

Fatigue life prediction and periodic operation and maintenance optimization of high-strength steel welded joints considering residual stress

Chaojing Jiang ^{*,a}

Department of Mechanical and Information Engineering, Sichuan University of Architectural Technology, Deyang, China

Article Info

Article History:

Received 16 Jul 2026

Accepted 05 Sep 2026

Keywords:

HSS;
Welded joint;
Residual stress;
Fatigue life prediction;
Stacked ensemble learning;
Periodic O&M

Abstract

Existing fatigue assessment methods for high-strength steel welded joints commonly provide isolated life estimates and cannot directly convert residual-stress effects and prediction uncertainty into operation and maintenance decisions. To address this limitation, an integrated RS-SN-Stack-RDF-MDP framework is developed. A total of 864 experimental fatigue-life records are divided by original test series into training, validation, and held-out test sets, while 196 records from non-overlapping publications are used for external validation. Residual-stress-corrected S-N life and crack-growth descriptors are combined with out-of-fold predictions from SVR, RF, XGBoost, LightGBM, and MLP, and fused by a physics-constrained Ridge meta-learner. Predicted life is converted into remaining life, damage, and next-period failure probability for five-state average-cost maintenance optimization. RS-SN-Stack achieves an RMSE of 0.058, MAE of 0.044, MAPE of 3.1%, and R² of 0.947 on the internal test set, while the corresponding external results are 0.084, 0.064, 4.4%, and 0.879. RDF-MDP opportunistic maintenance reduces the long-term average cost from 168.4×10³ to 126.5×10³, failure penalty cost from 39.4×10³ to 20.8×10³, and average downtime from 19.2 to 11.8 h. The framework establishes a continuous computational link from residual-stress-corrected fatigue prediction to maintenance action selection. Its validated applicability is currently limited to the investigated high-strength steel grades, joint geometries, and loading ranges.

© 2026 MIM Research Group. All rights reserved.

1. Introduction

High-strength steel (HSS) welded structures are widely used in bridges, rail transit, construction machinery, marine equipment, and large load-bearing components. Safety in service of such structures is mainly dependent upon the fatigue performance of welds. Welded joints typically possess discontinuous geometries, heterogeneous microstructures, welding flaws, and higher levels of residual stress compared to those of their base metals. For HSS welded joints, tensile residual stress, stress concentration, and defects act in combination to negate the inherent strength benefits of HSS; thus, an increase in material strength does not always yield a corresponding rise in joint fatigue strength [1]. Conventional fatigue assessment relies mainly on nominal-stress, hot-spot-stress, structural-stress, S-N-curve, and fracture-mechanics methods [2]. These methods provide fatigue-life or damage estimates but generally do not produce the remaining-life distribution, next-period failure probability, or transition probabilities required for maintenance-state optimization. While conventional approaches lack this capability, machine learning can recognize intricate nonlinear relations among material properties, geometry, loading conditions, residual stresses, and fatigue life information from experiments and can enhance its accuracy and

^{*}Corresponding author: jiangchaojing2026@163.com

^aorcid.org/0009-0003-0567-4143

DOI: <https://dx.doi.org/10.17515/resm2026-1892am0716rs>

Res. Eng. Struct. Mat. Vol. x Iss. x (xxxx) xx-xx

generalization ability using ensemble learning techniques, support vector regression, trees, and neural networks with physics constraints [3-4]. Combining residual-stress-corrected master S-N curves and crack propagation mechanisms with machine learning models retains the physical interpretability of fatigue mechanisms while enhancing life prediction capability under complex loading conditions.

Fatigue life prediction, however, is not the end goal in engineering applications. For HSS welded structures in actual service, operators are more concerned with when joints should be inspected, when repairs are needed, whether maintenance can be synchronized with adjacent components, and how to balance safety risk, maintenance cost, and downtime losses [5-6]. Enhances the practicality of HSS welded joint fatigue life evaluation in engineering Operation and Maintenance (O&M) practice, current research mainly focuses on fatigue life prediction for HSS welded joints using machine learning-based methodologies, incorporating residual stress while utilizing remaining life, damage indices, and failure probability as state variables in O&M optimization practice.

2. Related Studies

The foundation of fatigue life assessment for HSS welded joints still rests on welding fatigue theory, local stress methods, and residual stress correction approaches. Ono et al. studied local residual stress relaxation in high-frequency mechanical impact (HFMI)-treated HSS welded joints under high peak loads and showed that surface morphology, material properties, and residual stress jointly govern the distribution of fatigue damage in the joints [7]. Ahola et al. further incorporated material strength, stress ratio, residual stress, and weld toe radius into the four-parameter fatigue assessment method (4R method) for evaluating both as-welded and HFMI-treated HSS joints under variable amplitude loading, demonstrating that residual stress serves as an important parameter for mean stress correction and local stress characterization [8]. According to Song et al., the presence of welding defects as well as geometric variations influences the local values of stress concentration, scatter in fatigue life, and failure probability, suggesting that models for life prediction must also account for these parameters [9].

Recent studies have moved from isolated S-N fitting toward data-physics hybrid fatigue prediction. Ghanadi et al. used K-means and DBSCAN to identify local stress peaks automatically, but their method did not directly predict fatigue life or maintenance states [10]. Schubnell et al. transferred knowledge from non-welded steel specimens to welded joints, thereby reducing dependence on expensive welded-joint tests; however, residual stress and initial defects were not included among the available features [3]. Wang et al. demonstrated that neglecting welding residual stress can substantially overestimate high-cycle fatigue life, although their damage-coupled thermal-mechanical calculation remains computationally intensive [11]. Kim et al. embedded fracture-mechanics features into a physics-informed Gaussian process regression model [12], while Li et al. developed a modified physics-informed neural network for deck-rib double-sided welded joints [13]. These hybrid models improve prediction accuracy or physical consistency, but they still treat fatigue-life prediction as the terminal output and do not convert it into remaining life, cumulative damage, and next-cycle failure probability states for maintenance optimization. From an O&M perspective, fatigue predictions must be converted into decision-oriented health states before they can support periodic or opportunistic maintenance. Karakaya et al. integrated sensor-based condition prediction with maintenance routing, but their health indicators were derived from monitoring data rather than residual-stress-corrected weld fatigue life [14]. Xu et al. formulated group maintenance as a factored MDP, although the application concerned generic multicomponent degradation and did not define weld-specific life, damage, and failure states [15]. Chai et al. combined a Gamma degradation process with periodic inspection and component reassignment [16], while Eggertsson et al. considered incomplete and environment-dependent maintenance information [17]. These studies provide effective decision mechanisms, but they generally assume that degradation states are already available rather than constructing them from a physically corrected fatigue model.

Taken together, three gaps remain. First, existing fatigue studies rarely combine residual-stress relaxation, physical baseline life, and nonlinear ensemble correction within one prediction model. Second, predicted life is seldom converted into remaining life, damage, and next-cycle failure probability with quantified uncertainty.

Third, maintenance models usually assume state inputs instead of constructing them from fatigue-mechanism variables. This study addresses these gaps by linking residual-stress-corrected life prediction with RDF-MDP state construction and periodic maintenance optimization. Its main contribution is therefore the traceable conversion from fatigue-mechanism inputs to prediction outputs and subsequently to maintenance actions, rather than the isolated use of either stacked ensemble learning or an MDP.

3. Corrected Life Prediction and State-Driven Periodic Modeling Method

3.1 Fatigue Life Prediction Model Considering Residual Stress

The fatigue failure of HSS welded joints typically results from a combination of factors including the weld toe, defects, and the existence of residual tensile stress areas. This variability in fatigue lives impacts the determination of inspection cycles and maintenance periods. Accordingly, a fatigue life prediction model incorporating residual stress is first developed. The resulting RS-SN-Stack model uses a residual-stress-corrected S-N curve to generate the physical baseline life and then employs stacking learning to correct nonlinear prediction bias.

To avoid disconnecting data-driven prediction from fatigue mechanisms, the model comprises four sequential steps: residual-stress evolution, equivalent structural-stress correction, physical baseline-life calculation, and stacking-based fatigue-life prediction. Under cyclic loading, welding residual stress gradually relaxes or redistributes and must therefore be updated before fatigue-life correction [18]. The exponential form is obtained by assuming that the residual-stress relaxation rate is proportional to the current residual stress, $z_i/dn = -\lambda_i z_i$. Integration gives $z_i/dn = -\lambda_i z_i$, representing the experimentally observed rapid initial relaxation followed by gradual stabilization [7,18]. The relaxation relationship at the critical location is given by Equation (1).

$$z_i(n) = z_{0,i} \exp(-\lambda_i n_i) \quad (1)$$

In Equation (1), $z_i(n)$ is the residual stress at the critical weld-toe position after n_i loading blocks, and $z_{0,i}$ is the corresponding initial residual stress. The coefficient λ_i was fitted by nonlinear least squares using residual-stress measurements before loading, after the first peak load, and after the first complete loading block for each joint and loading group. For records without repeated measurements, λ_i was sampled from the lognormal distribution of the corresponding material-joint-loading group. Its uncertainty was subsequently propagated through the Monte Carlo analysis in Section 4.5. Equivalent mean stress and structural stress must be computed from residual stress before calculating fatigue damage. This correction accounts for the influence of residual tensile stress on the stress ratio, structural stress amplitude, and fatigue life [19-20]. The expressions are shown in Equation (2).

$$\begin{cases} M_i = m_i + \eta_i z_i(n) \\ R_i = \frac{M_i - S_i}{M_i + S_i} \\ E_i = S_i(1 - R_i)^{-\gamma} \end{cases} \quad (2)$$

In Equation (2), M_i represents the equivalent mean stress considering residual stress. m_i represents the mean stress induced by external load. η_i represents the effective coefficient of residual stress. R_i represents the equivalent stress ratio. S_i represents the structural stress amplitude. E_i represents the equivalent structural stress after residual stress correction. γ represents the stress ratio correction factor. The stress relaxation and equivalent structural stress correction mechanism describe the transfer path of residual stress from the initial welding state to life calculation variables, as shown in Figure 1.

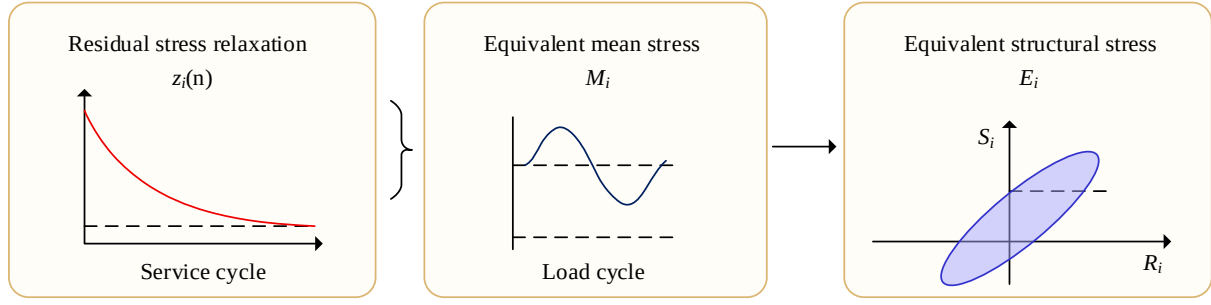


Fig. 1. Residual stress relaxation and equivalent structural stress correction mechanism

As illustrated in Figure 1 the attenuation curve is a depiction of the relaxation of residual stresses when subjected to cyclical loads, further impacting the calculation of corrected mean stresses, stress ratio, and structural equivalent stress. This corrects the values to account for the effect of residual stress on the fatigue life of components. This study uses the residual-stress-corrected main S-N curve to calculate the physical baseline life, as shown in Equation (3).

$$\log N_{p,i} = A - B \log E_i \quad (3)$$

In Equation (3), $N_{p,i}$ represents the predicted life of the i -th welded joint obtained from the physical baseline model. A and B represent the main S-N curve parameters. For joints with high sensitivity to initial defects, crack growth parameters are also included as auxiliary features in defect characterization. This study constructs the machine learning input matrix and establishes the life prediction mapping as shown in Equation (4).

$$\begin{cases} N_{prop,i} = \int_{a_{0,i}}^{a_{c,i}} \frac{da}{C_i (\Delta K_{eff,i})^{m_i}} \\ g_i = [a_{0,i}, a_{c,i}, \Delta K_{th,i}, C_i, m_i, N_{prop,i}] \\ X_i = [t_i, w_i, S_i, M_i, R_i, E_i, z_i, g_i, N_{p,i}] \\ y_i = \log N_i, \hat{y}_i = f_\theta(X_i) \end{cases} \quad (4)$$

In Equation (4), fatigue crack initiation is represented by the weld-toe geometry w_i , initial defect size $a_{0,i}$, and stress concentration embedded in E_i . Crack propagation is represented by the critical crack size $a_{c,i}$, propagation threshold $\Delta K_{th,i}$, Paris coefficients C_i and m_i , and the propagation-life descriptor $N_{prop,i}$. The measured target N_i is the total life to failure and therefore contains both crack-initiation and crack-propagation stages. The mapping in Equation (4) is implemented using five heterogeneous base learners and a Ridge regression meta-learner. The standardized feature vector is supplied to Support Vector Regression (SVR), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Multilayer Perceptron (MLP) base learners. Five-fold grouped training produces the out-of-fold prediction vector $\hat{y}_i = [\hat{y}_i^{(1)}, \hat{y}_i^{(2)}, \hat{y}_i^{(3)}, \hat{y}_i^{(4)}, \hat{y}_i^{(5)}]^T$ for each training record. Ridge regression then fuses these predictions to obtain the logarithmic fatigue-life estimate. Remaining life, damage, and failure probability are not directly predicted by the meta-learner but are calculated from the fused life estimate using Equation (6). Ridge regression is selected because its L2 penalty suppresses collinearity among the five OOF prediction columns while retaining all base learners, which is preferable to using a high-capacity nonlinear meta-model for the present small dataset. To control both prediction error and deviation from the residual-stress-corrected physical baseline, the meta-learner is trained using Equation (5).

$$L(\beta) = \frac{1}{n} \sum_{i=1}^n (y_i - \beta_0 - \mathbf{z}_i^T \beta)^2 + \rho \frac{1}{n} \sum_{i=1}^n (\beta_0 + \mathbf{z}_i^T \beta - \log N_{p,i})^2 + \alpha \|\beta\|_2^2 \quad (5)$$

In Equation (5), $L(\beta)$ represents the training loss. The first term is the data-fitting error, the second term penalizes deviation from the residual-stress-corrected physical life $N_{p,i}$, and the third term is the ridge penalty. The parameters ρ and α control the physical constraint and coefficient

regularization, respectively. Validation-set optimization yielded $\rho = 0.20$ and $\alpha = 0.10$. The fused fatigue-life estimate is then converted into the state variables required for O&M, as shown in Equation (6).

$$\begin{cases} \hat{N}_i = 10^{\hat{y}_i} \\ U_i = \hat{N}_i - n_i \\ D_i = \frac{n_i}{\hat{N}_i} \\ P_i = \Phi\left(\frac{n_i + \tau - \hat{N}_i}{q_i}\right) \end{cases} \quad (6)$$

In Equation (6), $\hat{N}_i$ represents the predicted fatigue life. U_i represents the remaining life. D_i represents the damage index. P_i represents the failure probability within the next cycle. τ represents the cycle increment corresponding to one maintenance period. q_i represents the uncertainty scale of prediction. Φ represents the standard normal distribution function.

3.2 Periodic O&M Optimization Model with Life Prediction Output as State Variable

The complete prediction-maintenance workflow begins with data preprocessing and grouped data splitting, followed by RS-SN-Stack training and fatigue-life prediction. As summarized in Figure 2, the prediction outputs are subsequently converted into remaining-life, cumulative-damage, and failure-probability states, which are supplied to the RDF-MDP model for transition-probability estimation and maintenance-policy optimization.

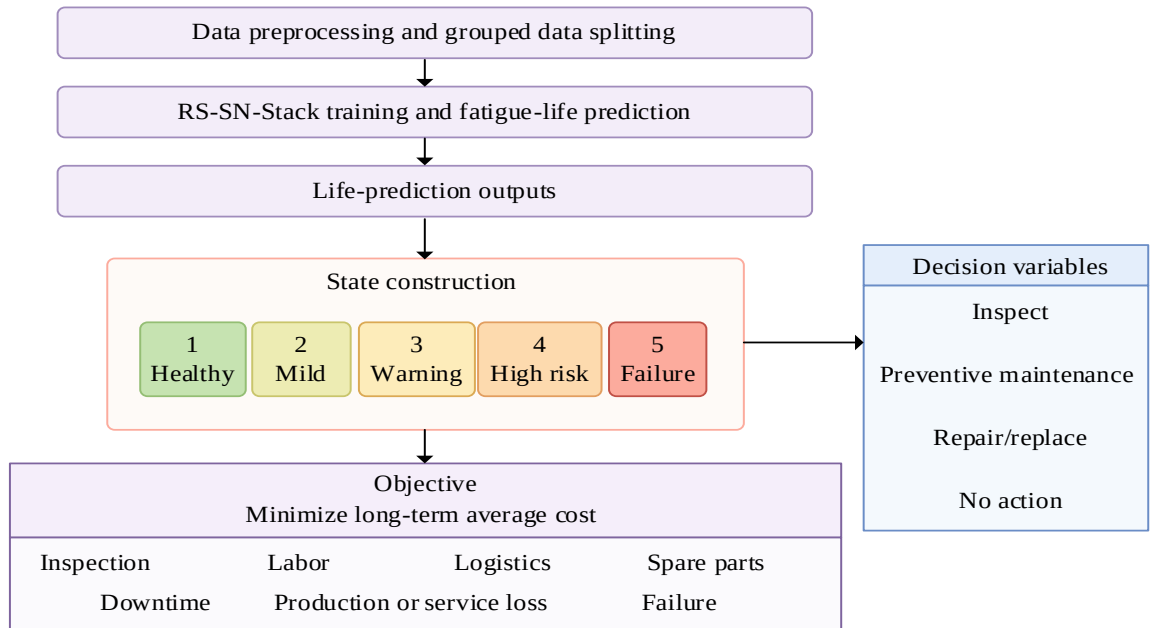


Fig. 2. RDF-MDP periodic O&M model

Figure 2 shows how the outputs of RS-SN-Stack are converted into discrete health states and subsequently used by RDF-MDP to determine maintenance actions under the long-term average-cost objective. This model uses remaining life, cumulative damage, and failure probability to represent the health states of welded joints and determines the optimal maintenance strategy by accounting for inspection, labor, logistics, spare parts, downtime, production or service loss, and failure costs. Compared with fixed-interval maintenance, this model adjusts inspection and repair timing according to the real degradation state of welded joints. It also triggers opportunistic maintenance when multiple joints fall within similar risk levels. The RDF-MDP assumes that the health state is observable at each inspection; the next state depends only on the current state and maintenance action; transition probabilities remain stationary within the analyzed horizon; maintenance effects are realized before the next inspection; and joints degrade conditionally

independent of one another, except for shared logistics and downtime during opportunistic maintenance. The state and action spaces are finite. These assumptions permit estimation of a stationary average-cost policy but may be violated when inspection error, environmental drift, or strong load redistribution exists. This study defines the health states of welded joints based on the damage index and failure probability, as shown in Equation (7).

$$x_i = \max\{h_U(\frac{U_i}{\tau}), h_D(D_i), h_P(P_i)\}, \quad x_i \in \{1,2,3,4,5\} \quad (7)$$

In Equation (7), x_i represents the health state of the i -th welded joint. The five states were defined as follows: healthy, $U_i/\tau > 8$, $D_i < 0.20$, and $P_i < 0.02$; mild, $5 < U_i/\tau \leq 8$, $0.20 \leq D_i < 0.40$, or $0.02 \leq P_i < 0.05$; warning, $3 < U_i/\tau \leq 5$, $0.40 \leq D_i < 0.65$, or $0.05 \leq P_i < 0.10$; high risk, $0 < U_i/\tau \leq 3$, $0.65 \leq D_i < 1$, or $0.10 \leq P_i < 0.25$; and failure, $U_i \leq 0$, $D_i \geq 1$, or $P_i \geq 0.25$. The maximum-risk mapping prevents a low value of one indicator from concealing a critical value of another. Under periodic inspection conditions, the state of welded joints transitions with service cycles and maintenance actions. The transition probability is shown in Equation (8).

$$\begin{cases} \pi_{ab}^c = P(x_{i,t+1} = b \mid x_{i,t} = a, c_t = c) \\ \hat{\pi}_{ab}^c = \frac{N_{ab}^c + 1}{\sum_{b'} N_{ab'}^c + 5} \end{cases} \quad (8)$$

In Equation (8), π_{ab}^c represents the probability that the welded joint transitions from state a to state b under maintenance action c . $x_{i,t}$ represents the state of the i -th welded joint in the t -th maintenance cycle. c_t represents the current maintenance action. N_{ab}^c is the number of transitions from state a to state b under action c . Transition counts were obtained from 10,000 degradation trajectories generated using Equation (6), the empirical out-of-fold residual distribution, and the action-dependent recovery rules. Laplace smoothing with a pseudo-count of 1 was used to avoid zero-probability transitions. Each row of the resulting transition matrix was normalized to one. To describe the economic effects of different maintenance actions, a single-cycle O&M cost function is defined, as shown in Equation (9).

$$C_t = c_i I_t + c_l L_t + c_g G_t + c_b B_t + c_d T_t^{\text{down}} + c_p Q_t^{\text{loss}} + c_f F_t \quad (9)$$

In Equation (9), C_t represents the total cost in the t -th cycle. I_t represents the inspection action variable. L_t is the required labor time. G_t is the number of logistics mobilizations. B_t represents spare-part consumption, T_t^{down} is downtime, Q_t^{loss} denotes production or service loss, and F_t is the failure event. All costs were converted to 10^3 monetary units using the same price basis. For the numerical case study, normalized scenario coefficients were adopted on a common price basis: 2.4 per inspection, 0.12 per labor-hour, 3.6 per logistics mobilization, 6.5 for preventive spare parts, 18.0 for replacement parts, 1.2 per downtime hour, 0.8 per service-loss hour, and 28.0 per failure event. Finally, the RDF-MDP periodic O&M model is formulated with the objective of minimizing long-term average cost, as shown in Equation (10).

$$J(\pi) = \lim_{T \rightarrow \infty} \frac{1}{T} E_\pi \left[\sum_{t=0}^{T-1} C_t \right] \quad (10)$$

In Equation (10), $J(\pi)$ represents the long-term average cost under policy π . E_π represents the expectation operator. To determine the cost-optimal maintenance action under uncertain state transitions, Figure 3 presents the solution structure of the Markov decision process. Figure 3 shows how the Markov decision process technique can be used for finding an optimal maintenance strategy. System state vector involves a number of different states of healthiness of welded joints. The discrete states are shown using circles, while the actions include inspection, preventive maintenance, and repair activities. The cost-minimal solution for a stochastic environment can be found via Bellman equations. In solving the optimal maintenance strategy, the Markov decision process is used to describe the recursive relationship among state, action, and cost, as shown in Equation (11).

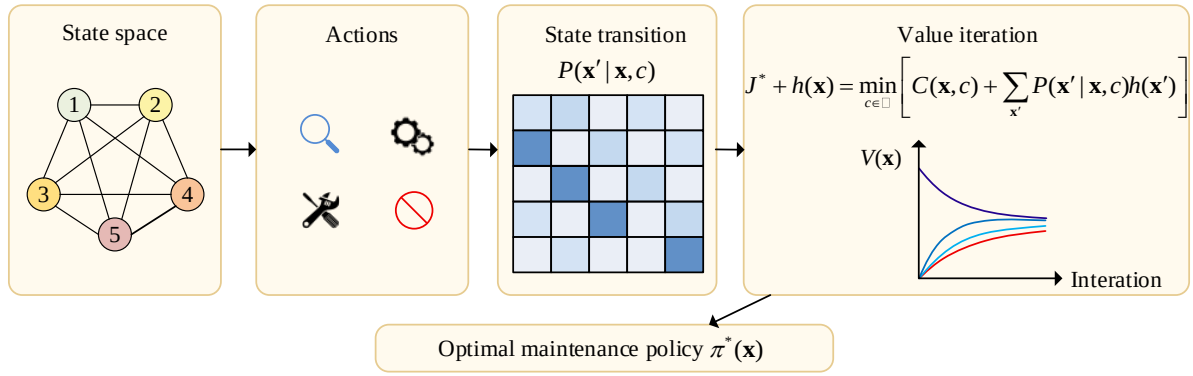


Fig. 3. Markov decision process for optimal maintenance strategy

$$J^* + h(x) = \min_{c \in \mathcal{A}} \left[C(x, c) + \sum_{x'} P(x'|x, c) h(x') \right] \quad (11)$$

In Equation (11), J^* is the optimal long-term average cost, $h(x)$ is the relative value function, and $\pi^*(x)$ is obtained from the minimizing action. x represents the system state vector composed of multiple welded joint health states. $\mathcal{A}$ represents the set of available actions. $C(x, c)$ represents the immediate cost. $P(x'|x, c)$ represents the transition probability from the current state to the next state x' . In multi-welded-joint structures, the opportunistic maintenance trigger rule can be further expressed as a state threshold constraint, as shown in Equation (12).

$$c_i = \begin{cases} 0, & D_i < q_1 \\ 1, & q_1 \leq D_i < q_2 \\ 2, & q_2 \leq D_i < q_3 \\ 3, & D_i \geq q_3 \text{ or } P_i \geq r \end{cases} \quad (12)$$

In Equation (12), c_i represents the maintenance action level of the i -th welded joint: 0 means no maintenance, 1 means routine inspection, 2 means preventive maintenance, and 3 means forced repair or replacement. r represents the failure probability threshold. If one joint triggers forced maintenance and adjacent joints satisfy $q_1 \leq D_i < q_3$, opportunistic maintenance can be executed within the same downtime window, reducing additional cost caused by repeated shutdowns and disassembly. When multiple welded joints have overlapping maintenance windows, the coupling logic of opportunistic maintenance triggering and cost is shown in Figure 4.

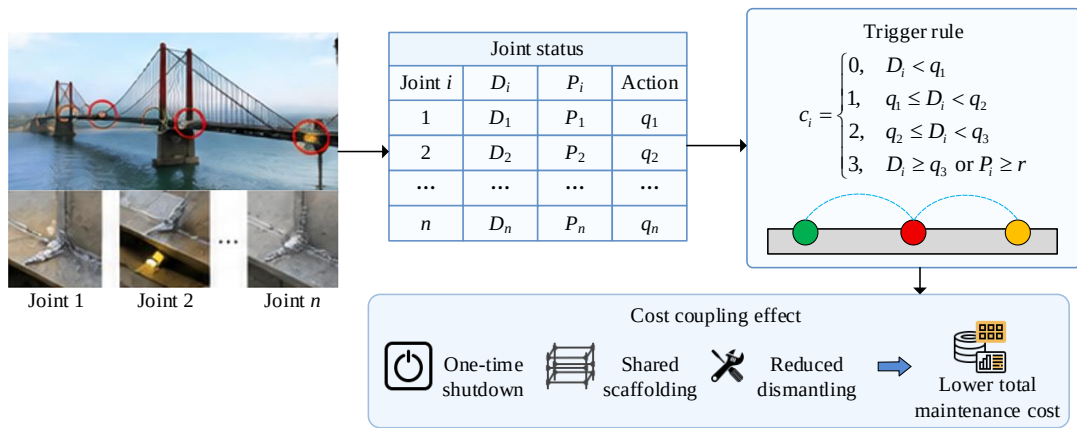


Fig. 4. Coupling mechanism of opportunistic maintenance triggering and cost

In Figure 4, the health states of each joint are monitored, and thresholds define neighboring joints, low-risk joints, and triggering joints. The model provides forced maintenance for critical joints, along with opportunistic maintenance for adjacent joints in the warning state in the same shutdown interval. The cost-comparison module evaluates whether coordinating eligible repairs within the same shutdown window reduces repeated mobilization and downtime costs without exceeding the prescribed risk threshold. For a representative joint with an initial residual stress of

325 MPa, $\lambda=0.06$, and ten loading blocks, Equation (1) yielded a relaxed residual stress of 178.3 MPa. Equations (2)-(5) subsequently produced an equivalent structural stress of 181.2 MPa, a physical baseline life of 1.06×10^6 cycles, and a fused fatigue-life estimate of 1.20×10^6 cycles. With 7.20×10^5 accumulated cycles, a maintenance interval of 5.0×10^4 cycles, and a prediction uncertainty scale of 2.94×10^5 cycles, Equation (6) yielded a remaining-life ratio of 9.6, cumulative damage of 0.60, and next-period failure probability of 0.072. According to Equation (7), the maximum-risk mapping assigned the joint to the warning state, after which Equations (8)-(12) were used to evaluate the preventive-maintenance action.

4. Results Analysis and Model Validation

4.1 Residual Stress Correction Propagation Analysis

The numerical experiments were conducted on Ubuntu 22.04 LTS using an Intel Core i9-13900K CPU, an NVIDIA RTX 4070 GPU with 12 GB memory, and 64 GB RAM. Fatigue-life records were collected from the NIMS Metallic Materials Database and screened from published literature on welded-joint fatigue [21-22]. After duplicate records, run-outs, and entries lacking material grade, joint geometry, loading condition, or fatigue life were removed, 864 experimental records were retained. The dataset covered S500, S690, S700, S780, S800, and S960 steels; butt, cruciform, longitudinal-attachment, and transverse-attachment joints; and GMAW, SMAW, SAW, and FCAW processes. Constant- and variable-amplitude loading conditions were included, with stress ratios ranging from -1.0 to 0.5. Initial residual stress ranged from -65 to 410 MPa. Among the 864 records, 286 contained directly measured residual stress obtained by X-ray diffraction, hole-drilling, or neutron diffraction, whereas the remaining 578 values were reconstructed from the reported test-series statistics.

Table 1. Data composition and computational settings

Category	Setting
Internal dataset	864 experimental fatigue-life records: S500, 118; S690, 156; S700, 144; S780, 112; S800, 178; S960, 156
Joint type	Butt, 236; cruciform, 198; longitudinal attachment, 224; transverse attachment, 206
Welding process	GMAW, 312; SMAW, 168; SAW, 214; FCAW, 170
Loading condition	Constant amplitude, 603; variable amplitude, 261; stress ratio, -1.0-0.5
Residual stress	-65-410 MPa; 286 measured records and 578 test-series-based reconstructed records
Data partition	Grouped by original test series: training, 604; validation, 130; held-out test, 130
External validation	196 experimental records from non-overlapping publications; no record was used for training or hyperparameter selection
Preprocessing	Continuous-variable scaling and categorical encoding were fitted using the training set only
SVR	RBF kernel, $C = 64$, $\gamma = 0.03125$, $\epsilon = 0.02$
RF	600 trees, maximum depth 12, minimum leaf size 2, feature ratio 0.70
XGBoost	700 trees, maximum depth 4, learning rate 0.035, subsample ratio 0.85
LightGBM	800 trees, 31 leaves, maximum depth 8, learning rate 0.030
MLP	128-64-32 hidden units, ReLU, learning rate 10^{-3} , batch size 32, maximum 300 epochs
Meta-learner	Ridge regression, $\alpha = 0.10$, physics weight $\rho = 0.20$, five-fold grouped OOF predictions
Computational cost	Complete training, 41.8 s; inference, 2.7 ms/sample; peak GPU memory, 1.84 GB
RDF-MDP	Five health states, four maintenance actions, average-cost value iteration, convergence tolerance 10^{-5}

The prediction dataset therefore comprised 864 experimentally measured fatigue-life targets, whereas the residual-stress relaxation values and maintenance-state trajectories were used only as model-derived variables. The detailed data composition and computational settings are summarized in Table 1. Records originating from the same specimen series, publication, or S-N curve were assigned to the same subset. Scaling, categorical encoding, residual-stress reconstruction, and hyperparameter selection were completed within the training data, thereby preventing information from the validation and test sets from entering model fitting. Hyperparameters were selected by Bayesian optimization with 100 trials on the validation set. The optimization objective was the mean RMSE obtained by five-fold grouped cross-validation within the training set. For the MLP, early stopping was activated when validation RMSE did not decrease for 30 consecutive epochs. The independent test set and the external dataset were evaluated only after all model and hyperparameter choices had been fixed. On the hardware listed in Table 1, five-fold OOF training of the complete framework required 41.8 s of CPU time, reached a peak GPU memory of 1.84 GB, and consumed 2.7 ms for single-record inference. When the training set was duplicated to 1,728 records for the scaling test, training time increased to 79.6 s, whereas inference time remained below 3.0 ms per record. The five base learners can be trained in parallel, but OOF generation increases training cost approximately in proportion to the number of folds. The framework is therefore suitable for periodic batch updating, although large multi-joint systems require sparse-state or factored-MDP implementations to control state-space growth.

To verify how residual stress correction propagates through the foundational life inputs, this study analyzed the cyclic evolution of $z_i(n)$, M_i , R_i , and E_i , as shown in Figure 5.

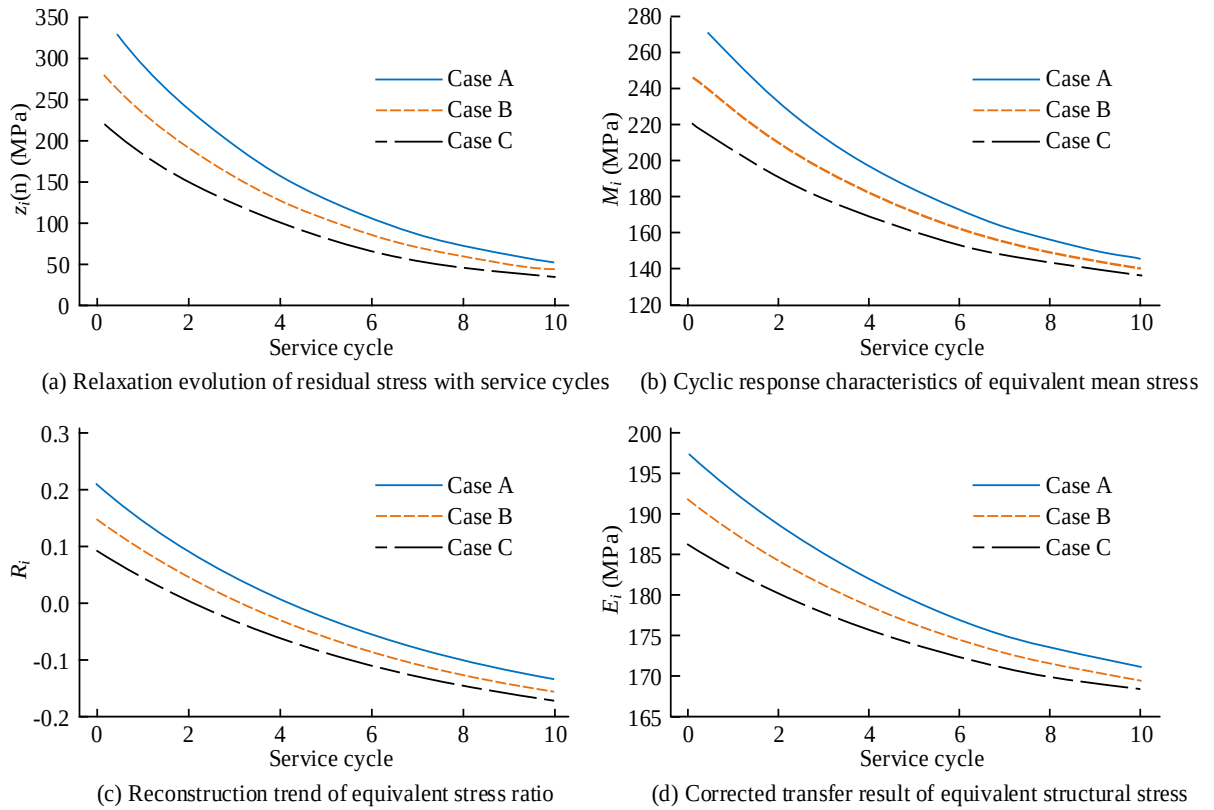


Fig. 5. Effect of residual stress correction on foundational life quantities

As shown in Figure 5, as service cycles increased, the critical-point residual stress $z_i(n)$ predicted by RS-SN-Stack showed a continuous declining trend across all three loading cases: Case A decreased from 325 MPa to 55 MPa, and Case C from 220 MPa to 40 MPa. The equivalent mean stress M_i decreased in tandem with residual stress relaxation, converging from an initial range of 220-268 MPa to 137-145 MPa. The value of the equivalent stress ratio R_i became increasingly negative as the tensile nature of the residual stress made its impact on the mean stress decrease. The equivalent structural stress E_i also decreased from the range of 186–197 MPa to 168–171 MPa.

This demonstrates that residual stress affects master S-N curve life predictions through mean stress, stress ratio, and equivalent structural stress.

4.2 Life Prediction Performance Comparison

This study then compared the predicted master S-N curves of different models against measured data, as shown in Figure 6.

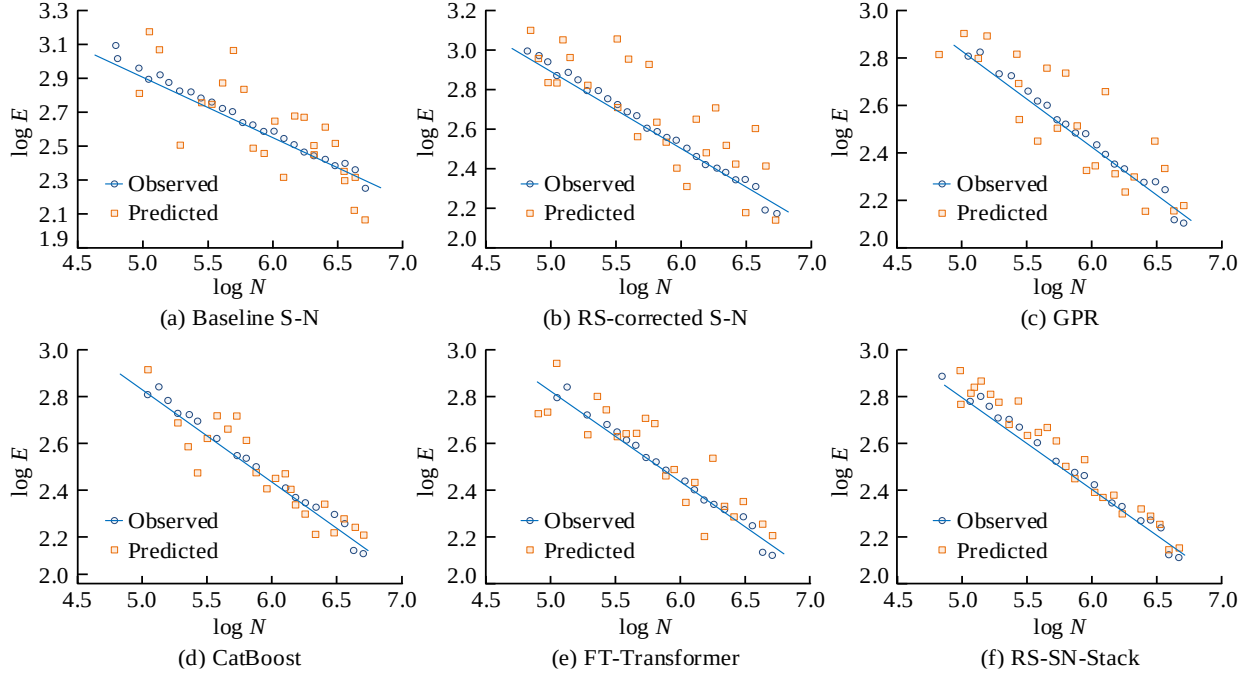


Fig. 6. Comparison of master S-N curves across different models

In Figure 6, the scatter point predictions of both S-N and RS-corrected S-N showed large deviations compared to the measured master S-N curve, whereas GPR, CatBoost, and FT-Transformer resulted in curves close to the measured one. In contrast, the predictions of the proposed model matched well with the measured master S-N curve, indicating that the proposed model achieves higher prediction accuracy while consistently obeying the physical law that fatigue life increases as stress amplitude decreases. These fitting results indicate that combining a physics-based reference life, residual stress correction, and stacking ensemble learning can simultaneously reduce local prediction error and slope deviation of the S-N curve. RMSE, MAE, MAPE, and R2 were calculated on the held-out internal test set and the independent external dataset. The benchmark set comprised the IIW and Eurocode 3 design curves, the fatigue-indicator-based physics-informed neural network reported in, and the Bayesian neural network reported in [13, 23-25]. Across five repetitions of five-fold grouped cross-validation, RS-SN-Stack achieved an RMSE of 0.064 ± 0.006 , an MAE of 0.049 ± 0.004 , a MAPE of $3.42\% \pm 0.31\%$, and an R2 of 0.938 ± 0.008 ; the narrow variations indicate stable performance across source-level partitions. The final test results are summarized in Table 2.

In Table 2, RS-SN-Stack achieved an RMSE of 0.058 and an R2 of 0.947 on the internal test set. On the independent external dataset, its RMSE increased to 0.084 and R2 decreased to 0.879, indicating a measurable effect of differences in material grade, weld geometry, and loading distribution. Nevertheless, it retained lower errors than the deterministic, PINN, and BNN comparison methods. The external result supports transfer across non-overlapping data sources but does not imply unrestricted generalization to all welded materials.

Under the same internal test split, RS-SN-Stack achieved an RMSE of 0.058, representing reductions of 37.0% and 32.6% relative to the reimplemented PINN model (0.092) and BNN model (0.086), respectively [13, 23]. Recent physics-informed GPR, modified PINN, and BNN studies have shown that incorporating fatigue mechanisms or probabilistic learning can improve physical consistency and uncertainty representation under limited-data conditions [12-13,23]. The present results

further indicate that heterogeneous model fusion can reduce prediction error while retaining the residual-stress-dependent fatigue trend. Unlike existing studies that terminate at fatigue-life prediction, the proposed framework converts the prediction outputs into remaining-life, cumulative-damage, and failure-probability states for maintenance optimization. However, direct numerical comparison with published errors should be interpreted cautiously because joint geometries, stress ratios, loading spectra, and dataset partitions differ among studies.

Table 2. Internal and external fatigue-life prediction results

Dataset	Model	RMSE	MAE	MAPE (%)	R ²
Internal test	IIW	0.244	0.193	13.4	0.122
	Eurocode 3	0.236	0.187	12.9	0.151
	Baseline S-N	0.230	0.184	12.6	0.176
	RS-corrected S-N	0.193	0.151	10.9	0.420
	GPR	0.125	0.098	6.7	0.759
	CatBoost	0.104	0.082	5.9	0.833
	FT-Transformer	0.110	0.087	6.2	0.811
	PINN	0.092	0.071	4.9	0.868
	BNN	0.086	0.066	4.5	0.884
	RS-SN-Stack	0.058	0.044	3.1	0.947
External	IIW	0.262	0.208	14.5	0.096
	Eurocode 3	0.251	0.199	13.9	0.128
	CatBoost	0.139	0.110	7.5	0.702
	PINN	0.124	0.096	6.6	0.758
	BNN	0.117	0.090	6.1	0.783
	RS-SN-Stack	0.084	0.064	4.4	0.879

4.3 Ablation and Interpretability Analysis

To quantify the contributions of residual stress, physical baseline life, and physical consistency, the ablation variants were evaluated on the same held-out test set used in Table 2. SHAP importance, dependence, and interaction analyses were then applied to the complete model, with the results shown in Figure 7.

In Figure 7(a), adding residual stress, physical baseline life, and the physics-constrained loss successively reduced RMSE from 0.168 to 0.131, 0.104, and 0.058. MAPE decreased from 10.8% to 3.1%, while R² increased from 0.862 to 0.947. Figure 7(b) shows mean absolute SHAP values of 0.241, 0.200, and 0.161 for equivalent structural stress, structural stress amplitude, and physical baseline life, respectively. In Figure 7(c), increasing equivalent structural stress from 168 to 198 MPa reduced the predicted logarithmic life by approximately 0.31. Figure 7(d) further shows that residual stress above 220 MPa strengthened the negative contribution of equivalent structural stress, indicating that the two stress variables acted jointly rather than independently. The meta-learner comparison produced RMSE values of 0.064, 0.067, 0.062, 0.061, and 0.058 for linear regression, Lasso, random forest, XGBoost, and Ridge, respectively. Ridge also required only 0.02 s for second-level fitting, compared with 1.84 s for random forest and 2.16 s for XGBoost. It was therefore retained because it achieved the lowest error without adding substantial second-level complexity.

4.4 Maintenance Strategy Comparison

This study then compared different maintenance strategies across three dimensions: cost, risk, and efficiency. The O&M optimization results are presented in Table 3. Table 3 shows that RDF-MDP opportunistic maintenance achieved the lowest long-term average cost, failure penalty, average failure probability, high-risk-state ratio, and downtime among the four strategies. Compared with fixed periodic maintenance, it reduced the long-term average cost from 168.4×10^3 to 126.5×10^3 , the failure penalty from 39.4×10^3 to 20.8×10^3 , and average downtime from 19.2 to 11.8 h. These improvements resulted from coordinating adjacent maintenance tasks within a common logistics mobilization and shutdown window, thereby reducing repeated mobilization and disassembly.

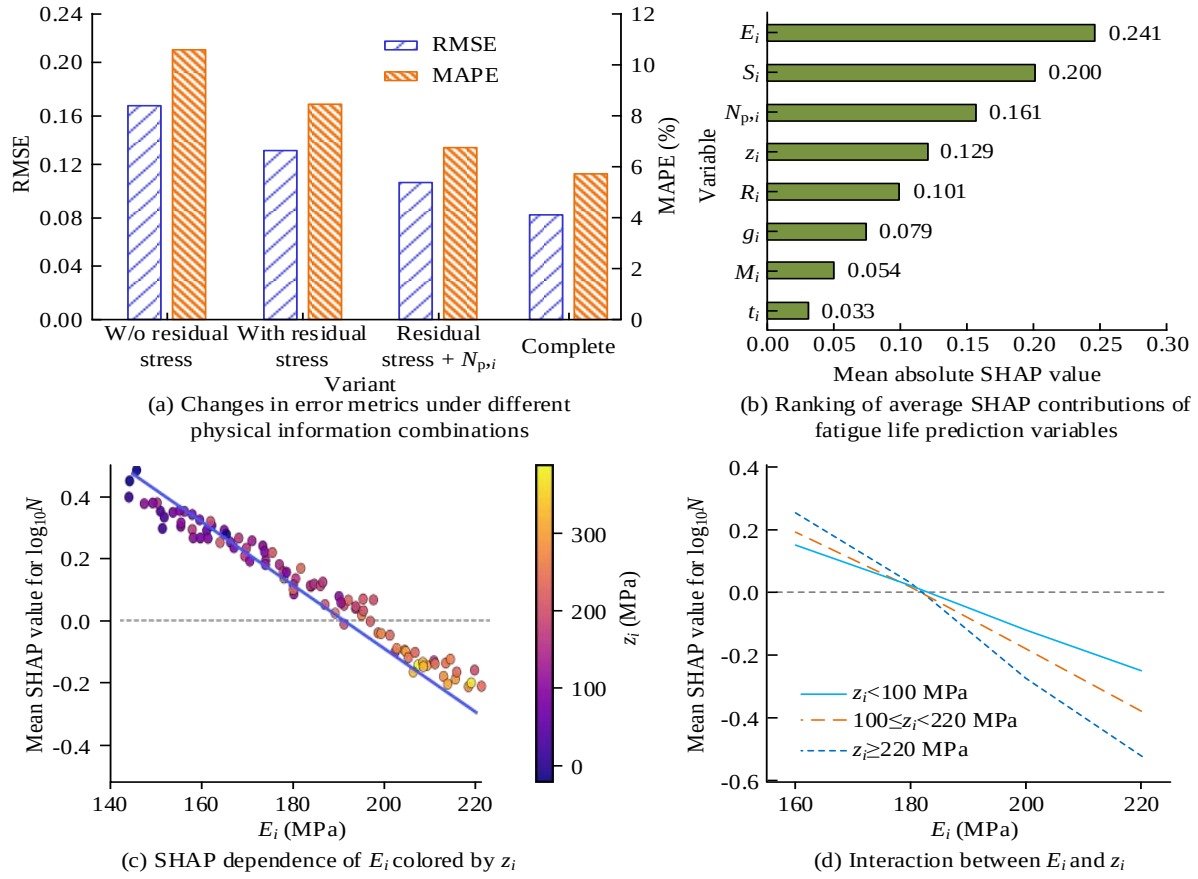


Fig. 7. Ablation study and variable contribution analysis of RS-SN-Stack

Note: Complete=Residual stress + $N_{p,i}$ + physics-guided loss

Table 3. Comprehensive comparison of periodic maintenance strategies

Strategy	$J(\pi)$ ($\times 103$)	Failure penalty ($\times 103$)	Average failure probability	High-risk state ratio	Average downtime (h)	OM triggers
Fixed periodic	168.4	39.4	0.083	0.214	19.2	0
Threshold- based	149.7	30.7	0.061	0.173	16.4	0
State-driven w/o OM	137.2	25.9	0.049	0.152	14.1	0
RDF-MDP OM	126.5	20.8	0.036	0.121	11.8	8

For industrial deployment, the trained RS-SN-Stack model can receive periodically updated inputs from load monitoring, weld-toe inspection, and residual-stress measurements. At each inspection cycle, the model updates fatigue life, remaining life, cumulative damage, and failure probability, after which RDF-MDP selects no action, inspection, preventive maintenance, or repair/replacement. When the condition in Equation (12) is satisfied, eligible neighboring joints are coordinated within the same shutdown window. The inference time of 2.7 ms per joint permits cycle-level deployment without continuous online computation. Nevertheless, field implementation requires site-specific calibration of residual-stress relaxation and cost coefficients, and the recommended action should remain subject to engineering review.

4.5 Uncertainty Sensitivity Analysis

To quantify uncertainty propagation, 10,000 Monte Carlo trials were performed by jointly sampling the residual-stress relaxation coefficient, life-prediction residual, and each row of the Markov

transition matrix. The effects of physical weight and state discretization were evaluated within the same analysis, as shown in Figure 8.

In Figure 8(a), increasing the uncertainty scale from 0.04 to 0.16 raised average failure probability from 0.016-0.032 to 0.096-0.112. The Monte Carlo results in Figure 8(b) yielded a long-term average cost of 127.1 with a 95% interval of 121.6-133.9, an average failure probability of 0.038 with a 95% interval of 0.030-0.049, and downtime of 12.1 h with a 95% interval of 10.7-13.8 h. Figure 8(c) shows that $\rho=0.20$ provided the minimum test RMSE of 0.058 while reducing the mean physical-baseline deviation from 0.132 at $\rho=0$ to 0.091. Increasing ρ to 1.00 further reduced the physical deviation to 0.051 but increased RMSE to 0.079. In Figure 8(d), the three-, five-, and ten-state systems produced average costs of 132.8, 126.5, and 124.9 and solution times of 0.42, 1.38, and 9.84 s. The five-state system retained most of the cost and risk improvement of the ten-state system while avoiding its substantially greater computational cost.

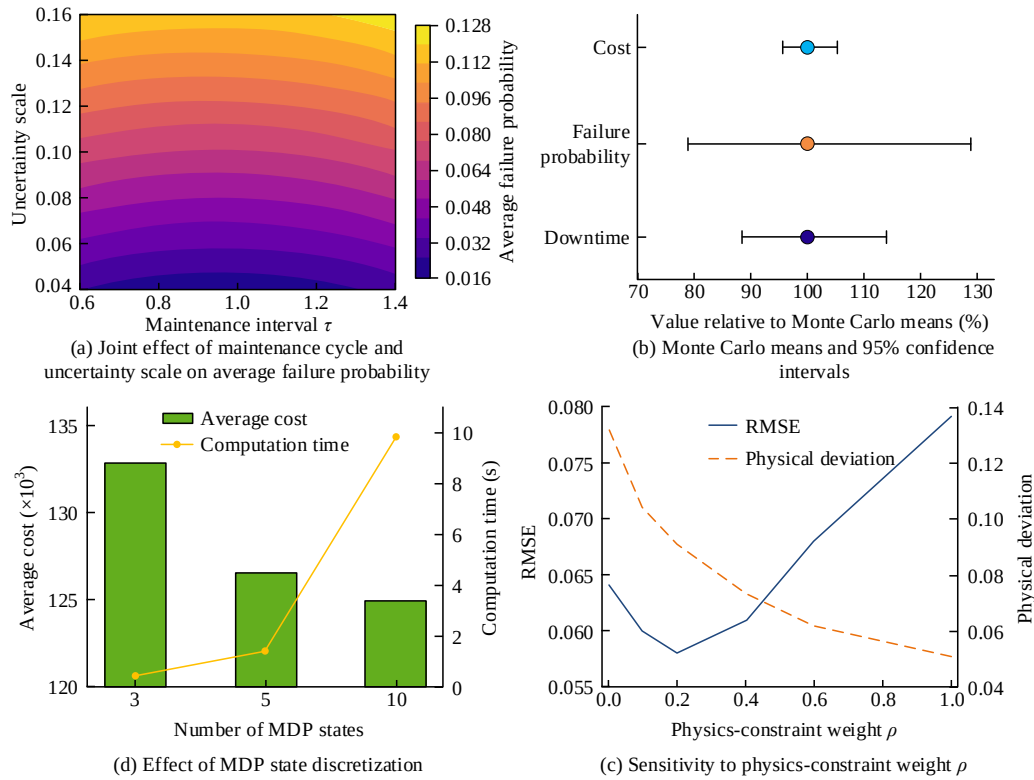


Fig. 8. Uncertainty propagation and parameter sensitivity of RS-SN-Stack and RDF-MDP

The residual-stress measurement technique affects the input $z_{0,i}$ and the fitted relaxation coefficient λ_i . X-ray diffraction mainly measures near-surface stress and is sensitive to surface preparation, whereas hole-drilling provides a depth-dependent local estimate but introduces limited material removal. Neutron diffraction can characterize through-thickness stress but has lower accessibility and coarser spatial resolution compared to X-ray diffraction. In the Monte Carlo analysis, the corresponding measurement standard deviations were set to 20, 30, and 15 MPa, respectively. Repeating the prediction with these error levels changed test RMSE from 0.058 to 0.061, 0.064, and 0.060, indicating that measurement inconsistency should be propagated rather than ignored. Because fatigue parameters and residual-stress evolution are material-dependent, application of the computational framework to aluminum or stainless-steel welded joints requires material-specific recalibration. Transfer requires recalibration of the S-N parameters, residual-stress relaxation distribution, elastic-plastic stress-ratio correction, and crack-growth coefficients. Aluminum joints additionally require explicit treatment of the absence of a conventional fatigue limit, whereas stainless-steel joints require material-specific cyclic hardening and mean-stress parameters. Accordingly, the reported predictive and maintenance performance applies to the HSS welded joints represented in the present datasets. Within the validated HSS domain, prediction robustness was also evaluated across different stress-ratio ranges and loading spectra, as summarized in Table 4.

Table 4. Prediction robustness under different stress ratios and loading spectra

Grouping variable	Level	Samples	RMSE	MAE	MAPE (%)	R ²
Stress ratio	$R \leq 0$	238	0.061	0.046	3.4	0.941
Stress ratio	$0 < R \leq 0.2$	412	0.056	0.043	3.0	0.951
Stress ratio	$R > 0.2$	214	0.067	0.052	3.8	0.927
Loading spectrum	Constant amplitude	603	0.057	0.044	3.1	0.949
Loading spectrum	Variable amplitude	261	0.066	0.051	3.7	0.932

In Table 4, the error increase at $R > 0.2$ and under variable-amplitude loading indicates that high mean stress and load-sequence variation remain the principal loading-related sources of generalization error. Nevertheless, all subgroup RMSE values remained below 0.07.

5. Summary and Outlook

To connect fatigue-life prediction with periodic O&M decisions for high-strength steel welded joints, an integrated RS-SN-Stack-RDF-MDP framework has been developed. Residual-stress relaxation, equivalent structural stress, physical S-N life, weld geometry, and crack-initiation and propagation descriptors were first combined to construct fatigue inputs. Five base learners generated grouped out-of-fold predictions, which were fused by a physics-constrained Ridge meta-learner. The fused life estimate was subsequently converted into remaining life, cumulative damage, and next-period failure probability and supplied to a five-state average-cost Markov decision model.

The internal held-out test results show that RS-SN-Stack achieved an RMSE of 0.058, MAE of 0.044, MAPE of 3.1%, and R² of 0.947. Five repetitions of grouped cross-validation produced an RMSE of 0.064 ± 0.006 , confirming that the result was stable across source-level partitions. On an external dataset containing 196 records from non-overlapping publications, RMSE increased to 0.084 and R² decreased to 0.879. This decrease indicates that differences in steel grade, weld geometry, and loading distribution remain relevant, although the proposed model retained lower errors than the PINN, BNN, and deterministic comparison methods. Ablation analysis further showed that residual stress, physical baseline life, and physical consistency reduced RMSE from 0.168 to 0.058. Equivalent structural stress, structural stress amplitude, and physical baseline life produced mean absolute SHAP values of 0.241, 0.200, and 0.161, respectively.

Under the complete cost formulation incorporating labor, logistics, spare parts, downtime, production loss, and failure penalties, RDF-MDP opportunistic maintenance reduced long-term average cost from 168.4×10^3 to 126.5×10^3 , failure penalty cost from 39.4×10^3 to 20.8×10^3 , average failure probability from 0.083 to 0.036, and average downtime from 19.2 to 11.8 h. Monte Carlo propagation yielded a mean failure probability of 0.038 with a 95% interval of 0.030-0.049. The five-state system retained most of the risk reduction of the ten-state system while requiring substantially less solution time.

These results demonstrate that residual-stress-corrected fatigue prediction can be converted into quantitative maintenance states without separating prediction and decision models. Nevertheless, the present validation is limited to the investigated HSS grades and available residual-stress measurements. Transfer to aluminum, stainless steel, or long-term field monitoring requires material-specific recalibration, time-varying transition probabilities, and prospective validation using continuously observed inspection and maintenance records.

Appendix

Table A1. Meta-learner comparison

Meta-learner	RMSE	Fitting time (s)
Linear regression	0.064	0.01
Lasso	0.067	0.02
Random forest	0.062	1.84
XGBoost	0.061	2.16
Ridge	0.058	0.02

References

- [1] Grönlund K, Ahola A, Lipiäinen K, Pesonen T, Björk T, Moshtaghi M. Fatigue performance of welded joints made of high-strength steel under variable amplitude loading conditions and varying mean stress. *Weld World*. 2025;69(8):2253-2268. <https://doi.org/10.1007/s40194-025-01989-5>
- [2] Schubnell J, Konidena SK, Jung M, Braun M, Ehlers S, Madia M, et al. Approach for the probabilistic fatigue assessment of welded joints based on the local geometry of the weld seam. *Fatigue Fract Eng Mater Struct*. 2024;47(1):88-107. <https://doi.org/10.1111/ffe.14170>
- [3] Schubnell J, Fliegenger S, Rosenberger J, Feth S, Braun M, Beiler M, et al. Data-driven fatigue assessment of welded steel joints based on transfer learning. *Weld World*. 2025;69(8):2223-2238. <https://doi.org/10.1007/s40194-025-01967-x>
- [4] Hills MA, Becker TH. Machine learning-based prediction of fatigue strength in additively manufactured Ti-6Al-4V parts: a sensitivity analysis of input features. *J Mater Eng Perform*. 2025;34(15):16801-16821. <https://doi.org/10.1007/s11665-025-11115-7>
- [5] Rozas H, Xie W, Gebraeel N. Condition-based maintenance for wind farms using a distributionally robust chance constrained program. *IEEE Trans Power Syst*. 2024;40(1):231-243. <https://doi.org/10.1109/TPWRS.2024.3406335>
- [6] Cai Y, de Jonge B, Teunter RH. Data-driven condition-based maintenance optimization given limited data. *Eur J Oper Res*. 2025;324(1):324-334. <https://doi.org/10.1016/j.ejor.2025.01.010>
- [7] Ono Y, Remes H, Kinoshita K, Yildirim HC, Nussbaumer A. Local relaxation of residual stress in high-strength steel welded joints treated by HFMI. *Weld World*. 2024;68(8):2187-2202. <https://doi.org/10.1007/s40194-024-01789-3>
- [8] Ahola A, Leitner M, Grönlund K, Brunnhofer P, Buzzi C, Moshtaghi M, et al. Fatigue assessment of as-welded and HFMI-treated high-strength steel joints under variable amplitude loading using local approaches. *Weld World*. 2025;69(3):687-700. <https://doi.org/10.1007/s40194-024-01919-x>
- [9] Song W, Man Z, Xu J, Fan Y, Xia X, He M, et al. Fatigue reliability assessment of load-carrying cruciform welded joints with undercuts and misalignments. *Fatigue Fract Eng Mater Struct*. 2024;47(2):511-531. <https://doi.org/10.1111/ffe.14205>
- [10] Ghanadi M, Kumar M, Danielsson PO, Hultgren G, Barsoum Z. Unsupervised machine learning for local stress identification in fatigue analysis of welded joints. *Weld World*. 2025;69(1):213-226. <https://doi.org/10.1007/s40194-024-01868-5>
- [11] Wang L, Qian X, Feng L. Effect of welding residual stresses on the fatigue life assessment of welded connections. *Int J Fatigue*. 2024;189:108570. <https://doi.org/10.1016/j.ijfatigue.2024.108570>
- [12] Kim D, Kim DY, Ko T, Lee SH. Physics-informed Gaussian process regression model for predicting the fatigue life of welded joints. *Int J Fatigue*. 2025;190:108644. <https://doi.org/10.1016/j.ijfatigue.2024.108644>
- [13] Li X, Fu Z, Shu J, Ji B, Ji B. A modified physics-informed neural network to fatigue life prediction of deck-rib double-side welded joints. *Int J Fatigue*. 2024;189:108566. <https://doi.org/10.1016/j.ijfatigue.2024.108566>
- [14] Karakaya Ş, Yildirim M, Gebraeel N, Xia T. A sensor-driven operations and maintenance planning approach for large-scale leased manufacturing systems. *Int J Prod Res*. 2024;62(24):8701-8718. <https://doi.org/10.1080/00207543.2024.2347564>
- [15] Xu J, Liu B, Zhao X, Wang XL. Online reinforcement learning for condition-based group maintenance using factored Markov decision processes. *Eur J Oper Res*. 2024;315(1):176-190. <https://doi.org/10.1016/j.ejor.2023.11.039>
- [16] Chai X, Kilic OA, Veldman J, Teunter RH, Zhao X. Condition-based reallocation and maintenance for a 1-out-of-2 pairs balanced system. *Eur J Oper Res*. 2024;318(2):618-628. <https://doi.org/10.1016/j.ejor.2024.05.028>
- [17] Eggertsson R, Basten R, van Houtum GJ. Maintenance optimization for capital goods when information is incomplete and environment-dependent. *IIE Trans*. 2024;56(11):1146-1161. <https://doi.org/10.1080/24725854.2023.2257245>
- [18] Grönlund K, Ahola A, Riski J, Pesonen T, Lipiäinen K, Björk T. Overload and variable amplitude load effects on the fatigue strength of welded joints. *Weld World*. 2024;68(2):411-425. <https://doi.org/10.1007/s40194-023-01642-z>
- [19] Yoshihara Y, Osawa N, Murakawa H, Dai P. Study on the estimation method for the mean stress effect on the fatigue strength of welded joints with various failure modes and joint types. *Weld World*. 2025;69(1):151-167. <https://doi.org/10.1007/s40194-024-01815-4>
- [20] Kim JH, Jang SW, Kim MH. An investigation of compressive mean stress effect on fillet welded joints with high strength steel. *Weld World*. 2025;69(8):2311-2322. <https://doi.org/10.1007/s40194-025-02028-z>
- [21] Deng C, Luo C, Gong B, Liu Y, Zhang Z, Wang S, et al. A dataset of fatigue properties for welded joints. *Sci Data*. 2025;12:1815. <https://doi.org/10.1038/s41597-025-06067-5>

- [22] Matsuda A, Demura M, Yamazaki M, Kadohira T, Ashino T, Furuya Y, et al. A unified model for metallic materials reliability data and its application in NIMS metallic materials database (Kinzoku). *Sci Technol Adv Mater Methods*. 2025;5(1):2518745. <https://doi.org/10.1080/27660400.2025.2518745>
- [23] Centola A, Ciampaglia A, Paolino DS, Tridello A. Probabilistic machine learning for preventing fatigue failures in additively manufactured SS316L. *Eng Fail Anal*. 2025;168:109081. <https://doi.org/10.1016/j.engfailanal.2024.109081>
- [24] Hobbacher AF, Baumgartner J. Recommendations for fatigue design of welded joints and components. 3rd ed. Cham: Springer; 2024. <https://doi.org/10.1007/978-3-031-57667-6>
- [25] European Committee for Standardization. EN 1993-1-9:2025 Eurocode 3: Design of Steel Structures-Part 1-9: Fatigue. Brussels: CEN; 2025.