Using Reinforcement Learning to Optimize the Formulation Design of Water-Based Coatings

Xiaoli Zhao, Min Li, Wan Li, Kun Yang, Yanlin Liu

Abstract—**Traditional solvent based coatings generate a large amount of volatile organic compounds during application, which can cause secondary environmental pollution when released into the atmosphere. However, the application performance of water-based coatings is often not ideal. To address these challenges, this study employs reinforcement learning (RL) algorithms to optimize the formulation of water-based coatings, parameterize the formulation process, and create intelligent design models. The object detection function in RL is used to identify key elements in water-based paint formulations. By combining data analysis with RL algorithms, the design process was optimized and key factors affecting coating performance were identified. In order to improve the corrosion resistance and high-temperature stability of water-based coatings, an optimization model combining neural networks with RL is proposed. Further optimize the formulation process using orthogonal method. The RL algorithm was used to customize and optimize the parameters of water-based coating formulations, and the effectiveness of these optimizations was verified through experiments. Specifically, the resin derived from organic modified acrylic acid is degreased and cleaned, maintaining a surface viscosity of 12 to 16 seconds at room temperature of 25 ° C. The viscosity content of the coating exceeds 27%. The results indicate that water-based coatings optimized by RL algorithm are not only more environmentally friendly, but also exhibit excellent performance, including enhanced corrosion resistance, high temperature stability, coverage strength, impact strength, and overall performance.**

Index Terms—**Reinforcement Learning, Water-based Coatings, Paint Formulation, Performance Testing, Parameter Optimization**

I. Introduction

COATINGS have a history of over 4,000 years in China, with vegetable oil serving as the primary component in early formulations. These coatings are commonly referred to as paints and have been in use for a long time [1]. In China, the performance of coatings is defined as follows: Coatings are materials applied to the surface of objects using various processes. Once applied, they will form solid thin films with specific strength and resistance characteristics, providing protection and decoration. The key components of coatings include resins, solvents, pigments, and additives [2]. Resin plays a crucial role in determining the basic performance of coatings. Traditional coatings mainly rely on natural resins, resulting in significant resource consumption and costs [3]. However, modern coatings utilize synthetic resins to improve the efficiency and stability of formulations. Solvents facilitate the dissolution and uniform distribution of resins during the production process, including substances such as water, inorganic compounds, and organic compounds [4]. Pigments help improve the heat resistance, coverage ability, strength, and corrosion resistance of coatings, enabling them to meet various performance requirements. Meanwhile, additives are auxiliary substances added during the production process to reduce defects and improve the performance of the final product [5].

Traditional formula design methods often rely on empirical methods and trial and error experiments, which are not only inefficient but also fail to fully explore the vast formula space. This limitation makes it challenging to identify high-performance formulas in a timely manner. Therefore, developing efficient and intelligent formulation design methods has become a key area of research in coatings, especially in addressing the complexity associated with water-based coating formulations. RL is an advanced machine learning technique that demonstrates enormous potential in solving complex optimization problems due to its powerful decision-making ability and adaptability. By redefining the formulation design of water-based coatings as a decision-making process within the RL framework, this method utilizes the intelligent search function of RL to quickly identify high-performance formulation combinations in a wide formulation space. In addition, RL improves and optimizes decision-making strategies through continuous interaction with the environment (represented by coating performance testing), thereby enhancing the efficiency and accuracy of formulation design. Coatings can be classified into various types based on their composition and form, such as solvent based coatings, water-based coatings, and solid coatings. Among them, water-based coatings have received extensive research attention due to their unique properties and formulation challenges. Most studies focus more on the performance and formulation design of water-based coatings. Water based coatings use water molecules as a medium instead of organic solvents, and have become an important component of industrial products [6]. Their use greatly reduces the release of harmful substances and allows for the use of low pollution additives, which have advantages over traditional solvent based coatings. In addition to environmental protection, water-based coatings also provide

Manuscript received September 29, 2024; revised March 18, 2025.

Xiaoli Zhao is an engineer of Shandong Yong'an Rubber Industry Co., Ltd, Weifang 262600, China (corresponding author, e-mail: 18705365167@163.com).

Min Li is an engineer of Shandong Yong'an Rubber Industry Co., Ltd, Weifang 262600, China. (e-mail: min.li@china-yongan.com.cn).

Wan Li is an engineer of Shandong Yong'an Rubber Industry Co., Ltd, Weifang 262600, China. (e-mail: min.li@china-yongan.com.cn).

Kun Yang is an engineer of Shandong Yong'an Rubber Industry Co., Ltd, Weifang 262600, China. (e-mail: kun.yang@china-yongan.com.cn).

Yanlin Liu is an engineer of Shandong Yong'an Rubber Industry Co., Ltd, Weifang 262600, China. (e-mail: yanlin.liu@china-yongan.com.cn).

health benefits and reduce fire risks during the manufacturing process by eliminating the use of flammable solvents, thereby improving the safety of coating production. However, the dependence on water as a dispersing medium brings certain limitations. The high evaporation potential and surface tension of water pose challenges to water-based coatings, leading to specific drawbacks [7]. For example, water-based coatings have lower construction tolerances and stricter requirements for environmental conditions during application. These coatings are highly sensitive to factors such as temperature and humidity, thus requiring strict material preparation standards. Therefore, compared to solvent based coatings, water-based coatings have higher formulation and application costs [8]. In addition, water-based coatings often exhibit performance defects, including poor corrosion resistance, limited temperature resistance, and shortened service life. For example, most water-based coatings have difficulty maintaining long-term corrosion resistance in submerged or marine environments, which is particularly evident in large-scale applications [9]. Similarly, the high temperature environment exacerbates application challenges, making it difficult for water-based coatings to achieve the performance level of solvent based coatings. These limitations have prompted researchers to explore advanced technologies such as intelligent algorithms to optimize water-based coating formulations. These methods aim to address performance deficiencies by improving performance such as corrosion resistance, temperature resistance, and durability.

Traditional coatings are mainly composed of solvents, which account for over 50% of their composition [10]. Solvent based volatile organic compounds (VOCs) have a high concentration and undergo photochemical reactions when exposed to sunlight. These reactions allow volatile organic compounds to enter the atmosphere, causing serious environmental damage [11]. For example, ozone and aldehydes produced in the lower atmosphere pose significant risks to human health and the environment. With the rapid development of industry, the use of coatings in modern buildings is increasing, highlighting the necessity of addressing environmental and safety issues related to coatings [12]. Water based coatings are recognized for their environmental friendliness, as they use water molecules as transfer media, combined with resins and additives, to create low-risk alternatives to traditional solvent based coatings [13]. However, water-based coatings face significant drawbacks, including easy evaporation, poor corrosion resistance, low high temperature resistance, and long drying time. Addressing these limitations has become a focus of research [14]. RL is a prominent artificial intelligence (AI) technology that has significant advantages in data processing, analysis, and detection. Through continuous interaction with the environment during data training, RL extracts insights from experience and iteratively improves the model based on this feedback [15]. Recently, the application of RL algorithm in optimizing the formulation parameters of water-based coatings has attracted widespread attention, triggering a surge in research activities in this field. This study combines the latest advances in RL and coating formulation optimization to enhance its academic rigor and depth. For example, Maity et al. analyzed the dynamics of unstable thin liquid film flow on a porous stretched cylinder, providing key mathematical modeling references and information for the RL based optimization algorithm used in this study [16]. Similarly, the dynamic modeling method developed by Suganya G and Senthamarai R in pest control research provides valuable theoretical support for dynamically controlling the coating performance of water-based coatings [17]. In addition, Zuo H et al. introduced a hybrid approach that combines extended tree enhanced naive Bayes classifiers with generative adversarial networks (GANs). Although their research focuses on classification problems in machine learning, their proposed GAN based approach provides valuable perspectives for parameter tuning and optimization, which are applied to the RL model used in this study [18].

Traditional data analysis and experiential learning algorithms typically update parameters through random extraction and calculate weight step sizes during the interaction process, while ignoring the informational value of interaction on learning. This limitation leads to low reinforcement efficiency in learning [19]. However, reinforcement learning algorithms have become a superior solution among various technologies and are becoming increasingly popular in various fields [20]. The development of RL algorithms has a long history in the United States, where they are widely applied in machine learning and artificial intelligence to solve complex problems and facilitate effective decision-making processes [21]. Researchers have determined that in the optimization process of RL, the expected value of the solution can be accurately estimated, thereby enabling more effective problem-solving. Meanwhile, British researchers have demonstrated that combining RL with neural networks can create virtual models of images and videos, resulting in immersive animation effects. Although neural networks enhance dimension computation in RL, problems such as unreliable computational fitting still exist [22]. To address this issue, the separation of behavior and evaluation was introduced, significantly reducing the instability of RL in data processing. Based on these findings, reinforcement learning algorithms have demonstrated the ability to optimize and adjust parameters through periodic response learning without the need to identify specific content in test data [23]. This method is applied to optimize the formulation design of water-based coatings, allowing for a deeper exploration of their potential.

This study introduces advanced artificial intelligence technology RL to optimize the formulation of water-based coatings. Developed and validated an intelligent optimization model based on RL algorithm, enabling automatic adjustment of key parameters in water-based coating formulations. This method aims to improve the overall coating performance, including corrosion resistance, high temperature resistance, coverage strength, and impact resistance. By effectively applying the RL algorithm to the complex environment of water-based coating formulation design, precise optimization of the formulation can be achieved. The key issues addressed in this study include how the RL model identifies and prioritizes key factors that affect the performance of water-based coatings during the optimization process.

The main contributions and innovations of this study are as follows:

(1) This study innovatively designed an advanced optimization model that integrates neural networks and reinforcement learning (RL). This model cleverly integrates the excellent nonlinear data fitting characteristics of neural networks with the intelligent decision-making and optimization potential of reinforcement learning, opening up a novel and efficient path for coating formulation optimization.

(2) This study introduced the orthogonal experimental design method. This method significantly reduces the necessary number of tests through scientific experimental arrangements, while ensuring the comprehensiveness and representativeness of experimental data, laying a solid foundation for quickly locating the optimal formula combination.

(3) In the process of formula optimization, this study not only focused on improving a single performance indicator, but also conducted comprehensive performance tests on the coating formula before and after optimization, covering multiple key dimensions such as corrosion resistance, high temperature resistance, coverage strength, and impact resistance. This comprehensive and multi-scale evaluation strategy not only verifies the effectiveness and stability of the optimization scheme, but also greatly enhances the practical application value and social influence of the research results.

II. Methods

A. Research on water-based coating formulation performance based on RL object detection

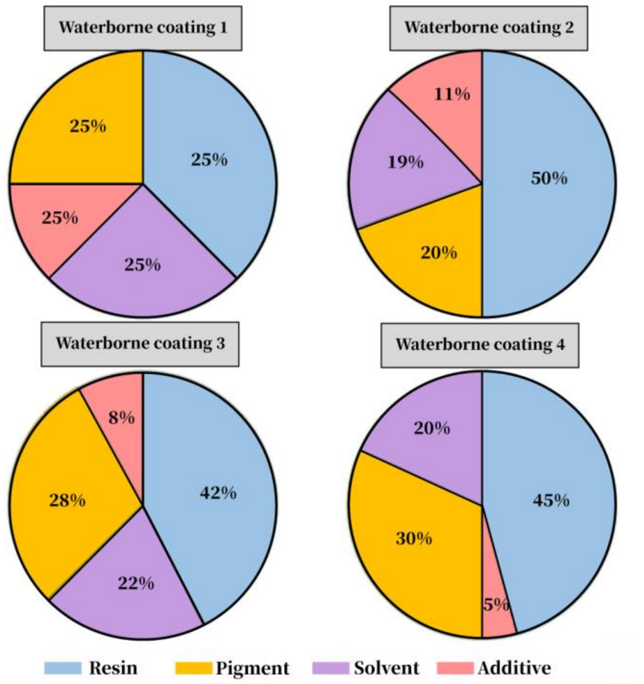

Fig. 1. Proportional importance of constituent elements in water-based coatings

Water based coatings can be divided into two main types based on their drying mechanism and internal components: automatic drying and machine drying. Most water-based coatings are composed of inorganic and organic compounds, including acrylic acid, modified epoxy esters, and water-based fluorocarbon compounds. Resin media play a crucial role in determining the basic properties of water-based coatings, influenced by the resin composition and its interactions with solvents and pigments. Conduct statistical analysis to evaluate the proportion importance of key components (resin, solvent, pigment, and additive) in water-based coatings. The results of this analysis are shown in Fig. 1.

As shown in Fig. 1, four types of water-based coatings were selected, each with a different proportion of internal components. However, a common trend is that resin media account for the largest proportion, followed by pigments, solvents, and finally additives. According to existing literature, water-based acrylic acid has good corrosion resistance and is abundant in water-based alkyd resin formulations. Some water-based coatings dry slowly but have excellent performance after film formation, while others have fast drying time and fast mechanical properties, but their durability is moderate. These differences highlight how changes in formula design can lead to different application outcomes. The quality of water-based coating formula design directly affects the service life and performance of the coating after application. The formulation design process of water-based coatings is similar to that of solvent based coatings, but it also has its unique features in practical applications. When optimizing formula design, several key factors must be considered. Firstly, it is necessary to analyze the relationship between the formulation of water-based coatings and their expected applications. This includes determining the performance requirements of the coating based on the internal system of the product, as a guide for formulation design. The original formula ratio will continue to evolve to meet the needs of the covering body. Secondly, selecting the appropriate proportion of resin, pigment, and additive is crucial for controlling the performance of the coating. After formulation, the design system should be continuously improved through evaluation. Finally, the complexity of the coating formula must be considered. Some low complexity coatings do not require additives, while others may only require a small amount of additives to achieve the desired coating performance. Designing all formulas based on high proportions and standards will unnecessarily increase the cost of water-based coatings. In addition, even for coatings with similar requirements and types, there are significant differences in formulation ratios, especially in the selection of other media. Therefore, compared with solvent based coatings, the production process and formulation design of water-based coatings are more sensitive and complex. This study aims to optimize the formulation design through a systematic approach, and explore the factors that affect the performance of water-based coatings by analyzing data and parameter changes.

The RL algorithm interacts with the analyzed object and its surrounding environment to determine the most suitable formula design through trial and error. By incorporating the decision feedback from these interactions into the learning model, the matching probability of the formula increases when the target object receives positive incentives during the matching process. On the contrary, if the target object receives a negative penalty after completing the matching feedback, the probability of choosing this formula will decrease, thereby optimizing the design. In this process, the interaction between the target object and its environment involves analyzing the rewards or punishments received in a specific context, guiding the development of a model to adjust its behavior. Based on this feedback, the formulation design transitions from one state to another and continuously receives performance feedback, optimizing to achieve the highest performance of water-based coatings. RL uses Markov decision processes to determine the state of the target before and after an action. The process is shown in equation (1).

$$P(S_{t+1} \mid S_t S_{t+1} S_{t-2} \ldots S_0) = P(S_{t+1} \mid S_t) \tag{1}$$

The random calculation process is represented by a combination of elements, where all sets represent the target state. All possible actions of the target formula are integrated together, and the reward function provides incentives in the formula optimization environment, promoting data transmission and computation in the conventional environment. Starting from a certain point, select transition probability for strategy determination, and then provide environmental feedback. The cumulative probability of the target object is represented by equation (2):

$$R = \sum_{S}^{T} t - \gamma^t r_t \tag{2}$$

In equation (2), T represents the time step or total time range, represents the number of cycles in which the algorithm makes decisions and receives feedback, and S represents the state space that contains all possible states. In RL, agents are in a specific state at each time step and make decisions based on that state. In addition, T is used to represent the current time step or iteration count. During algorithm execution, t starts from an initial value (such as 0 or 1) and increases at each iteration until the predetermined termination condition is reached. R is a reward function that guides the learning process of intelligent agents by optimizing their behavior to achieve higher cumulative rewards.

The parameter $\gamma^t \in [0,1]$ represents the impact of future feedback. The decrease in the influence coefficient is calculated and integrated into the mathematical convergence process to yield the final result. Upon completion of the expected value of a state, it is termed the total value. This decision value can be expressed in functional form, denoted as Eq. (3):

$$V(s) = E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots \mid s_t = s] \tag{3}$$

In Eq. (3), *V* (*s*) represents the value function or expected value at state *s*. *E*[...| *S*=*s*=*s* denotes the expected value of the expression in square brackets, conditioned on state *s* being the current state. The expected value is the weighted average of all possible values of a random variable, specifically referring to the expected cumulative value of a series of rewards and states in the future under state *s*. *r* represents the immediate reward obtained after performing an action in state *s*. The terms *Yr+1*, *yr+2,...* represent the values associated with states or actions at various time steps (*t+1*, *t+2*,...) after state *s*. The symbol γ is used universally to represent the evaluation values corresponding to these subsequent states or actions. The recursive process is derived from the definition of the value function:

$$V(s) = E[r_t + \gamma V(s_{t+1}) \mid s_t = s] \tag{4}$$

In Eq. (4), *V* (*s*) refers to the state value function, representing the expected return (i.e., cumulative sum of future rewards) that can be obtained by following a certain strategy in state *s*; *Q*(*s*, *a*) is the action value function, representing the expected return obtainable by performing action *a* in state *s* and then following a specific strategy; *r* denotes the immediate reward received after executing action *a* in state *s*; *S'* or *s*+1 represents the transition from the current state *s* to a new state, typically due to the execution of action *a*; *E*[...] is the expectation operator, used to calculate the average or expected value of a random variable, typically to determine expected returns in RL; γ is the discount factor, a value between 0 and 1, used to discount future rewards when calculating expected returns. It reflects the relative importance of future rewards compared to current rewards.

The update of the value function and gradient calculation are dynamically updated, with the formula design requiring a basic function value to estimate the action state. The effectiveness of RL algorithms is evident in this calculation process. Even if the target data sample size is large, the output probability distribution can be accurately determined in the experience pool. The determination formula for network parameters and data update loss is as follows:

$$L(\theta) = \tfrac{1}{m} \sum_{i=1}^{m} (y_i - Q(s_i, a_i; \theta))^2 \tag{5}$$

$$y_i = r_i, d_t = True \tag{6}$$

$$y_i = r_i + y \max_a, Q(s_i, a, \theta) d_t = False \tag{7}$$

In this context, *L*(θ) is the loss function, commonly used to measure the discrepancy between the predicted and actual values of a model. In RL, it evaluates the difference between the estimated Q-value under the current strategy and the actual Q-value obtained via sampling. θ represents the network parameters, optimized during training by minimizing the loss function. *Q*(*s, a;* θ) is the action value function, representing the expected return from executing action *a* in a given state *s*, with the strategy corresponding to the current parameter θ. Here, θ is explicitly included in the Q-function to indicate that it is parameterized, meaning the form of the Q-function depends on the network parameters. y_i is a target value used to guide the update of the Q-function during training. It is calculated based on the current immediate reward *r*, the Q-value for the optimal action in the next state *s*' (i.e. $y\max_{a'} Q(s, a'; \theta)$), and a logical variable *d* indicating whether the termination state has been reached. The target data parameters θ, obtained via iterative loss updates, are defined by the following optimization formula for RL to express the expected reward across all states:

$$Q(s_t, a_t) = R[r_t + u r_{t+1} + u^2 r_{t+2} + \ldots \mid s_t, a_t] \tag{8}$$

In Eq. (8), *Q*(*t*, *a*) is the action value function, also known as the Q-function, representing the expected reward (or "value") achievable through a particular strategy after performing action *a* in a given state *t*. *S* is the current state; *a* is the current action; *r* is the immediate reward; γ is the discount factor; *s* is the next state. This formula reflects the basic RL principle: selecting the optimal action in a given state to maximize the sum of future discounted rewards. By combining a simple gradient strategy to optimize the function and learning mechanisms to increase matching probability, the corresponding formula scheme is selected through the policy gradient function in the training data, and the learned parameters are updated according to the reward mechanism. The formula for defining a strategy is as follows:

$$h(s_t, a_t, o) = o^T(s_t, a_t) \tag{9}$$

In Eq. (9), $h(s, a, o)$ is a function that takes three parameters: state s, action a, and a fixed value o, which may also represent a more complex set of constants or parameters. The output of this function is often associated with a specific property or response of the system or environment resulting from performing action a in a given state s. Within the context of RL, this function typically relates to evaluating the performance of a particular strategy. The term state refers to the configuration or condition of a system or environment at a specific moment. In RL, the agent observes the current state to decide on the next action. An action is a behavior or decision that an intelligent agent can execute while in a given state, with the primary objective being to maximize future rewards or achieve a predefined goal. In the equation, the term refers to the "target state object," a key component that interacts dynamically with the function to influence the optimization process. In Eq. (9), the term (s_t, a_t) represents the target state object. The random strategy can be expressed Eq. (10).

$$\pi(a \mid s; o) = \frac{e^h(s_t, b, o)}{\sum_b e^h(s_t, b, o)} \tag{10}$$

In Eq. (10), S_t represents the state at time point t; b is a parameter associated with the function, the specific meaning of which depends on the function's context; o acts as a placeholder or a fixed constant/parameter, and in this context, it might not directly contribute to the calculation of the e^h function. Instead, it could represent a simplified expression or set of variables. The policy function $\pi(a \mid s; \theta)$ is a common representation of a policy function, where π represents the policy, a is the action chosen in state s, and θ represents the parameters of the policy function. The summation symbol $\sum$ indicates the aggregation of elements within a sequence or set. The function $e^h(s, b, o)$ is associated with summation operations, with s, b, and o retaining the same definitions as previously outlined. The overarching aim is to incorporate as many relevant functions as possible into training and learning processes to achieve a comprehensive improvement in strategy performance. To further analyze factors influencing the performance of water-based coatings, RL is applied to object detection. Traditional object detection methods often rely on allocation algorithms to determine optimal solutions based on time complexity. While these methods are relatively easy to execute, they frequently suffer from accuracy loss, particularly in scenarios involving large-scale data processing. As algorithmic complexity escalates with increased data volumes, significant computational resources are consumed, making model optimization more challenging. In contrast, RL is capable of sustaining stable monitoring calculations even with dynamic and large datasets. It offers distinct advantages in both efficiency and accuracy. The RL structure and object detection process are outlined in Fig. 2.

As illustrated in Fig. 2, detection performance is maintained while filters are employed to enhance the detection speed of the RL model. The system integrates multi-objective state estimation, allocation matrices, and neural network algorithms to optimize overall application effectiveness.

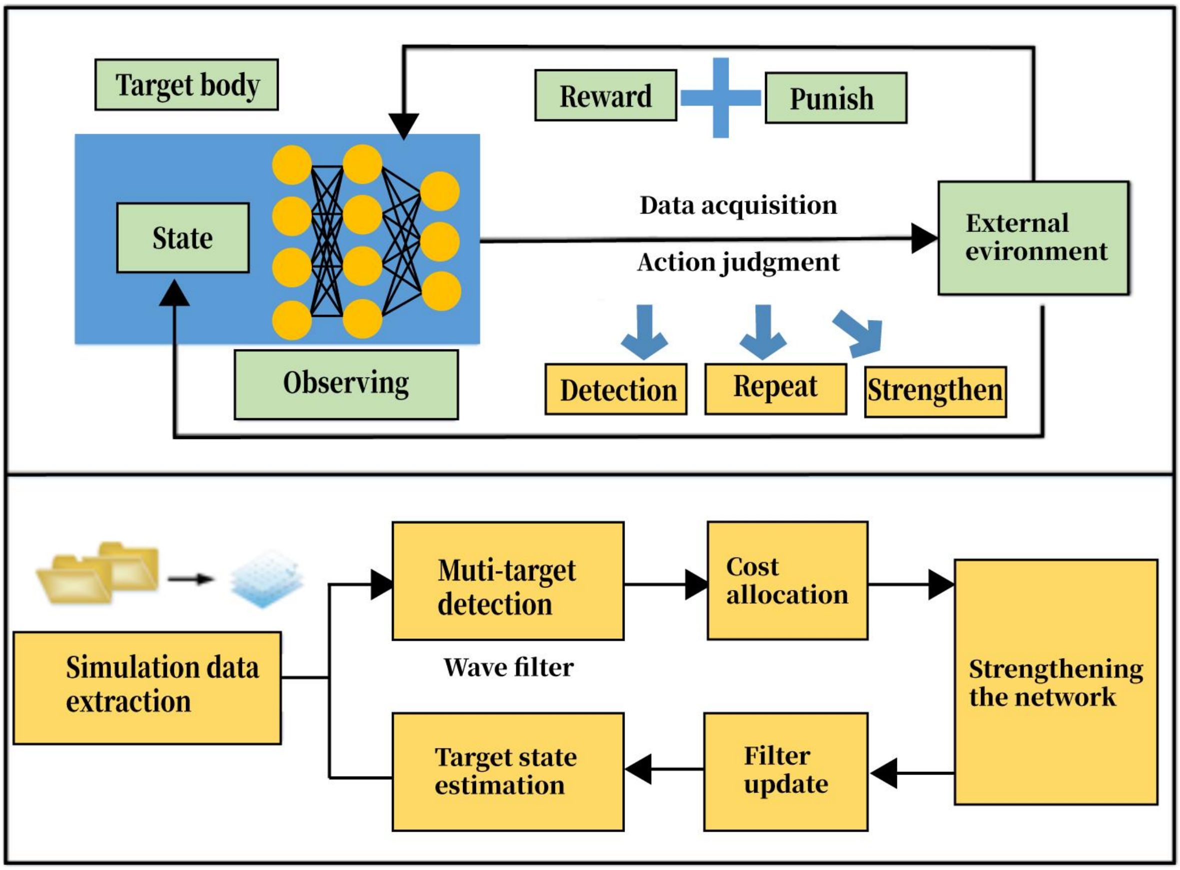

Fig. 2. RL structure and object detection process

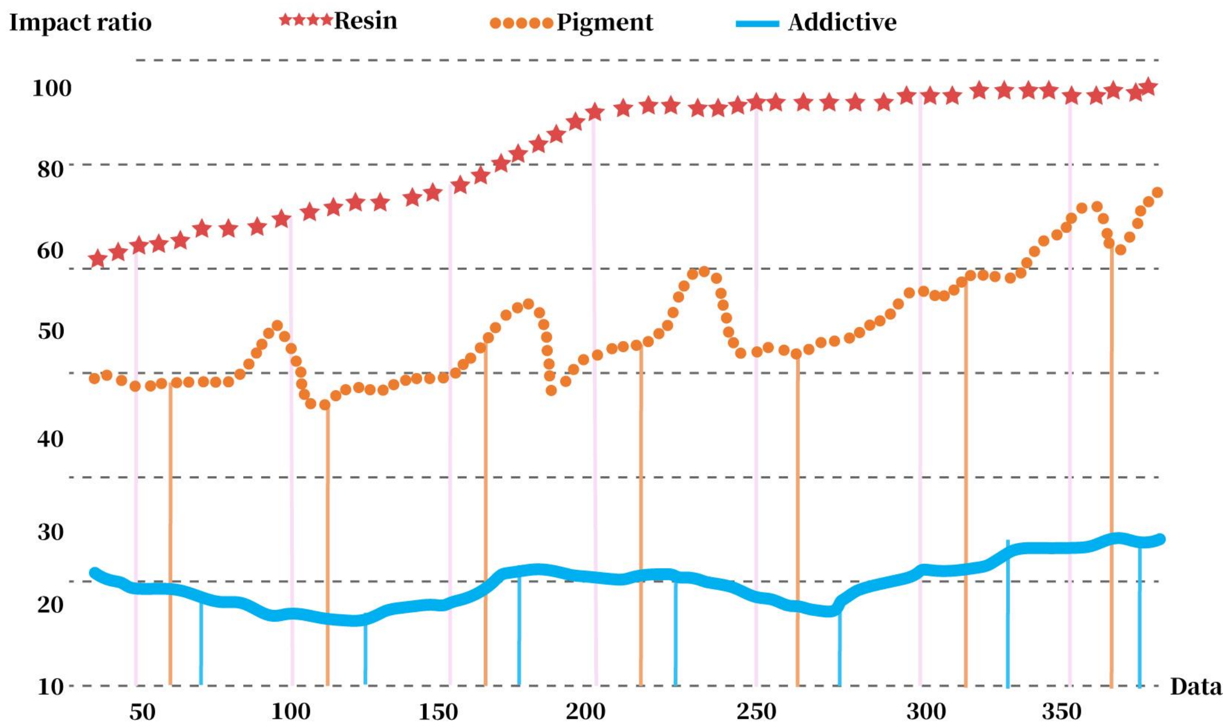

Fig. 3. Comparison of influence coefficients

B. Influence of data-driven additives on the performance of water-based coatings

Currently, most coatings used for materials are solvent based, which poses significant environmental issues. These coatings not only harm the surrounding environment, but also pose a safety hazard to human health. In contrast, water-based coatings have enormous potential for development as a sustainable alternative to the protective coating industry. Although water-based coatings have high environmental compatibility, there are certain limitations in terms of corrosion resistance, high temperature resistance, water resistance, and other properties, which limit their wider application and adoption. Waterborne coatings are usually composed of lotion resins, additives, pigments and other substances. The formulations of these coatings vary depending on specific performance requirements. Due to the complex internal composition of water-based coatings, optimizing component combinations to achieve high-performance materials remains a key research challenge. The traditional optimization methods for water-based coating formulations typically rely on methods such as linear regression and nonlinear regression to establish mathematical models related to coating performance indicators. Then, these models are used to optimize the formula based on the calculated data. However, the complex dynamic relationship between a large amount of data and water-based coating constraints results in low accuracy of the model and limited applicability in scheme generation. To solve this problem, data analysis methods are used to determine the influence coefficients of key components such as resin, pigment, and additives on the performance of water-based coatings. Fig. 3 illustrates these relationships.

As shown in Fig. 3, the structure of resin material plays a crucial role in the performance of water-based coatings. The film-forming speed and method of these resins significantly affect the coverage and adhesion performance of coatings, resulting in a significant increase in their coefficient of influence. Considering the use of water molecules as dispersion media, selecting low-density resin materials is crucial to ensure stable film formation and enhance hydrophilicity in aqueous environments. The second major influencing factor is pigments. Pigments contribute to the auxiliary functions of water-based coatings, such as improving performance when interacting with substrate surfaces. However, the coefficient of influence of pigments is variable as it depends on the specific functional requirements of the coating. In addition to resins and pigments, the drying rate also plays a crucial role in the film-forming process of water-based coatings. As the temperature increases, the viscosity of the coating decreases, affecting the drying process. The drying speed in turn affects the film-forming density, which is a key performance indicator in coating production. In order to optimize the formulation of water-based coatings, various influencing factors were integrated into the RL optimization model. By adjusting dynamic performance and other indicators to control parameter variables, the quality of the formulation can be evaluated. The dynamic state is incorporated into the calculation as a reward function, and the formula generation process is refined through regular online learning and gradient updates. This method achieves parameter customization and automatic optimization, effectively solving the challenges of formula design. The problem of parameter customization and optimization is represented by inequality constraints as follows:

$$Q = \begin{cases} \delta(X) \le \Omega_1 \\ t_r(X)\Omega_2 \\ t_s(X) \le \Omega_3 \end{cases} \tag{11}$$

In Eq. (11), Ω represents a constraint value or upper bound, which limits the threshold that the result of the variable X, processed by a specific function, should not exceed. This value Ω dynamically adjusts based on specific conditions. X denotes a parameter variable subject to optimization, while $\delta(X)$ specifies the permissible range for the parameter variables and represents the constraint values imposed during optimization. The algorithm incorporates these elements into the reward function, defined as follows:

$$R(X)=\frac{1}{\delta(X)^2+t_r(X)^2+t_s(X)^2} \tag{12}$$

$$X=a\cdot\nabla X+\sigma \tag{13}$$

Here, X is a vector or matrix representing the result of the parameter variables undergoing optimization. The value of X is calculated by multiplying the constant a with the matrix ∇X and adding another constant term o. The constant a is a scalar used to scale the matrix ∇X in the multiplication process. The matrix ∇X represents a transformation or mapping function that converts an input vector (or matrix) into an output vector (or matrix). To enhance the effectiveness of global parameter optimization, a noise removal mechanism is introduced during the custom training process of the RL model. This mechanism improves the reliability of the calculated results. The formula for incorporating the noise removal steps is defined as follows:

$$\frac{Q(s)}{U(s)}=\frac{K_1}{(T_1s+1)}\cdot e \tag{14}$$

$$\frac{H(s)}{U(s)}=\frac{K_1K_2}{(T_1s+1)(T_2s+2)}\cdot e^{(-r+s)} \tag{15}$$

Here, $Q(s)$ is a function or value associated with state s, usually used in RL to represent the optimal action-value function in that state. Specifically, it denotes the maximum expected reward achievable by performing a particular action in state s. While the exact definition may vary based on the context, it is generally understood as a quantity closely tied to s, which is an optimization objective or evaluation metric. $U(s, (T_1s+1))$ is the value of a certain quantity U in state s at a specific time series $Ts+1$. It reflects a cumulative utility, cost, or error that evolves over time or as iterations progress. Here, T_1s denotes a current time step or iteration, while T_1s+1 corresponds to the next.

$H(s)$ represents a performance metric or condition associated with state s. K_1 and K_2 are constants independent of s or the time step Ts. These coefficients play a critical role in the calculation of $U(s, (T_1s+1))$ or related terms, typically used to adjust the relative importance of various components. The term (T_1s+1) (T_2s+2) is employed in the computation of $U(s, (T_1s+1))$, incorporating information from two distinct time steps, T_1s+1 and T_2s+2. Based on the provided formula, the RL optimization model produces parameter adjustments after training. When integrated with diverse performance requirements for water-based coatings, the model is capable of autonomously generating an optimal formulation design scheme.

III. Results and discussion

A. Analysis of research results on water-based coating formulation performance using RL-based object detection

In the formulation design system of water-based coatings, the first step is to understand the expected application environment and usage requirements. This includes identifying the main factors that affect performance. According to various usage conditions and environmental requirements, this formula contains different structures, film-forming substances, pigments, additives, and other components. These are customized through specific film-forming methods to demonstrate the required characteristics of water-based coatings. This study emphasizes the corrosion resistance and high temperature resistance of water-based coatings. Modified acrylic acid is the main resin material, and other materials have been added to develop the formula. RL is used to optimize the detection process of other parameters by integrating influencing factors as control variables into the training model. Randomly implement various material combination strategies and introduce reward and punishment functions at each training stage. This method increases the likelihood of successful formulation design for water-based coatings. To evaluate the reliability of the proposed method, a comparison was made between traditional data detection models and RL based detection models. The comparison results are shown in Fig. 4.

As shown in Fig. 4, the performance of water-based coatings is significantly influenced by the material ratio data. With the inclusion of a large volume of data, the detection efficiency of conventional data detection models decreases substantially. In contrast, the RL algorithm used in this study demonstrates superior efficiency in detecting key objects, providing reliable data support for the design of water-based coating formulations. The resin, constructed with organic modified acrylic acid, was degreased and cleaned, with the coating's surface viscosity maintained at room temperature (25°C) for 12 to 16 seconds. The viscosity content of the coating exceeded 27%, ensuring an adequate coating thickness.

To evaluate the effectiveness of RL technology, we conducted comparative experiments with traditional data detection models. As shown in Fig. 4, with the increase of data volume, traditional detection models are difficult to handle complex formula data [24], resulting in a significant decrease in detection efficiency. In contrast, due to its advanced decision-making ability and adaptability, RL algorithms exhibit higher efficiency and accuracy in detecting key objects. These findings indicate that RL technology has great potential in optimizing water-based coating formulations. Further analysis was conducted on the molecular structure of organic modified acrylic resin and its impact on coating properties [25]. As shown in Fig. 5, a mixture of hydroxyl silicone oil (molecular weight 300-5000) and methyl silicone oil (molecular weight 0-2000) is mixed with methyl acrylic acid to produce a silicone modified acrylic resin coating with a unique chemical structure. This modified resin coating retains the beneficial properties of acrylic resin, while improving its weather resistance, wear resistance, and stain resistance by incorporating silicone oil. Specifically, the hydroxyl groups in hydroxyl silicone oil react with carboxyl or ester groups on the acrylic molecular chain to form stable chemical bonds, enhancing the cohesion and adhesion of the coating [26]. Methyl silicone oil has low surface tension and excellent lubricity, which can reduce the surface energy of coatings and improve their water resistance and stain resistance. The chemical structure of the two molecules and the resulting coating are shown in Fig. 5.

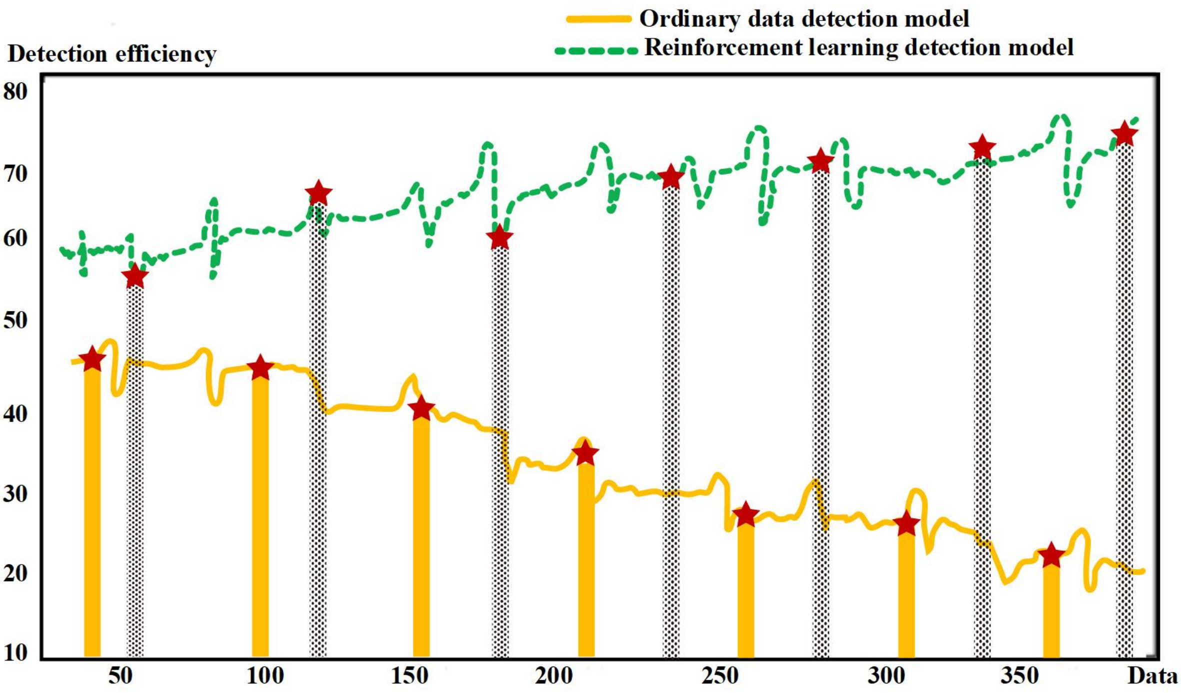

Fig. 4. Comparison of testing efficiency between two technologies

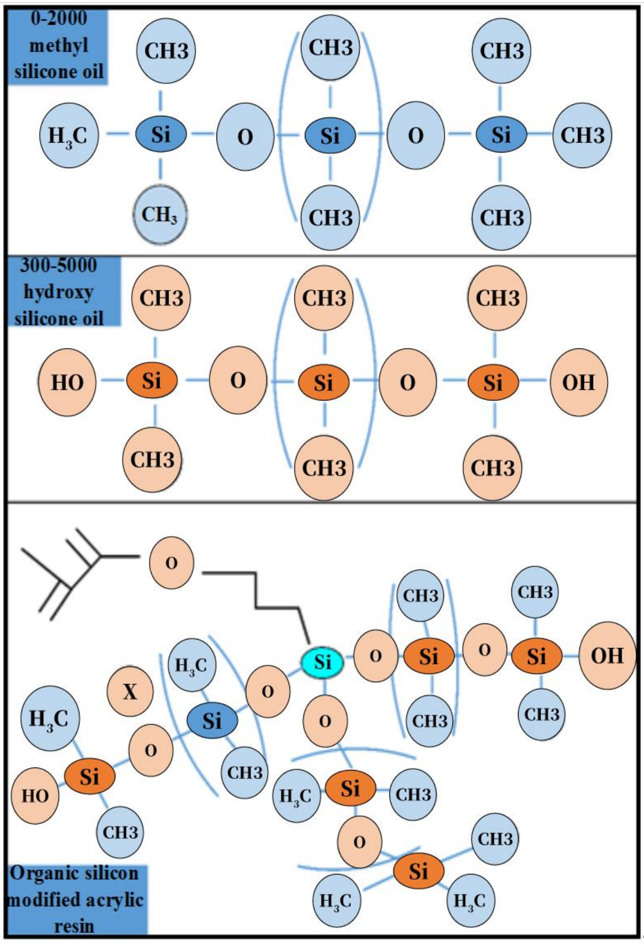

Fig. 5. Chemical molecular structure of coatings

The experiment revealed that the total proportion of silicone oil in the resin should not exceed 15%, as higher concentrations lead to solvent incompatibility. Additionally, increasing the proportion of methacrylic acid results in smaller particle sizes for the coating, and the finished product tends to exhibit a semi-transparent color. These findings establish the foundation for preparing water-based coatings with high stability, providing a basis for adjusting the formulation proportions in subsequent experiments. Building on the initial preparation, the following subsection focuses on exploring the impact of the mixing ratio of water, pigments, additives, and other substances on the corrosion resistance and high-temperature resistance of water-based coatings.

In order to more intuitively demonstrate the differences in data processing efficiency and accuracy between RL technology and traditional data detection models, we conducted a detailed statistical analysis and drew the following chart:

TABLE I
COMPARISON OF DATA PROCESSING EFFICIENCY AND ACCURACY

Data Volume (GB)	Traditional Detection Model Processing Time (hours)	RL Algorithm Processing Time (hours)	Accuracy of Traditional Detection Models (%)	Accuracy of RL Algorithm (%)
1	2.5	1.2	90	95
5	12.8	3.5	85	98
10	28.6	6.0	80	99

TABLE II
EFFECT OF FORMULA PROPORTION ON CORROSION RESISTANCE AND HIGH TEMPERATURE RESISTANCE OF COATINGS

Water Content (%)	Pigment Content (%)	Additive Content (%)	Corrosion Resistance (Rating)	High Temperature Resistance (℃)
10	15	5	4	120
15	10	8	5	130
20	5	10	3	110

From Table I, it can be seen that as the amount of data increases, the processing time of traditional detection models significantly increases, while the processing time of RL algorithms increases relatively slowly. At the same time, RL algorithms have shown significant advantages in accuracy, especially in the case of large amounts of data.

On the basis of preliminary preparation of organosilicon modified acrylic resin coatings, we further investigated the influence of the mixing ratio of water, pigments, additives, and other substances on the corrosion resistance and high temperature resistance of water-based coatings. The following is a chart display of some experimental results:

From Table II, it can be seen that the content of water, pigments, and additives has a significant impact on the corrosion resistance and high temperature resistance of the coating. By adjusting the formula ratio, we can obtain water-based coatings with excellent corrosion resistance and high temperature resistance.

B. Research results analysis of water-based coating formulation design model based on RL optimization

The formulation design of water-based coatings involves various components, including resins, polyvinyl chloride (PVC) values, pigments, and additives. The detection of resin properties requires a reliable and high-precision system. In this study, RL algorithm was used to detect targets and ensure the reliability and accuracy of the results. The selection of additives is crucial, as different additives can cause significant performance changes in the same resin matrix. In order to achieve high anti-corrosion performance in water-based coatings, it is necessary to reduce the PVC value in the formula and consider the compatibility between additives and pigments. In addition, the interaction between additives can affect the overall state of the coating. To enhance compatibility, material wetting agents and balancing agents are added to the formula. In traditional formula design, changing one material can affect the concentration of other ingredients. The formula design process was optimized based on RL, and the performance of water-based coatings was improved by automatically adjusting parameters during the reinforcement training process. Fig. 6 shows the changes in corrosion resistance and high temperature resistance of water-based coatings before and after RL optimization.

Fig. 6 shows that compared to the pre optimized formula, the water-based coating formula optimized through RL training exhibits significantly enhanced corrosion resistance and high temperature resistance. In pigment selection, in addition to the stability of water-based coatings, factors such as dispersibility and coverage must also be considered. Emphasis should be placed on extending the service life of these coatings during use.

In order to further verify the experimental results, the service life of water-based coatings was tested, as shown in Fig. 7. Fig. 7 shows that compared to unoptimized coatings, RL optimized water-based coating formulations have a longer service life, especially in high temperature and corrosive marine environments. This highlights that the scientific and systematic formulation design of water-based coatings directly affects the effectiveness and durability of coating coverage.

Table III shows a comparison of the performance of water-based coatings before and after optimization. The optimized formula (B) is significantly better than the unoptimized formula (A) in terms of corrosion resistance, high temperature resistance, and viscosity stability. Specifically, the corrosion resistance in the salt spray test increased from 300 hours in formula A to 600 hours in formula B, while the high temperature resistance increased from 200 hours to 400 hours. In addition, the viscosity stability increased from 10-12 seconds to 12-16 seconds, and the coverage rate of formulation B increased from 85% to 90%, indicating better uniformity and adhesion of the coating.

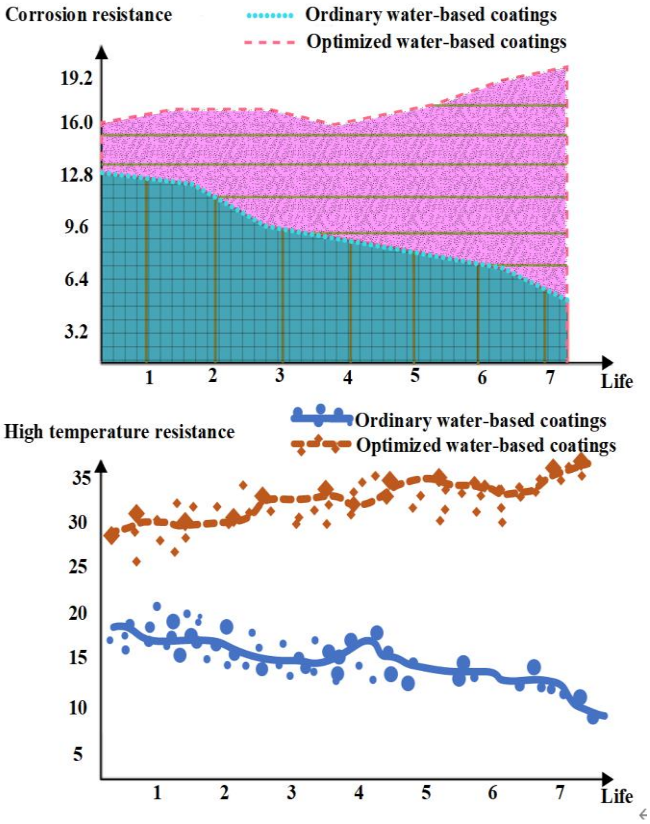

Fig. 6. Comparison of anti-corrosion and high-temperature resistance

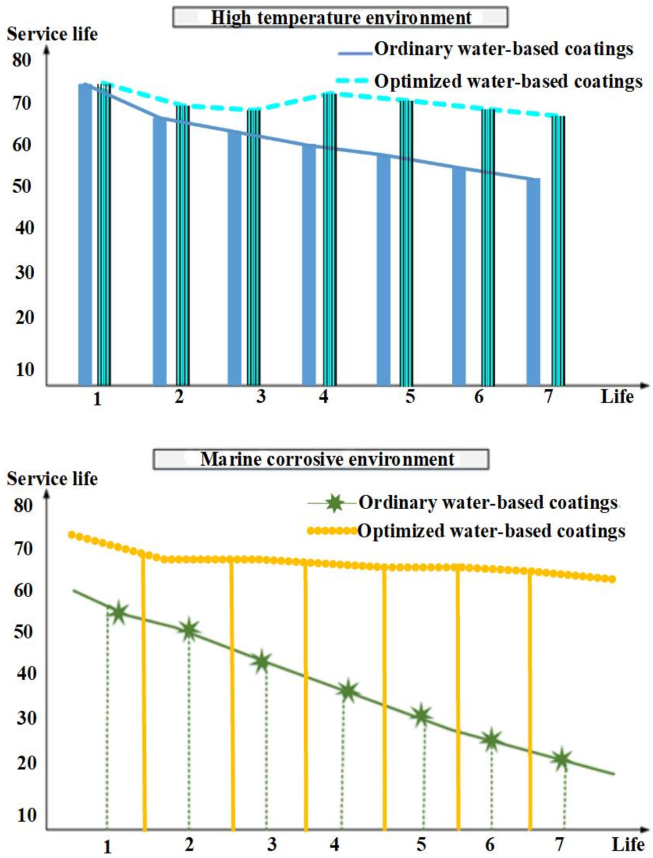

Fig. 7. Service life testing of two water-based coatings

TABLE III
COMPARISON OF WATER-BASED COATING PERFORMANCE BEFORE AND AFTER OPTIMIZATION

Formulation ID	Resin Type	Pigment Ratio	Additive Type	Corrosion Resistance (Salt Spray Test, Hours)	High-Temperature Resistance (60°C, Hours)	Viscosity Stability (25°C, Seconds)	Coverage Rate (%)
Formulation A (Unoptimized)	Modified Acrylic Resin	25%	Surfactant	300	200	10-12	85%
Formulation B (Optimized)	Modified Acrylic Resin	20%	Surfactant	600	400	12-16	90%

TABLE IV
IMPACT OF PARAMETER ADJUSTMENTS ON COATING PERFORMANCE DURING OPTIMIZATION

Formulation ID	Resin Type	Pigment Ratio	Additive Type	Initial Viscosity (mPa·s)	Additive Ratio (%)	Optimized Viscosity (mPa·s)	Corrosion Resistance Improvement (Hours)	High-Temperature Resistance Improvement (Hours)
Formulation A (Unoptimized)	Modified Acrylic Resin	25%	Surfactant	500	5%	450	300	200
Formulation B (Optimized)	Modified Acrylic Resin	20%	Surfactant	480	7%	460	600	400

TABLE V
COMPARISON OF WATER-BASED COATING PERFORMANCE UNDER DIFFERENT OPTIMIZATION METHODS

Optimization Method	Corrosion Resistance (Salt Spray Test, Hours)	High-Temperature Resistance (180°C, Hours)	Surface Viscosity Stability (25°C, Seconds)	Coating Thickness Uniformity (%)	Coverage Rate (cm^2)
Traditional Experimental Design Optimization	120	100	10	85	120
RL Optimization	200	160	14	95	150
Patti et al. [27]	130	110	12	88	125
Zhong et al. [28]	140	120	13	90	130

Table IV further illustrates the impact of parameter adjustments on coating performance during the optimization process. By increasing the proportion of surfactants from 5% to 7%, the viscosity of the coating decreased from 500 mPa · s to 480 mPa · s, indicating an improvement in flowability after viscosity optimization. Regarding corrosion resistance, Formula B has improved by 300 hours compared to Formula A, while its high temperature resistance has increased by 200 hours. The proportion of resin and additives was optimized through RL algorithm, ensuring the long-term stability and reliability of the coating in practical applications.

The research results in Tables III and IV indicate that formula adjustment and automatic optimization using RL significantly improve the performance of water-based coatings. It is worth noting that its corrosion resistance, high temperature resistance, and viscosity stability exceed those of unoptimized formulations. These results validate the effectiveness and potential of RL in optimizing water-based coating formulations.

C. Discussion

This study optimized the formulation design of water-based coatings using RL and introduced an optimization model that combines neural networks with RL technology. In addition, applying the orthogonal method to the formulation process significantly improved the performance of the coating in terms of corrosion resistance, high temperature resistance, and coverage strength. To verify the effectiveness of the proposed method, comparative experiments were conducted on optimized and non optimized coatings under various environmental conditions. The results indicate that the optimized coating outperforms traditional designed coatings in terms of performance and service life. Table V shows a comparison of the performance of water-based coatings using different optimization methods. The RL optimization method proposed in this study has substantial advantages over traditional experimental design methods in multiple performance indicators. It is worth noting that the corrosion resistance measured through salt spray testing has significantly improved, increasing from 120 hours using traditional experimental design methods to 200 hours using RL optimization, an increase of 66.7%. Similarly, the high-temperature resistance of the optimized coating increased by 60% from 100 hours to 160 hours. The surface viscosity stability also showed significant improvement, rising from 10 seconds to 14 seconds, while the coating thickness uniformity increased from 85% to 95%. These advances indicate that RL optimization not only improves the physical and chemical stability of coatings, but also ensures better application uniformity.

This analysis examines the advantages, innovations, and limitations of this study in comparison to existing research. Firstly, regarding the design of traditional coating formulations, Patti et al. studied the basic optimization strategies for water-based coatings. They used experimental design methods to study the interaction between additives and resins and their impact on coating performance [27]. Their findings indicate that adjusting the proportion of additives can improve the durability of coatings. However, their method heavily relies on a large amount of experimental data, which makes it both expensive and inefficient. In contrast, this study utilized RL to automatically adjust parameter ratios. By combining reward and punishment functions, the need for large-scale experimental validation is minimized to the greatest extent possible, significantly improving optimization efficiency. Secondly, in terms of improving corrosion resistance, Zhong et al. explored organosilicon modification technology, synthesizing organosilicon modified resins through copolymerization of hydroxy silicone oil and methyl methacrylate. Their work has demonstrated the effectiveness of this method in improving durability and adhesion [28]. However, their research lacks systematic exploration of component ratios and mainly relies on empirical adjustments. Based on their findings, this study applied RL algorithm to automatically adjust the proportion of silicone and optimize the overall formula. The results indicate that while solving the solvent incompatibility problem related to high silicone content, the corrosion resistance has been significantly improved. Thirdly, in the optimization of high temperature resistance performance, Zhang et al. used molecular dynamics simulations to study the relationship between molecular structure and coating properties. They found that viscosity stability is a key factor affecting the service life of coatings under high temperature conditions [29]. This study further validated these findings by optimizing key formulation parameters, such as the ratio of methyl silicone oil to methyl methacrylate, using RL algorithm. The optimized water-based coating exhibits better viscosity stability (12-16 seconds) and consistent coating thickness at 25 ° C, thereby improving high temperature resistance. Despite its advantages, this study also has certain limitations. For example, the initial parameter settings of RL models can significantly affect the optimization results, and further research is needed to determine the optimal initial parameters. In addition, the training process of RL models heavily relies on high-quality input data. Challenges related to insufficient data or noise remain unresolved, which may limit the generalizability of the model in real-world applications.

In summary, this study effectively combines RL algorithm with data-driven methods, providing a systematic and effective approach for optimizing water-based coatings. Compared to previous studies, its performance has significantly improved. Although there are still areas for improvement, these findings provide valuable insights and technical support for the design of high-performance water-based coatings.

IV. CONCLUSIONS

This study marks a pioneering effort in applying RL algorithms to optimize water-based coating formulations. By utilizing intelligent design models, this study systematically parameterized and optimized the composition of these coatings. Traditional methods such as regression analysis and traditional mathematical modeling techniques have been used for formula adjustment. However, these methods are often insufficient to address the inherent complexity of coating composition and the non-linear relationship between formulation parameters and performance results. Therefore, their accuracy is limited, and the resulting formulas are often not optimal. The RL based optimization model introduced in this study combines innovative mechanisms, including reward functions, dynamic state updates, and regular online learning. These features enable the model to adaptively adjust and automatically optimize within a complex recipe space, significantly improving accuracy and efficiency. By applying the RL algorithm, the model dynamically determines the influence coefficients of each component in the formula. This method has achieved substantial improvements in key performance indicators such as corrosion resistance and high-temperature durability. This study further combines neural networks with reinforcement learning, enhancing the model's adaptability and predictive ability when dealing with complex datasets. The experimental results highlight the excellent performance, stability, and durability of RL optimized water-based coatings, especially in challenging environments under high temperature and corrosive conditions. These findings not only provide a new way to improve the performance of water-based coatings, but also lay a solid theoretical foundation for the advancement of future intelligent coating designs. The unique contribution of this research lies in its successful integration of artificial intelligence technology into the coatings industry. This advancement simplifies and improves the automation and efficiency of water-based coating formulation optimization, representing a transformative step in intelligent and high-performance design of coatings. This work has enormous potential for practical application and widespread adoption, marking a key milestone in the development of intelligent material design.

REFERENCES

[1] Yunusaliev, M. (2023) The value of watercolor and gouache material in the technique of water-based paint. Science and Innovation 2(C9), 5-10.

[2] Nasirzadeh, M., Yahyaei, H., & Mohseni, M. (2023) Effects of inorganic fillers on the performance of the water-based intumescent fire-retardant coating. *Fire and Materials* 47(1), 51-61.

[3] Mahendru, P., Tembely, M., & Dolatabadi, A. (2023) Artificial intelligence models for analyzing thermally sprayed functional coatings. *Journal of Thermal Spray Technology* 32(2), 388-400.

[4] Gao, F., Liu, Y., & Jiao, C. (2023) Fluorine-phosphate copolymerization waterborne acrylic resin coating with enhanced anticorrosive performance. *Journal of Polymer Science* 61(21), 2677-2687.

[5] Mucci, V. L., Hormaiztegui, M. E. V., & Amalvy, J. I. (2024) Formulation, structure and properties of waterborne polyurethane coatings: A brief review. *Journal of Adhesion Science and Technology* 38(4), 489-516.

[6] Nakajima, H., Matsumoto, N., & Ogura, T. (2024) Hidden correlation between rheological dynamics and crack formation in water-based slurry. *Journal of the European Ceramic Society* 44(6), 4141-4149.

[7] Li, B., Njuko, D., & Meng, M. (2022) Designing smart microcapsules with natural polyelectrolytes to improve self-healing performance for water-based polyurethane coatings. *ACS Applied Materials & Interfaces* 14(47), 53370-53379.

[8] Xu, G., Nigmatullin, R., & Koev, T. T. (2022) Octylamine-modified cellulose nanocrystal-enhanced stabilization of Pickering emulsions for self-healing composite coatings. *ACS Applied Materials & Interfaces* 14(10), 12722-12733.

[9] Park, B. K., Ahn, Y. K., & Kwon, Y. K. (2022) Developing eco-friendly ceramic composite separator with competitive electrochemical properties using water-based polymer binder for lithium-ion batteries. *International Journal of Energy Research* 46(2), 1398-1406.

[10] Arya, R. K., Giribabu, D., & Sharma, J. (2022) Effect of surfactant on drying and rheology of poly (vinyl alcohol)–water coatings. Drying Technology 40(10), 2140-2152.

[11] An, C., Huang, B., & Jiang, J. (2023) Design and synthesis of a water-based Nano delivery pesticide system for improved efficacy and safety. ACS Nano 18(1), 662-679.

[12] Kang, J., Chang, N., & Wang, F. (2022) Selective solid-phase microextraction of polycyclic aromatic hydrocarbons in water based on oriented phosphorus-containing titanium oxide nanofibers grown on titanium support prior to high-performance liquid chromatography with ultraviolet detection. Journal of Separation Science 45(6), 1273-1281.

[13] Sharma, H., Sharma, B., & Raghuvanshi, S. (2024) Development of acrylic based antimicrobial water-based ink using curcumin for active

packaging applications. Polymer Engineering & Science 64(1), 350-362.
[14] Wang, Q., Zeng, J., & Shao, K. (2023) Wear-resistant/antifouling integrated coating protection technology of polyethylene cable. Materialwissenschaft und Werkstofftechnik 54(7), 850-856.
[15] Gan, Q., Qin, N., & Guo, H. (2024) Consolidating the vulnerable interphase of Ni-rich layered cathode by multifunctional water-based binder. ACS Energy Letters 9, 1562-1571.
[16] Susanta Maity, Palky Handique, and Swatilekha Nag, "Unsteady Thin Liquid Film Flow on a Porous Stretching Cylinder," IAENG International Journal of Applied Mathematics, vol. 52, no.1, pp213-218, 2022.
[17] G. Suganya, and R. Senthamarai, "Impact of Awareness on the Dynamics of Pest Control in Coconut Trees - A Mathematical Model," Engineering Letters, vol. 30, no.4, pp1199-1209, 2022.
[18] Zuo H, Wang H, Zhou Y. (2023). Improving Extended Tree Augmented Naıve Classifier with GAN. IMECS,25-30.
[19] Huo, Y., Zhou, J., & Xu, Z. (2023) Preparation and properties of water-based photochromic polyurethane coatings containing Er (III)-doped WO3. Surface and Interface Analysis 55(4), 288-295.
[20] Kim, T., Samuel, E., & Seol, J. (2022) High-energy-density supercapacitors using supersonically sprayed water-based precursors comprising cobalt iron oxide and reduced graphene oxide nanosheets. International Journal of Energy Research 46(10), 14305-14317.
[21] Boinovich, L. B., Emelyanenko, A. M., & Emelyanenko, K. A. (2024) Effect of water adsorption on lubricating film stability in slippery coatings. Langmuir 40(3), 1633-1645.
[22] He, H., Yang, T., & Liu, Y. (2023) Self-stratifying coating with robust and durable superhydrophobic underlayer constructed by amphiphilic Janus nanoparticles. Chemistry of Materials 35(15), 6083-6096.
[23] Cheng, W., Qiao, T., & Xue, C. (2022) Computation of binding energy of MCS and GO-grafted MCS with waterborne epoxy resin using density functional theory method: Investigating the corrosion resistance of the composite coatings. ACS Omega 7(44), 40374-40386.
[24] Liu, W., Huang, S., Ye, C., Jia, L., Kang, Y., Sha, J., Chen, B. Q., Wu, Y., & Xiong, H. (2023) Progress in Nb-Si ultra-high temperature structural materials: A review. Journal of Materials Science & Technology 149, 127-153.
[25] Li, K., Ji, C., Bai, S., Jiang, B., & Pan, F. (2023) Selective laser melting of magnesium alloys: Necessity, formability, performance, optimization and applications. Journal of Materials Science & Technology 154(1), 65-93.
[26] Liu, C., Zhang, Q., & Zou, Z. (2023) Recent advances in designing ZnIn2S4-based heterostructured photocatalysts for hydrogen evolution. Journal of Materials Science & Technology 139(1), 167-188.
[27] Patti, A., & Acierno, D. (2023) Structure-property relationships of waterborne polyurethane (WPU) in aqueous formulations. Journal of Vinyl and Additive Technology 29(4), 589-606.
[28] Zhong, S., Li, J., & Cai, Y. et al. (2019) Novel surfactant-free waterborne acrylic-silicone modified alkyd hybrid resin coatings containing nano-silica for the corrosion protection of carbon steel. Polymer-Plastics Technology and Materials 58(8), 866-878.
[29] Zhang, X., Yang, Z., & Deng, Y. (2024) Molecular dynamics study on wear resistance of high entropy alloy coatings considering the effect of temperature. Materials 17(16), 3911.