, which is then designated as the year of failure (t_{fail}). The designation of $\beta = 0.5$ as the point of failure is primarily based on standards set by the USACE,³⁷ where $\beta = 1.0$ is defined as hazardous. Since the failure probability at the time of destruction should be higher (implying a smaller β), this study extends this concept by setting $\beta = 0.5$ as the point of failure.

$$\beta_{\text{FORM}} = \frac{\mu_c - \mu_d}{\sqrt{\sigma_c^2 + \sigma_d^2}} \quad (5)$$

$$\mu_c = \mu_{c0} - (100e^{(40-2t_{\text{life}})^{-0.05}} + 15). \quad (6)$$

Step 4: The failure year derived via FORM is employed to ascertain the corresponding α_w , with the assumption that this failure year serves as the new mean life while retaining the previously determined β_w . This process facilitates the calculation of a corresponding α_w , henceforth referred to as $c\alpha_w$.

Step 5: Revise the inverse gamma parameter. Given that α_w is assumed to be an inverse gamma random variable in Step 2, the posterior distribution of α_w is delineated in Equation (7).

$$f(\alpha_w | c\alpha_w) = \frac{f(c\alpha_w | \alpha_w, \beta_w) f(\alpha_w)}{\int f(c\alpha_w | \alpha_w, \beta_w) f(\alpha_w) d\alpha_w}$$

$$f(\alpha_w | c\alpha_w) = \frac{\left[\frac{\beta_w}{\alpha_w} \left(\frac{c\alpha_w}{\alpha_w} \right)^{\beta_w - 1} e^{-\left(\frac{c\alpha_w}{\alpha_w} \right)^{\beta_w}} \right] \left[\frac{b^a}{\Gamma(a)} \left(\frac{1}{\alpha_w} \right)^{a+1} e^{-\left(\frac{b}{\alpha_w} \right)} \right]}{\int \left[\frac{\beta_w}{\alpha_w} \left(\frac{c\alpha_w}{\alpha_w} \right)^{\beta_w - 1} e^{-\left(\frac{c\alpha_w}{\alpha_w} \right)^{\beta_w}} \right] \left[\frac{b^a}{\Gamma(a)} \left(\frac{1}{\alpha_w} \right)^{a+1} e^{-\left(\frac{b}{\alpha_w} \right)} \right] d\alpha_w} \quad (7)$$

$$\hat{a} = a + n$$

$$\hat{b} = b + \sum_{i=1}^n c\alpha_w^b \quad (8)$$

where posterior parameter of α_w remains characterized as an inverse gamma random variable; the updated inverse gamma parameter can be obtained using Equation (8), which is formulated based on the Bayesian approach. This equation utilizes the updated failure probability calculated in the previous step to determine the updated parameter. a and \hat{a} are the prior

and posterior shape parameters of the inverse gamma random variable, respectively. b and \hat{b} are the prior and posterior scale parameters of the inverse gamma random variable, respectively. n is the number of observed values and $c\alpha_w$ is the observed value. More specifically, “ n ” represents the number of times FORM is executed, and “ $c\alpha_w$ ” is the calculated “ a ” value based on the CS and FORM.

Step 6: Upon obtaining the necessary parameters, \hat{a} and \hat{b} , the updated α_w can be computed, given the prior assumption that α_w represents the mean value of the inverse gamma distribution parameters (\hat{a} and \hat{b}). As delineated from Steps 1–5, this computation is executed across various values of β_w , indicating the existence of a set of updated α_w ($\hat{\alpha}_w$). The posterior Weibull CDF is generated using values of $\hat{\alpha}_w$ and β_w , as illustrated in Equation (9). It is seen that the probability at fail year ($P(\hat{\alpha}_w, \beta_w)$), computed in Step 3, is utilized as a weighting factor. Formulation of $P(\hat{\alpha}_w, \beta_w)$ is shown in Equation (10).

$$F_w(t) = \frac{\sum_{\hat{\alpha}_w, \beta_w} \left(1 - e^{-\left(\frac{t}{\hat{\alpha}_w}\right)^{\beta_w}}\right) P(\hat{\alpha}_w, \beta_w)}{\sum_{\hat{\alpha}_w, \beta_w} P(\hat{\alpha}_w, \beta_w)} \quad (9)$$

$$P(\hat{\alpha}_w, \beta_w) = 1 - \text{abs}\left(\left[1 - e^{-\left(\frac{t_{\text{fail}}}{\hat{\alpha}_w}\right)^{\beta_w}}\right] - 0.632\right). \quad (10)$$

As illustrated in Figure 2, the scale and shape parameters derived from Equation (9) will be employed in the calculation of the latent failure probability in section “Latent failure limit state” and the semi-Markov transition probability in section “Semi-Markov process.” Note that Equation (9) represents a set of Weibull distributions, not a single Weibull distribution. For clarity and simplicity in expressing the trends between general and latent limit states, subsequent sections of this paper will approximate Equation (9) using a single Weibull distribution.

Latent failure limit state

The concept of a latent failure limit state encompasses hidden failures or undetected errors that result in the failure to activate a system. It is crucial to accurately differentiate between failures caused by wear and aging and those attributed to latent failure. Failures stemming from diminished performance due to component aging are classified within this study as general failure limit state. Conversely, issues arising from inadequate inspection, such as mismanagement in electrical circuitry or insufficient lubrication changes leading to activation failures despite the system’s design life still being within its productive years, can often be mitigated

through optimally scheduled inspections. For the purposes of this research, a straightforward methodology is employed to model the latent failure limit state. Similar to the approach outlined in section “General limit state using BW-FORM,” the failure probability is formulated for each CS and then integrated with semi-Markov transition probabilities. The latent failure limit state is postulated to adhere to a Weibull CDF, as indicated in Equation (11). The Weibull PDF is widely utilized in reliability analysis due to its flexibility, enabling effective modeling of various life data types. This study adopts it for the latent limit state, establishing Weibull parameters based on results from the general limit state. That is, there is a measurable relative relationship in risk between the two limit states. The correlation in risk is established via the scale parameter (α_l) for latent failures as outlined in Equation (11). As shown, the scale parameter (α_l) for latent failures is calculated as a function of the general limit state’s scale parameter (α_s) from the BW-FORM, adjusted by the risk ratio (n_l). When $n_l = 1$, it signifies that the CDFs for both general and latent failures are almost identical. In contrast, a higher n_l value indicates a comparatively greater risk of latent failures. The shape parameter (β_l) is determined based on expert judgment, with higher values assigned to systems with greater uncertainty, and lower values to those with less. In the numerical example, varying values of n_l are employed to compare and demonstrate their impact between the general and latent limit states.

$$F_l(t) = 1 - e^{-\left(\frac{t}{\alpha_l}\right)^{\beta_l}} \quad (11)$$

$$\alpha_l = \frac{\alpha_s}{n_l}.$$

Semi-Markov process

Multiple methodologies for computing the semi-Markov transition probability are documented in the literature.^{38,39} This research adopts the approach utilized in Sobanjo’s study,²³ where the semi-Markov transition probability is formulated as presented in Equation (12).

$$P_{ij}(t, t_0) = \delta_{ij} S_i(t, t_0) + \sum_k \left[p_{ik}^d(t_0) \int_0^t f_{ik}(x, t_0) P_{kj}(t - x) dx \right] \quad (12)$$

$$P_{kj}(t) = \delta_{kj} S_k(t, t_0) + \sum_r \left[p_{kr}^d(t) \int_0^t f_{kr}(x) P_{rj}(t - x) dx \right] \quad (13)$$

$$S_i(t) = 1 - \int_0^t f_i(t) dx \quad (14)$$

where $P_{ij}(t, t_0)$ denotes the transition probability from state i to j , given that the system has already spent t_0 time in state i prior to transitioning. The Kronecker delta constant, δ_{ij} , is introduced (where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise) to eliminate the equation of $\delta_{ij}S_i(t, t_0)$ during transitions. $\delta_{ij}S_i(t, t_0)$ is specifically reserved for scenarios where the system remains in its current state. $p_{ik}^d(t_0)$ is the PDF of the system moving from the current state at time t_0 , and $f_{ik}(x, t_0)$ denotes the PDF of the system transitioning from state i given that it has spent x time in state i . $P_{kj}(t - x)$ symbolizes the transition probability from state k to state j during the remaining time $(t - x)$. Equation (13) further elaborates $P_{kj}(t - x)$ in cases where there is an intermediary state between states k and j . Similar to Equation (12), δ_{kj} is the Kronecker delta constant, $S_k(t, t_0)$ represents the probability of the system remaining in state k , $p_{kr}^d(t)$ is the PDF of the system moving from the current state at time t , $f_{kr}(x)$ is the PDF of the system transitioning from state k given it has spent x time in state k , and $P_{rj}(t - x)$ indicates the transition probability from state r to state j during the remaining time $(t - x)$. The probability of the system remaining in its current state is defined in Equation (14), where $f_i(t)$ represents the PDF of the system transitioning from state i .

Equations (12) and (13) illustrate that the computation of semi-Markov transition probabilities involves a repetitive process across multiple states, lacking a closed-form solution when employing a Weibull PDF. Consequently, this research aligns with the assumptions made in several references,^{40,41} converting Equations (12) and (13) into a discrete format. This conversion is predicated on the premise that the probability of transitioning to more than one lower state simultaneously is negligibly small and thus can be disregarded. Equations (15) and (16) delineate the discrete methodology utilized in this study for calculating the transition probabilities to either a lower state or to remain in the current state. Within section “General limit state using BW-FORM,” an inverse calculation procedure is employed to determine the Weibull parameters for each CS level, which are subsequently applied within Equations (15) and (16).

$$P_{ii}(t) = S_i(t) = 1 - \int_0^t f_i(t) dx \quad (15)$$

$$P_{ij}(t) = \sum_k \sum_{x=1}^t [f_{ik}(x) F_{kj}(t - x)] \quad (16)$$

Optimization of maintenance schedule

The CDFs derived in Equations (10) and (11) are integrated with the semi-Markov transition probabilities determined in section “Semi-Markov process,” as illustrated in Equations (17) and (18). The CDFs in Equations (17) and (18) encapsulate the overall failure probability for both general and latent limit states, incorporating the occurrence probability of each CS and the corresponding failure probability for each CS. Figure 4 elucidates the maintenance procedures. It depicts the failure CDF of the system attributable to general, $F_w(t)$, and latent failures, $F_l(t)$, as calculated in Equation (17) or Equation (18). These are plotted against two types of x -axes: the first x -axis represents the expected system life in years, denoted by the lower-case t , and the second x -axis signifies the actual life in years, denoted by the uppercase T . This dual-axis representation facilitates a comprehensive understanding of the system’s maintenance schedule and its impact on mitigating failure risks.

$$F_{w-total}(t) = \sum_{CS=1}^4 F_w(t|CS) \times P_{4-CS}(t) \quad (17)$$

$$F_{l-total}(t) = \sum_{CS=1}^4 F_l(t|CS) \times P_{4-CS}(t). \quad (18)$$

Without maintenance Type I or II, the system’s age matches the actual year. Maintenance Type I rejuvenates the system’s age. As depicted in Figure 4, implementing maintenance Type I at T_{M2} , reduces the expected system age to t_{M1} , lowering the failure CDF accordingly. This principle also applies to subsequent maintenance at T_{M4} , allowing the system’s age to revert to t_{M2} . Figure 4 shows that without maintenance Type I or II, $F_w(t)$ or $F_l(t)$ approaches the failure threshold (F_w limit and F_l limit). Maintenance is scheduled when T reaches $r \times T_f$, where T is the actual year, T_f is the year failure thresholds are met, and r is a ratio within 0 and 1. The initial maintenance Type I or II is planned for T_{M2} (or T_{I2}), equating to $r_{M1} \times T_{Mf1}$. This necessitates predefining two sets of ratios, $\mathbf{r}_M = [r_{M1}, r_{M2}, \dots, r_{Mn}]$ for maintenance Type I and $\mathbf{r}_I = [r_{I1}, r_{I2}, \dots, r_{In}]$ for Type II, with n_M and n_I indicating the counts of each maintenance type. Importantly, maintenance Type I mandates concurrent Type II to ensure the system’s operational integrity. The total cost function is then derived using Equation (19).

$$C(T) = \sum_{i=1}^{n_M} \sum_{j=1}^{n_I} [C_i(1 + \phi)^{T_i} + C_j(1 + \phi)^{T_j}] \quad (19)$$

C_i and C_j represent the costs associated with maintenance Types I and II, respectively, with the inflation

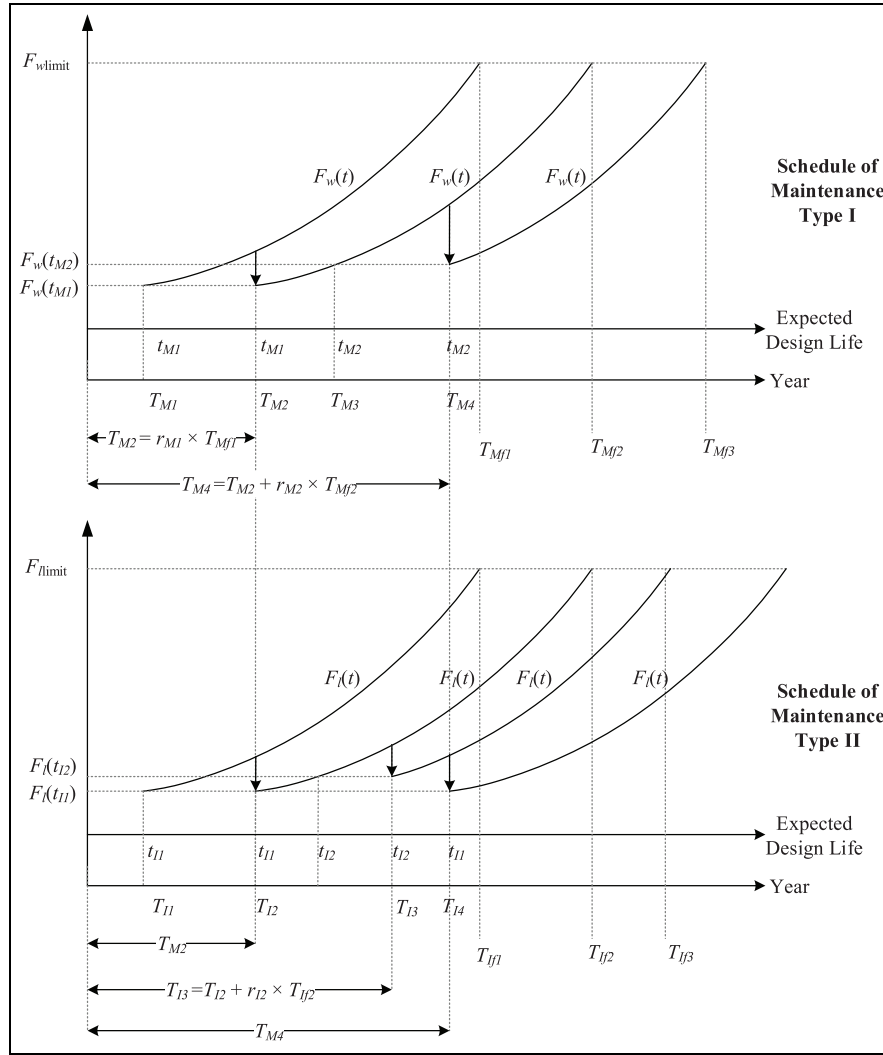


Figure 4. Effect demonstration of maintenance Type I and Type II.

rate accounted for to reflect cost escalation over time. The cost components C_i and C_j can be defined in monetary or normalized units. For simplicity in the numerical example in section “Numerical simulation of Tainter gate,” this study uses normalized cost values. GA⁴² is employed to minimize Equation (19). This study does not focus extensively on the optimization tool used. The choice of GA is due to its suitability for ordering optimization tasks, where it is often recognized for its effectiveness. However, it is important to acknowledge that using GA does not ensure the most efficient or optimal results. Any method that fits the proposed framework and meets user preferences can be considered appropriate. In addition to optimizing the maintenance schedule, the GA evaluates various strategies for maintenance Types I and II, indicated as M_{type} and I_{type} . The optimization parameters include r_M , r_I , M_{type} , and I_{type} , encompassing both the timing

and the strategies of maintenance. r_M is the ratio of maintenance time to failure time, ranging between 0 and 1. $r_M = 0$ means that immediately after a maintenance event, another maintenance is performed. Conversely, $r_M = 1$ means that maintenance is performed exactly when the system reaches the failure probability limit. For example, if the failure time is 10 years and $r_M = 0.1$, then Type I maintenance is performed in the first year. Since r_M is an optimized design variable, one can determine the optimal maintenance timing through r_M . The same explanation applies to r_I as well. For the GA optimization, a commonly adopted method is used, featuring two one-point crossovers. For the selection of parental chromosomes, a roulette wheel method is employed. Section “Numerical simulation of Tainter gate” will provide a numerical example to illustrate the application of this framework.

Table 2. Weibull parameters for latent failure for each CS.

CS from inspection	Statistical parameters	
	Scale (α_l)	Shape (β_l)
CS 4	18.7	4.0
CS 3	15.4	4.5
CS 2	11.8	4.5
CS 1	10.6	5.0

CS: condition score.

Table 3. Weibull parameters for Sojourn time on each CS.

Deterioration	Parameters derived from current example		Parameters for ideal conditions	
	Scale (α_s)	Shape (β_s)	Scale	Shape
CS 4	37.5	1.4	2.0	2.0
CS 3	30.9	1.2	4.0	1.5
CS 2	23.7	0.9	8.0	3.0
CS 1	21.2	0.8	6.0	3.0

CS: condition score.

Numerical simulation of Tainter gate

To validate the proposed framework, a numerical example is detailed in this section. Given the novelty of the semi-Markov approach and the concepts of general and latent limit states in Taiwan, simulating based on historical data poses a challenge due to the requirement for extensive infrastructure performance records. Therefore, this study employs a hypothetical scenario with plausible assumptions for demonstration. Specifically, a Tainter gate infrastructure with a designed lifespan of 40 years is selected for numerical simulation. Sections “System general limit state” and “System latent failure limit state” discuss the general and latent limit states of the system, respectively, with section “System general limit state” also elaborating on the construction of semi-Markov transition probabilities. Section “Optimization and discussion of maintenance scheduling” illustrates the application of metaheuristic optimization via GA to ascertain maintenance timings and strategies.

System general limit state

As outlined in section “Problem definition and research objective,” this study distinguishes between system failures caused by aging/wear and latent failures. The reliability assessment for wear-related failures does not consider latent failures, assuming the system will activate successfully when needed. This section defines the wear-related failure probability for each CS. Assumptions regarding the lifespan capacity and demand for each CS are summarized in Table 1. For evaluating the general limit state, the BW-FORM method introduced in section “Latent failure limit state” is applied. By integrating Equations (3) to (10) with the assumptions in Table 1, the failure probability for each CS is calculated, as depicted in Figure 5. The FORM calculation necessitates the input of the mean remaining life of the system (μ_{c0}), with a *cov* of 1.0 applied to both demand and capacity.

The system is projected to have a 40-year demand lifespan per CS. Annual capacity degradation is

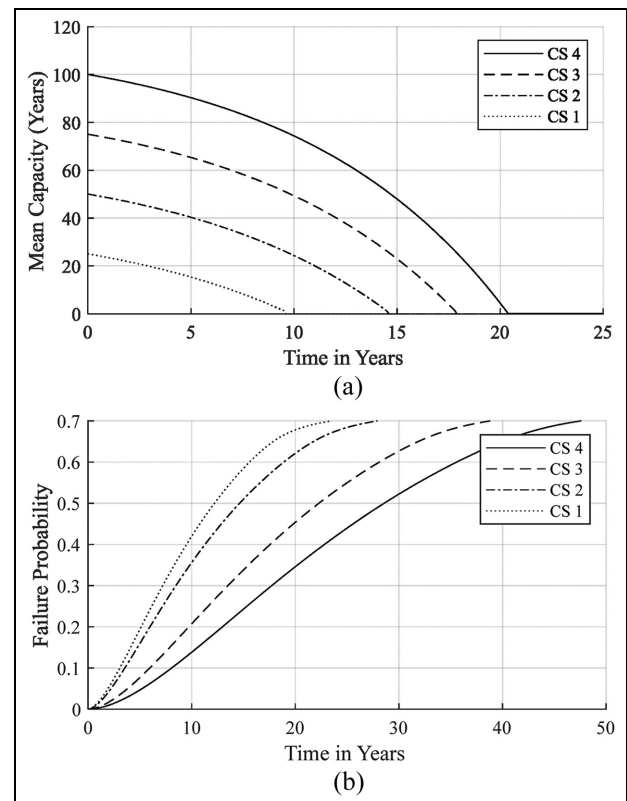


Figure 5. Construction of general failure probability using BW-FORM. (a) Mean remaining life (μ_c) at each year. (b) Weibull CDF failure probability computed from BW-FORM. BW-FORM: Bayesian–Weibull first-order reliability method; CDF: cumulative density function.

determined by Equation (6), as depicted in Figure 5(a), with the rate of mean capacity deterioration accelerating over time. Bayesian updated Weibull distributions for the designated CSs are illustrated in Figure 5(b). The aggregate failure probability, shown in Figure 8(a), is calculated by multiplying the Semi-Markov transition probabilities (detailed in Figure 7) with the Weibull failure probabilities from Figure 5(b). Furthermore, the results of the scale (α_s) and shape

(β_s) parameters presented in Figure 5(b) will be utilized in the calculation of the latent failure probability CDF in section “System latent failure limit state” and the semi-Markov transition probability in section “Semi-Markov transition probability.”

System latent failure limit state

Mirroring section “System general limit state” approach to general failure limit state, the latent failure limit state also accounts for the occurrence conditions of four distinct CSs. For each CS, failure probability is modeled using the Weibull distribution, with parameters for latent failure assumptions detailed in Table 2 and Figure 6. The scale parameter (α_l) is derived from the general limit state scale parameter (α_s), as formulated in Equation (11), with an assumption that $n_l = 4$. Meanwhile, the shape parameter is assigned based on reasonable assumptions. As shown in Figure 6, lower CS values correspond to increased

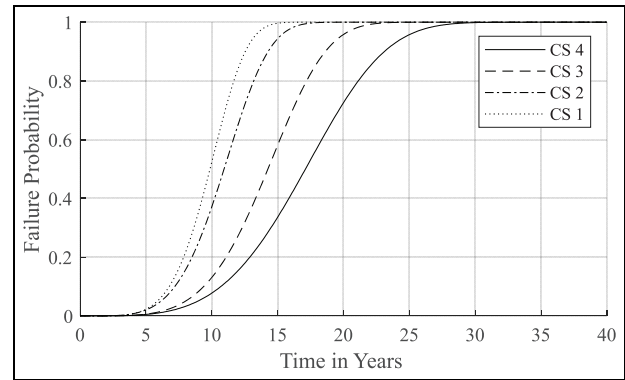


Figure 6. Construction of system latent failure probability using Weibull for each CS.
CS: condition score.

latent failure probabilities. These probabilities are then combined with semi-Markov transition probabilities to calculate the total/overall failure probability (Figure 8(b)).

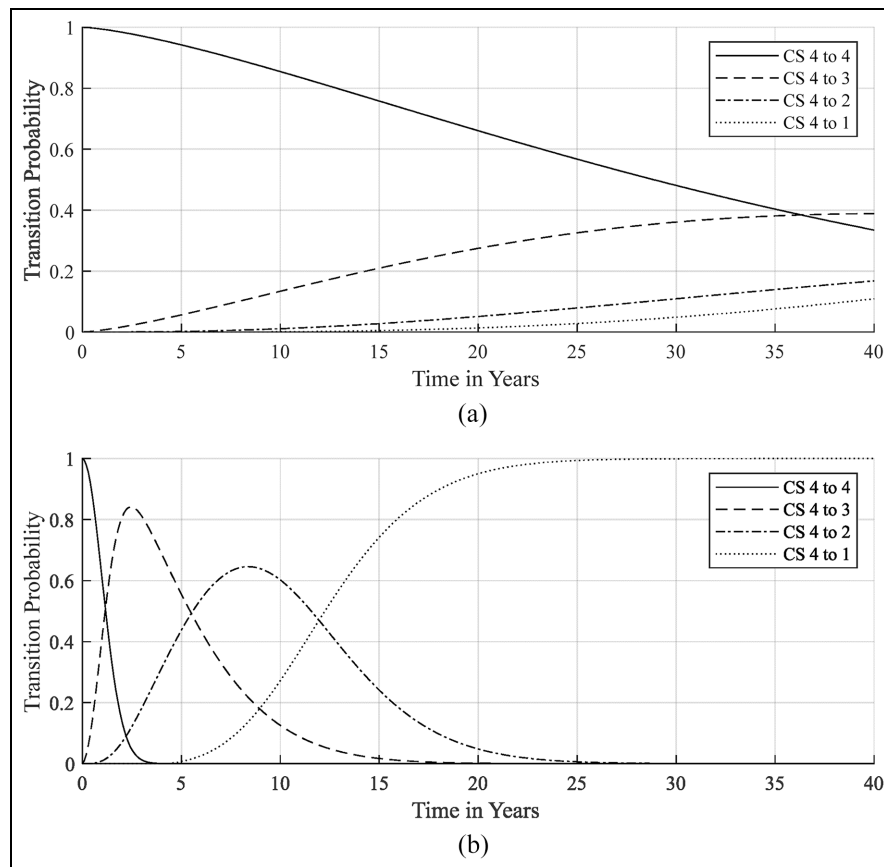


Figure 7. Transition probability from CS 1 to other CSs: (a) transition probability calculated using values from this example and (b) hypothetically derived ideal transition probability.
CS: condition score.

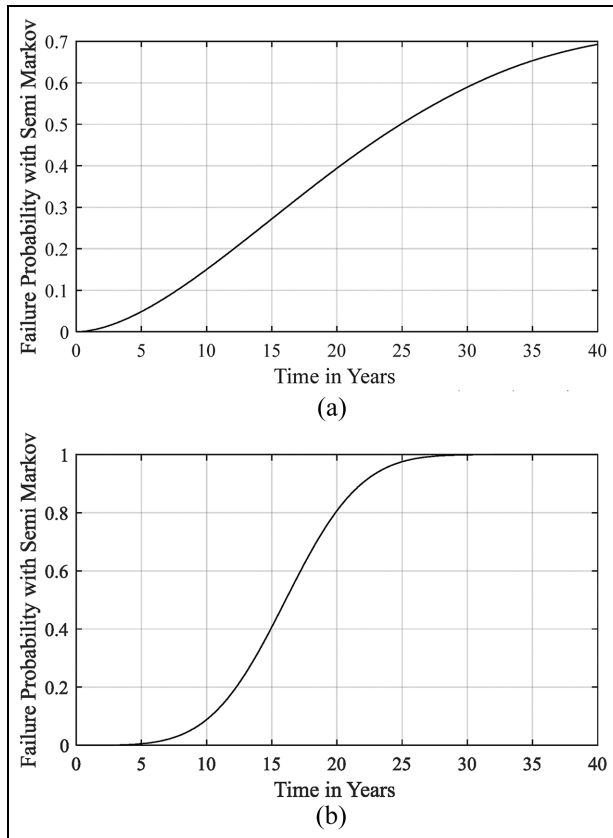


Figure 8. Construction of general and latent total failure probability. (a) Total failure probability from BW-FORM and Semi-Markov across CSs. (b) Total latent failure probability across CSs. CS: condition score.

Table 4. Maintenance parameters.

Maintenance Type I			Maintenance Type II		
Type	Effectiveness	Cost	Type	Effectiveness	Cost
I-1	0.95	1	2-1	0.95	1
I-2	0.85	2.5	2-2	0.85	2.5
I-3	0.75	3.5	2-3	0.75	3.5

Semi-Markov transition probability

This research posits that system deterioration is contingent on CS values throughout its lifespan. The distribution of each CS's sojourn time is presumed to follow a Weibull distribution, characterized by scale (α_s) and shape (β_s) parameters specified in Table 3, derived from inverse calculations of Figure 5(b). Employing Equations (15) and (16) with α_s and β_s allows for the construction of semi-Markov transition probabilities annually. Figure 7(a) illustrates the yearly distribution

Table 5. GA optimization description.

Description	
Objective	Minimized Equation (19) cost function
Constraint	- General failure limit <0.15 - Latent failure limit <0.15 - Concurrent execution of maintenance Types I and II
Design parameters	M_{type} , I_{type} , r_M , r_I

GA: genetic algorithm.

of transition probabilities. Initially, the Tainter gate system is certain (probability of 1.0) to be at CS 4. As time advances, the likelihood of remaining in CS 4 diminishes, with a concurrent rise in the transition probabilities to other CS levels. The shift toward CS 1 begins around year 5, under the assumption of no maintenance during the system's lifespan. Note that Figure 7(a) is based on the numerical values specific to this study; for a clearer understanding of the annual trend in transition probability distribution, Figure 7(b) calculates outcomes using hypothetical numbers (Table 3). Figure 7(b) offers a clearer view of the semi-Markov mechanism, maintaining the same trend as seen in Figure 7(a).

Figure 8(a) illustrates the cumulative failure probability resulting from integrating semi-Markov and BW-FORM for each CS, which will be applied in GA optimization in Section "Optimization and discussion of maintenance scheduling." Without maintenance, the total failure probability escalates, as demonstrated in Figure 8(a). However, performing maintenance Type I can reduce this probability significantly, effectively "rejuvenating" the system by a number of years depending on the maintenance conducted, as shown in Figure 4. Figure 8(b) presents the overall failure probability attributable to latent failures, considering all CSs, which increases over time. Yet, implementing maintenance Type II (e.g., inspection) can decrease this latent failure probability to a lower level, contingent upon the type of inspection conducted. The optimization of maintenance type and timing will be conducted using GA metaheuristic optimization in Section "Optimization and discussion of maintenance scheduling."

Optimization and discussion of maintenance scheduling

As previously mentioned, maintenance Type I targets weariness failure by altering or rejuvenating parts of the system, whereas Type II addresses latent failures. Consequently, optimization of maintenance scheduling

involves two distinct failure limit states, yet includes interactions between them. Type I maintenance is scheduled to prevent weariness failure probabilities in Figure 8(a) from exceeding specific thresholds. Similarly, Type II aims to keep failure probabilities in Figure 8(b) below predetermined limits. Effectiveness and cost are predefined for optimization. The goal, using GA, is to minimize the cost function outlined in Equation (19) (section “Methodology”), with maintenance costs and effectiveness detailed in Table 4. Effectiveness implies rejuvenation of the system to a “younger” state by a certain factor, based on assumptions in this study, though real-world data should ideally be experience-based. A lower effectiveness value indicates a younger state after maintenance. The optimization incorporates three constraints (Table 5): two relate to failure probability thresholds for both failure limit states, and the third ensures inspections follow any maintenance, as per common practice. The optimization adjusts M_{type} , I_{type} , r_M , r_I to minimize costs, assuming a 5% annual inflation rate for all cost components. GA optimization uses a standard approach with 200 chromosomes and 1000 iterations, employing a roulette wheel selection and single-point crossover for generating successive generations.

Figure 9 illustrates the results obtained from GA optimization. Figure 9(a) illustrates the cost function's convergence history. Figure 9(b) and (c) details the total failure probabilities for general and latent failure limit states, respectively. Maintenance reduces failure probabilities based on the type, as depicted in Figure 9(b) and (c). Figure 9 suggests the general limit state predominantly dictates system failure, frequently reaching the failure threshold. In this simulation, optimal maintenance is initiated around year 7 of operation, with a preference for a mix of maintenance Types 1-1, 1-2, and 1-3 for cost optimization.

Further analysis of Figure 9(b) reveals the recommended maintenance sequence as Types 1-1, 1-2, and 1-3, followed by a specific pattern, indicating Type 1-3, recognized as a larger scale maintenance operation according to Table 4's effectiveness, is scheduled for years 9.3, 11.8, 12.5, 13.2, 14.2, 16.2, 17.7, 19.5, 20.4, 23, 23.5, 25.7, 28.2, 30.7, 31.0, 33.5, 35.1, and 36.8 with intervals of 9.3, 2.5, 0.7, 0.7, 1, 2, 1.5, 1.8, 0.9, 2.6, 0.5, 2.2, 2.5, 2.5, 0.3, 2.5, 1.6, 1.7, and 3.2 years, respectively. Initially, extensive maintenance is not required, but the need for type 1-3 interventions grows over time. Although the numbers may not perfectly reflect this trend due to the effectiveness and cost parameters in Table 4 and the randomness in GA, this observation still underscores a significant conclusion.

Another critical observation is that while the general limit state predominates the constraints, by year 26, the system is primarily constrained by the latent limit

state. This observation highlights an important consideration: as expected, even with perfect execution of type I maintenance, the Tainter gate may still fail during the transition from dormancy to activation, necessitating special attention.

Comparatively, the outcomes from the GA were compared against scenarios employing singular maintenance strategies, specifically maintenance types 1-3, 2-3, 1-2, and 2-2. Analysis revealed that the implementation of maintenance types 1-3 and 2-3, triggered by thresholds of general or latent failure probabilities, incurred a cost function of 341.26 (as depicted in Figure 10). Conversely, the strategy involving maintenance types 1-2 and 2-2 resulted in a marginally higher cost function of 391.52, as illustrated in Figure 11. The cost disparity between these approaches was deemed negligible.

Three key points emerge from the analysis. First, despite the minimal cost variance between the two approaches, a significant difference in the frequency of maintenance activities was observed. Specifically, Type 3 maintenance was required 13 times over a 40-year period (Figure 10), whereas Type 2 maintenance was needed 21 times (Figure 11). This outcome aligns with expectations, given that Type 3 maintenance is more comprehensive and in-depth. A closer examination of the effectiveness and cost metrics in Table 4 reveals that, although Type 2 and Type 3 do not exhibit a direct proportional relationship in terms of numerical values, the optimized frequency of maintenance demonstrates a doubling effect (21 vs 13), indicating a nonlinear relationship in the system reliability assessment framework proposed in this study. Second, consistent with previous observations, the system's reliability is predominantly influenced by the general limit state. However, at certain junctures, the latent limit state becomes the governing factor in system reliability. Third, under a singular maintenance strategy, while the latent limit state does not dominate, the failure probability exhibits a monotonic increasing trend between two maintenance events. Although singular maintenance strategies are impractical in real-world applications, these findings reiterate the importance of paying closer attention to the latent limit state due to its potential to cause system failure and the increasing trend of failure probability over time.

The interaction between the two limit states is another focal point of this study, affecting maintenance management strategies. Figure 12 illustrates the optimization schedule when the risk ratio (n_l) is set to 4. As anticipated, it demonstrates that the latent limit state gathers more concern compared to the general limit state. A comparison between Figures 9 and 12 clearly reveals that, for maintenance Type II, not only does the frequency of implementation significantly increase

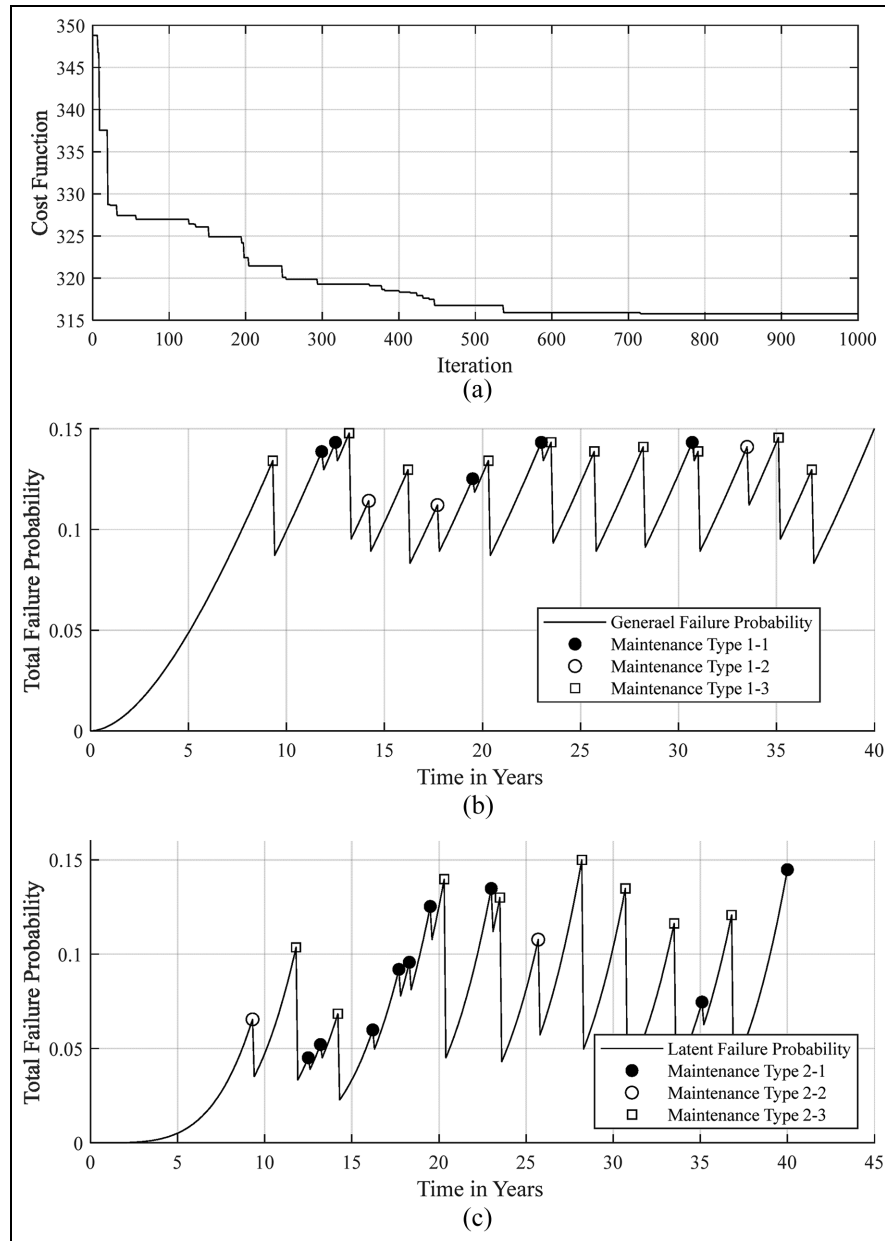


Figure 9. Summary of GA optimization results. (a) GA optimization convergence history (Best Cost Function = 316). (b) Results on maintenance Type I scheduling. (c) Results on maintenance Type II scheduling.
GA: genetic algorithm.

(from 19 to 39 times), but the corresponding failure probabilities also rise and are consistently close to the preset threshold values (0.15).

This study further inspects the impact of various n_l values on optimization schedule result. Figure 13 shows the optimization schedule when $n_l = 1, 2$, and 4. Similar to earlier analysis, a singular maintenance strategy using maintenance Type I-3 and maintenance Type II-3 is performed. Result of general limit state is same as the one shown in Figure 10(a). While variation

of latent limit state could be seen in Figure 13(a), it is shown that latent limit states become more concerning compared to general limit states as n_l value gets larger. This is due to the total latent probability increases as n_l increases, as shown in Figure 13(b). In Figure 13, maintenance Type I is performed 22 times, while maintenance Type II is performed 39 times due to riskier condition on latent limit state.

This study further examines the effects of various n_l values on the outcomes of optimization schedules.

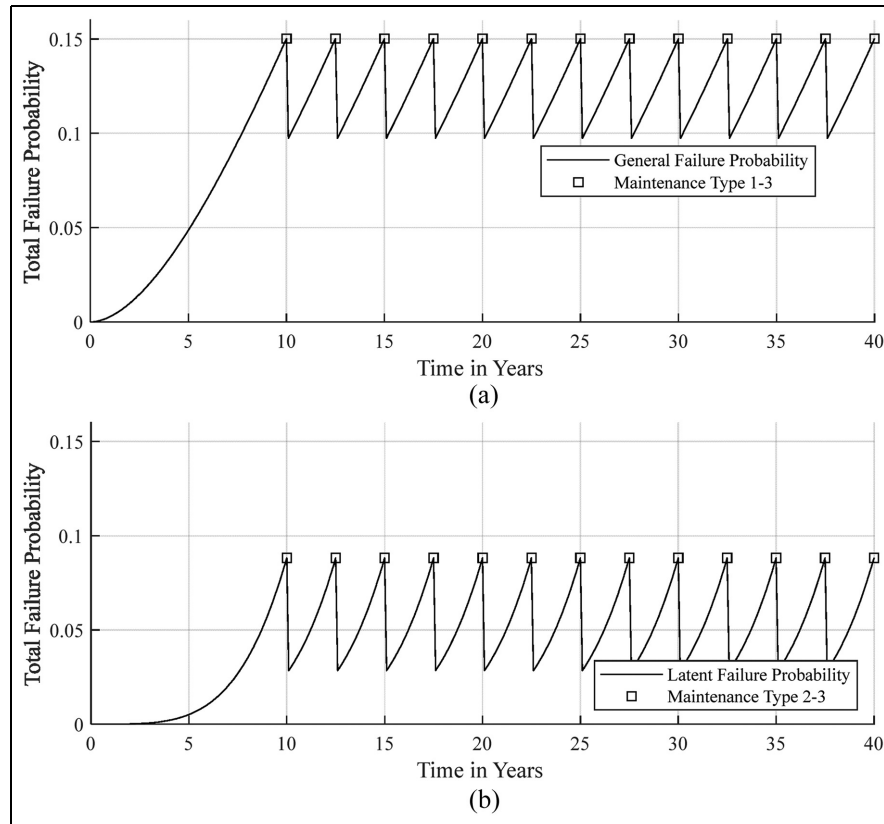


Figure 10. Type 3 maintenance and inspection strategy only (Cost Function = 341.26): (a) Type I-3 exclusive maintenance strategy. (b) Type 2-3 exclusive maintenance strategy.

Figure 13 illustrates the optimization schedule for $n_l = 1, 2$, and 4. Consistent with previous analyses, a single maintenance strategy utilizing maintenance Type I-3 and maintenance Type II-3 is implemented. The results for the general limit state remain identical to those presented in Figure 10(a). However, variations in the latent limit state are observable in Figure 13(a). Figure 13(a) demonstrates that as the value of n_l increases, the likelihood of the latent limit state being a dominant factor in system reliability also rises. When compared with Figure 11(b), it is evident that for $n_l = 4$, the failure probability of the latent limit state is generally higher and approaches the predefined threshold of 0.15. Figure 13(b) presents the total latent failure probability calculated using a semi-Markov process for various CS states. This figure clearly illustrates that an increase in n_l value corresponds to a higher probability of the latent limit state dominating the system's reliability. In Figure 13, maintenance Type I is conducted 22 times, while maintenance Type II is executed 39 times, reflecting the heightened risk associated with the latent limit state.

Conclusion

Considering the characteristics of Tainter gates, which include extended periods of dormancy, a requirement for high reliability upon activation, and safety during operation, this study devises a framework specifically designed to address these issues. For instance, this research introduces a semi-Markov process, wherein the duration time spent in any given state can follow a non-exponential distribution. Subsequently, the Weibull distribution is applied to depict the deterioration behavior of Tainter gates within each CS, and the semi-Markov process is employed to determine the transition probabilities between CSs. Additionally, optimization was conducted to determine the optimal maintenance strategy for the Tainter gate system, recognizing the potential for latent failure during extended dormancy periods and the critical requirement for the gate to remain in a state of readiness. This approach includes addressing both general and latent failure limit states, with maintenance serving as a mechanism to mitigate failures. By considering four

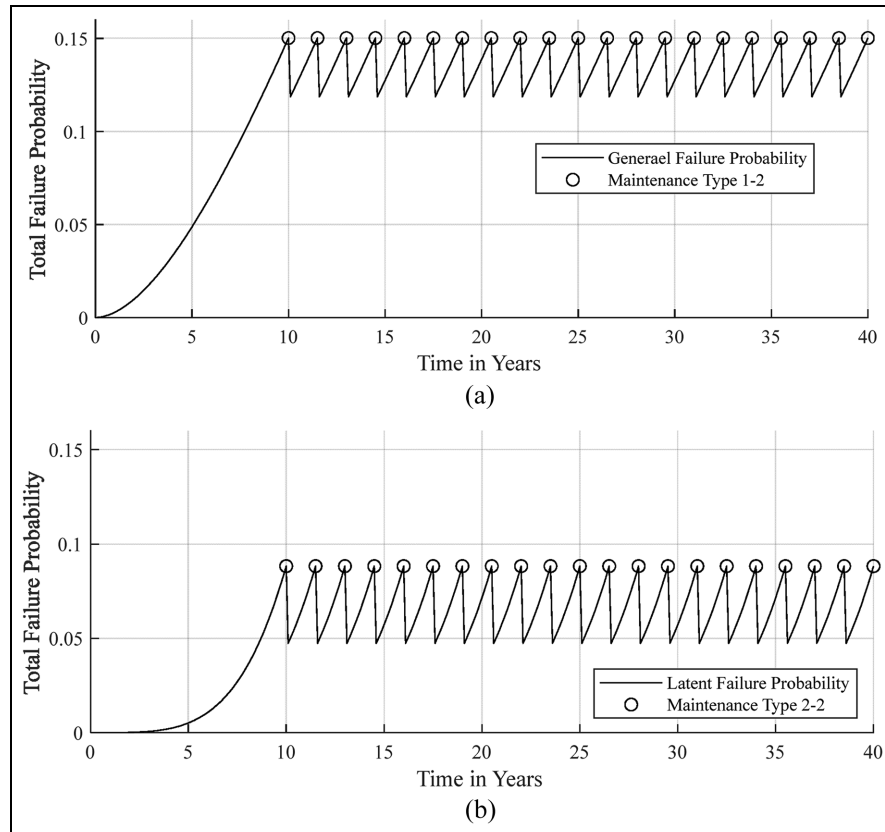


Figure 11. Type 2 maintenance and inspection strategy only (Cost Function = 391.52): (a) Type 1-2 exclusive maintenance strategy. (b) Type 2-2 exclusive maintenance strategy.

levels of CSs and combining failure probabilities for each limit state with transition probabilities via the semi-Markov process, a comprehensive failure probability is derived. GA optimization then identifies the optimal maintenance schedule to ensure system reliability thresholds are not exceeded. Key findings from this study include:

1. The general limit state is the primary determinant of system failure, though the latent limit state occasionally becomes critical to system reliability.
2. The requirement for comprehensive maintenance increases over time, despite initial minimal needs.
3. Even with flawless execution of Type I maintenance, Tainter gates risk failure during the shift from dormancy to activation, necessitating focused oversight.
4. The frequency of maintenance is not directly proportional to its effectiveness.
5. The latent limit state warrants special attention due to its failure potential and the growing probability of failure over time.
6. By analyzing the interaction between general and latent limit states, it is evident that inadequate routine maintenance can cause latent limit states to become the predominant factors controlling system reliability.

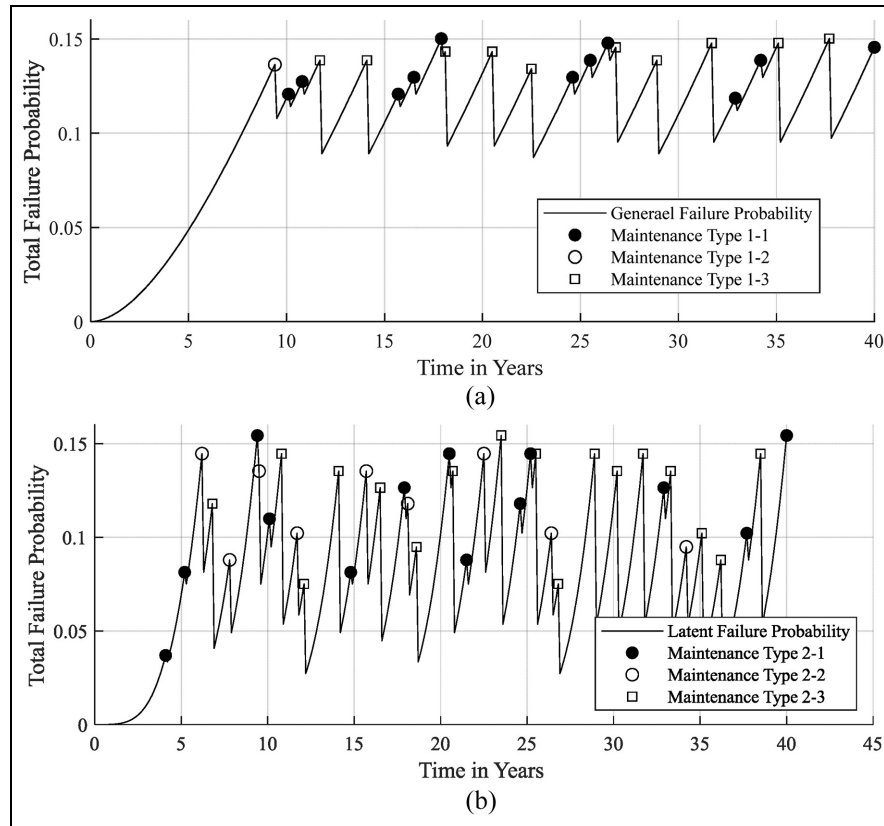


Figure 12. GA optimization results using $n_1 = 4$. (a) Results on maintenance Type I scheduling. (b) Results on maintenance Type II scheduling.

GA: genetic algorithm.

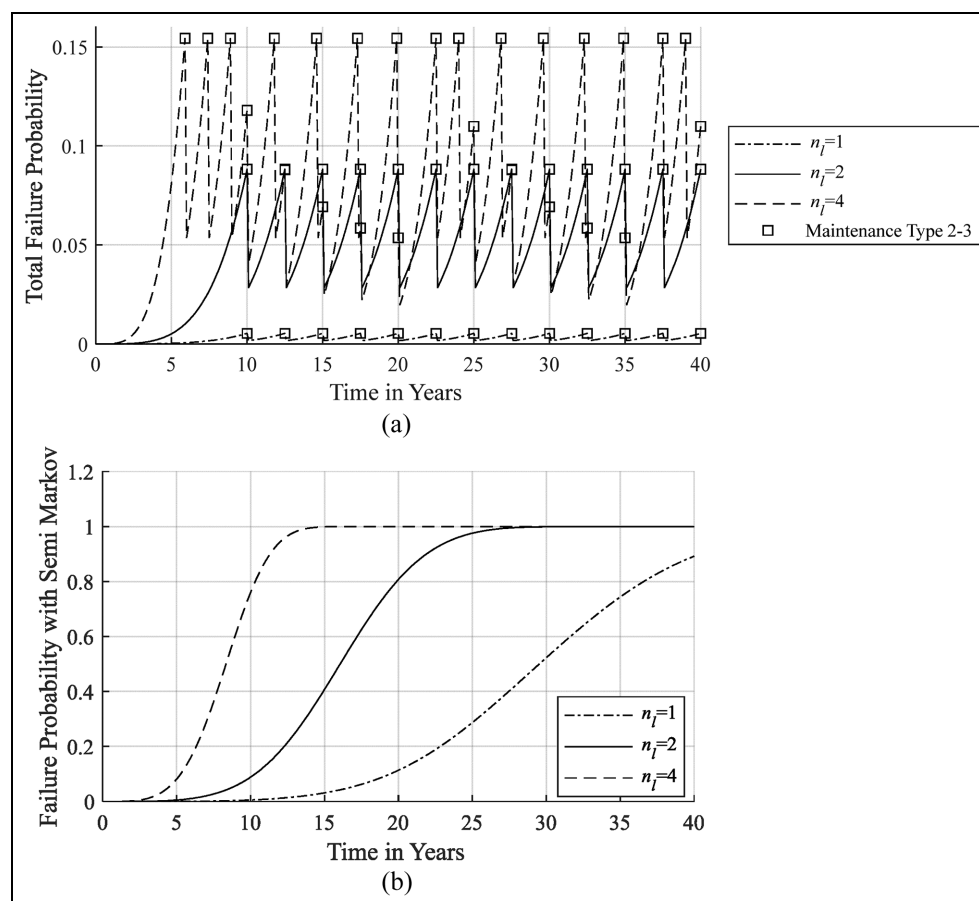


Figure 13. GA optimization results using various n_l : (a) results on maintenance Type 2-3 exclusive maintenance strategy on $n_l = 1$, 2, and 4 and (b) total latent failure probability across CSs using $n_l = 1$, 2, and 4.

GA: genetic algorithm; CS: condition score.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Southern Region Water resources office, Water Resources Agency, Ministry of Economic Affairs, Taiwan under grant number MOEA WRA1100407 and the National Science and Technology Council of Taiwan under grant number 112-2811-M-002-165. The support is gratefully acknowledged.

ORCID iD

Kuo-Wei Liao  <https://orcid.org/0000-0001-5274-291X>

References

1. Kirubakaran B and Ilankumaran M. Selection of optimum maintenance strategy based on FAHP integrated with GRA-TOPSIS. *Ann Oper Res* 2016; 245: 285–313.
2. Stenström C, Norrbin P, Parida A, et al. Preventive and corrective maintenance—cost comparison and cost-benefit analysis. *Struct Infrastruct Eng* 2016; 12: 603–617.
3. Ahmad R and Kamaruddin S. An overview of time-based and condition-based maintenance in industrial application. *Comput Ind Eng* 2012; 63: 135–149.
4. Jardine AKS, Lin D and Banjevic D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech Syst Signal Process* 2006; 20: 1483–1510.
5. Selcuk S. Predictive maintenance, its implementation and latest trends. *Proc Inst Mech Eng Part B J Eng Manuf* 2016; 231: 1670–1679.

6. Zhao H, Xu F, Liang B, et al. A condition-based opportunistic maintenance strategy for multi-component system. *Struct Health Monit* 2018; 18: 270–283.
7. Vu HC, Do P and Barros A. A stationary grouping maintenance strategy using mean residual life and the birnbaum importance measure for complex structures. *IEEE Trans Reliab* 2016; 65: 217–234.
8. Zhou Y, Lin TR, Sun Y, et al. An effective approach to reducing strategy space for maintenance optimisation of multistate series–parallel systems. *Reliab Eng Syst Saf* 2015; 138: 40–53.
9. Coolen-Schrijner P, Shaw SC and Coolen FPA. Opportunity-based age replacement with a one-cycle criterion. *J Oper Res Soc* 2009; 60: 1428–1438.
10. Wang J, Makis V and Zhao X. Optimal condition-based and age-based opportunistic maintenance policy for a two-unit series system. *Comput Ind Eng* 2019; 134: 1–10.
11. Tan Z, Li J, Wu Z, et al. An evaluation of maintenance strategy using risk based inspection. *Saf Sci* 2011; 49: 852–860.
12. Bevilacqua M and Braglia M. The analytic hierarchy process applied to maintenance strategy selection. *Reliab Eng Syst Saf* 2000; 70: 71–83.
13. Lipi TF, Lim JH, Zuo MJ, et al. A condition- and age-based replacement model using delay time modelling. *Proc Inst Mech Eng O J Risk Reliab* 2011; 226: 221–233.
14. Berrade MD, Scarf PA, Cavalcante CAV, et al. Imperfect inspection and replacement of a system with a defective state: a cost and reliability analysis. *Reliab Eng Syst Saf* 2013; 120: 80–87.
15. Vaurio JK. Optimization of test and maintenance intervals based on risk and cost. *Reliab Eng Syst Saf* 1995; 49: 23–36.
16. Vaurio JK. On time-dependent availability and maintenance optimization of standby units under various maintenance policies. *Reliab Eng Syst Saf* 1997; 56: 79–89.
17. Ahmadi A and Kumar U. Cost based risk analysis to identify inspection and restoration intervals of hidden failures subject to aging. *IEEE Trans Reliab* 2011; 60: 197–209.
18. Nguyen TPK, Yeung TG and Castanier B. Optimal maintenance and replacement decisions under technological change with consideration of spare parts inventories. *Int J Prod Econ* 2013; 143: 472–477.
19. Hu Z and Du X. Lifetime cost optimization with time-dependent reliability. *Eng Optimiz* 2014; 46(10): 1389–1410.
20. USACE (U.S. Army Corps of Engineers). *Reliability analysis of navigation lock and dam mechanical and electrical equipment*. Washington, DC: U.S. Army Corps of Engineers, 2001.
21. Khatib A. Hybrid hazard rate model for imperfect preventive maintenance of systems subject to random deterioration. *J Intell Manuf* 2015; 26: 601–608.
22. Samrout M, Châtelet E, Kouta R, et al. Optimization of maintenance policy using the proportional hazard model. *Reliab Eng Syst Saf* 2009; 94: 44–52.
23. Sobanjo JO. State transition probabilities in bridge deterioration based on Weibull sojourn times. *Struct Infrastruct Eng* 2011; 7: 747–764.
24. Fang Y and Sun L. Developing a Semi-Markov process model for bridge deterioration prediction in Shanghai. *Sustainability* 2019; 11: 5524.
25. Naga Srinivasa Rao P and Achutha Naikan VN. An algorithm for simultaneous optimization of parameters of condition-based preventive maintenance. *Struct Health Monit* 2008; 8: 83–94.
26. Le MD and Tan CM. Optimal maintenance strategy of deteriorating system under imperfect maintenance and inspection using mixed inspection scheduling. *Reliab Eng Syst Saf* 2013; 113: 21–29.
27. Kalantarnia M, Chouinard L and Foltz S. Objective procedure for optimization of inspection and testing strategies for spillways. *J Infrastruct Syst* 2016; 22: 04015011.
28. Belyi D, Popova E, Morton DP, et al. Bayesian failure-rate modeling and preventive maintenance optimization. *Eur J Oper Res* 2017; 262: 1085–1093.
29. Ching J and Leu S-S. Bayesian updating of reliability of civil infrastructure facilities based on condition-state data and fault-tree model. *Reliab Eng Syst Saf* 2009; 94: 1962–1974.
30. Vega MA, Hu Z and Todd MD. Optimal maintenance decisions for deteriorating quoin blocks in miter gates subject to uncertainty in the condition rating protocol. *Reliab Eng Syst Saf* 2020; 204: 107147.
31. Riveros GA and Rosario-Pérez ME. Deriving the transition probability matrix using computational mechanics. *Eng Comput* 2018; 35(2): 692–709.
32. Li YF, HuAng HZ, Liu Y, et al. A new fAult tree AnAl-ysis method: fuzzy dynAmic fAult tree AnAlizA metoD AnAlizy drzewA uszkodzeń: rozmytA AnAlizA dynAmicznego drzewA uszkodzeń. *Eksploat Niezawodn* 2012; 14(3): 208–214.
33. Jeong S, Kwon HH and Lee SO. Estimates of sediment pickup rate induced by surge wave within a multilevel Bayesian regression framework. *J Coast Res* 2018; 85: 286–290.
34. Leu SS, Fu YL and Wu PL. Dynamic civil facility degradation prediction for rare defects under imperfect maintenance. *J Qual Mainten Eng* 2024; 30(1): 81–100.
35. Am H and Lind N. An exact and invariant first order reliability format. *J Eng Mech*. 1974; 100: 111–121.
36. Tian H, Li DQ, Cao ZJ, et al. Auxiliary Bayesian updating of embankment settlement based on finite element model and response surface method. *Eng Geol* 2023; 323: 107244.
37. U.S. Army Corps of Engineers. Engineering and design: introduction to probability and reliability methods for use in geotechnical engineering. Engineer Technical Letter 1110-2-547, Department of the Army, Washington, DC, 1997.
38. Binda L and Molina C. Building materials durability: Semi-Markov approach. *J Mater Civ Eng* 1990; 2: 22–239.

39. Hastings NAJ. Dynamic probabilistic systems: Vol. 1: Markov models; Vol. II: Semi-Markov and decision processes. *J Oper Res Soc* 1973; 24: 327–329.
40. Black M, Brint AT and Brailsford JR. A Semi-Markov approach for modelling asset deterioration. *J Oper Res Soc* 2005; 56: 1241–1249.
41. Perman M, Senegacnik A and Tuma M. Semi-Markov models with an application to power-plant reliability analysis. *IEEE Trans Reliab* 1997; 46: 526–532.
42. Holland JH. Genetic algorithms. *Sci Am* 1992; 267: 66–73.