

# Application of Artificial Intelligence Optimization Algorithms Coupled with Machine Learning in Water Treatment Process Optimization and Water Quality Prediction

Xinwei Liu\*

*College of Environment and Resources, Taiyuan University of Science and Technology, Taiyuan 030024, China*

*\*Corresponding author: Xinwei Liu.*

---

## Abstract

The increasing scarcity of water resources and the continuous aggravation of water pollution have made the efficient optimization of water treatment processes and the accurate prediction of water quality parameters critical technical supports for ensuring urban and rural water supply safety. Traditional water treatment operation modes generally exhibit limitations such as unstable treatment performance, high energy and chemical consumption, and limited precision in automated control when addressing complex water quality variations and multi-objective synergistic optimization requirements. In recent years, the deep integration of artificial intelligence optimization algorithms with machine learning methods has provided a data-driven new pathway for the intelligent upgrading of water treatment systems. This paper systematically reviews the fundamental principles and application characteristics of typical intelligent optimization algorithms, including Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Harris Hawks Optimization (HHO). It focuses on analyzing the fusion strategies and operational mechanisms of these algorithms with predictive models such as deep neural networks (DNN), Support Vector Machines (SVM), and Random Forests (RF). Furthermore, the paper elaborates on the practical application effectiveness of this technical framework in process parameter optimization, water quality variation forecasting, and multi-objective system scheduling. Finally, future development directions are proposed to address existing challenges, including data reliability, model generalizability, and computational costs, providing theoretical references for related research and engineering implementation.

## Keywords

artificial intelligence optimization algorithms, machine learning, water treatment processes, water quality prediction, multi-objective optimization, coupled models

---

## 1. Introduction

## 1.1 Research Background and Significance

With the rapid advancement of global industrialization and urbanization, the types of pollutants in water bodies continue to increase, and the magnitude of water quality fluctuations has also grown significantly. Traditional water treatment facilities struggle to adapt to dynamically changing influent conditions while simultaneously balancing multiple objectives, including treatment efficiency, operational costs, and energy conservation. At present, the empirical control and single-objective optimization approaches commonly adopted by water treatment plants fail to achieve precise balance among various indicators under complex operating conditions. The problems of slow response and low accuracy are particularly evident when dealing with sudden pollution events. Against this backdrop, the development of efficient and intelligent water treatment optimization and prediction technologies has become an urgent requirement for improving system operational efficiency and ensuring drinking water safety.

Artificial intelligence optimization algorithms, drawing on nature-inspired heuristic rules, possess strong capabilities for searching in high-dimensional spaces and can solve complex constrained problems without relying on precise mathematical models. Machine learning, on the other hand, can extract hidden relationships between water quality and process parameters from historical monitoring data, enabling advance prediction. The coupled application of these two technologies can form an integrated closed-loop system encompassing prediction, optimization, and control. This approach not only enhances process operation efficiency but also enables dynamic forecasting of water quality, which holds significant practical value for promoting the digital and intelligent transformation of the water treatment industry.

## 1.2 Brief Review of Current Research Status

Scholars at home and abroad have conducted extensive explorations into the application of intelligent algorithms and machine learning in the field of water treatment. In the area of process optimization, classical algorithms such as Particle Swarm Optimization and Genetic Algorithm have been widely and maturely applied to parameter tuning in conventional treatment units, including coagulation, aeration, and sedimentation. Emerging algorithms such as Harris Hawks Optimization and Sine Cosine Algorithm, with their superior global search performance, have gradually been adopted in optimization studies for more complex processes such as membrane treatment and advanced oxidation. In the domain of water quality prediction, models including deep neural networks, Support Vector Machines, and Random Forests demonstrate excellent nonlinear fitting capabilities and can achieve high-precision prediction of key indicators such as COD, ammonia nitrogen, and total phosphorus.

The core research focus at present lies in the coupled integration of these two types of technologies. Optimization algorithms are employed to improve the parameters and structures of machine learning models, while the prediction results are used in turn to guide the optimization of process parameters, thereby forming a closed-loop control mechanism. However, existing achievements still face challenges such as inconsistent data quality, poor cross-scenario applicability of models, and high difficulty in engineering implementation. The overall theoretical and application framework still requires further improvement.

## 2. Application of Artificial Intelligence Optimization Algorithms in Water Treatment Process Optimization

### 2.1 Overview of Intelligent Optimization Algorithms

Intelligent optimization algorithms represent a class of global search methods inspired by natural phenomena or biological behaviors. These algorithms do not rely on analytical models of the research object and can effectively address complex engineering problems characterized by nonlinearity, multiple constraints, and multiple objectives. By appropriately balancing global exploration and local exploitation capabilities, such algorithms can rapidly approach optimal solutions within the parameter space. They have been widely applied in various aspects of water treatment, including process parameter adjustment, operational scheduling, and cost control, and have become a mainstream tool for solving complex water treatment optimization problems [1].

## 2.2 Typical Optimization Algorithms and Their Application Characteristics

The Genetic Algorithm simulates the mechanisms of biological genetic evolution and performs global search through selection, crossover, and mutation operations. It exhibits strong robustness and is particularly effective for multi-objective optimization. This algorithm can simultaneously optimize parameters across multiple process units, such as aeration, sedimentation, and filtration, while providing multiple feasible solutions and thereby enhancing the flexibility of system scheduling.

The Particle Swarm Optimization algorithm mimics the foraging behavior of biological populations. With its simple structure and fast convergence speed, it is the most commonly applied method for optimizing parameters such as chemical dosage, reaction time, and aeration intensity. It can quickly determine the optimal operating range and effectively reduce the consumption of chemicals and electricity [2].

The Harris Hawks Optimization algorithm draws on the cooperative predation strategies of hawk populations. It dynamically adjusts its search patterns and demonstrates a strong ability to avoid becoming trapped in local optima. Consequently, it achieves higher optimization accuracy in complex scenarios with numerous constraints, such as membrane separation and advanced oxidation processes.

The Simulated Annealing algorithm is based on the physical principles of solid-state cooling and annealing. It escapes local optima through probabilistic jumps and offers good stability along with a wide range of applicability. It is suitable for optimization tasks involving chemical dosing schemes, energy consumption management, and pipe network distribution.

The Sine Cosine Algorithm uses trigonometric functions as the basis for iteration. Its computational logic is clear and its operational efficiency is high, making it particularly suitable for real-time parameter optimization scenarios that require rapid response.

The Ant Colony Optimization algorithm relies on positive feedback of pheromones to achieve path optimization. It holds unique advantages in combinatorial optimization problems such as chemical dosing path planning, pipe network scheduling, and resource allocation [3].

## 2.3 Multi-Objective Optimization and Hybrid Algorithms

Actual water treatment systems must simultaneously satisfy multiple objectives, including compliance with water quality standards, minimization of energy consumption, and optimization of costs. Single-objective algorithms are often inadequate for such requirements. Multi-objective optimization algorithms can generate Pareto optimal solution sets, thereby providing decision-making support for plant managers.

To overcome the limitations of individual algorithms, hybrid optimization approaches have gradually become a research hotspot. For example, combining the fast convergence of Particle Swarm Optimization with the global search capability of Simulated Annealing can improve optimization stability under complex operating conditions. Similarly, integrating Genetic Algorithm with Ant Colony Optimization allows for both multi-objective search and path planning, making it well-suited for the overall regulation of multi-stage water treatment systems [4].

## 3. Application of Machine Learning in Water Quality Prediction

Machine learning is fundamentally data-driven. It constructs fitting models using historical water quality and process data to achieve dynamic prediction of key indicators, thereby providing a basis for proactive regulation. The mainstream models fall into the following three categories.

Deep neural networks employ multi-layer nonlinear structures to extract data features and exhibit strong fitting capabilities. They are well suited for handling high-dimensional and highly nonlinear water quality data. In scenarios with large influent fluctuations, these networks can accurately predict changes in indicators such as COD and ammonia nitrogen. When combined with optimization algorithms to adjust network weights and structures, their prediction accuracy can be further improved.

Support Vector Machines are based on statistical learning theory and demonstrate excellent performance with small-sample and high-dimensional data. By mapping data through kernel functions, they achieve nonlinear fitting with stable generalization ability. These models can rapidly predict effluent water quality

based on influent parameters, thereby providing reliable references for adjusting chemical dosage and reaction time.

Random Forest is an ensemble learning model that makes joint decisions through multiple decision trees. It possesses strong resistance to noise and missing data, and delivers stable and reliable prediction results. The model can effectively uncover higher-order relationships among water quality parameters, making it particularly suitable for real-world water plant operations where data imperfections are common [5].

## **4. Coupled Application of Artificial Intelligence Optimization Algorithms and Machine Learning**

### **4.1 Coupling Mechanism**

The integration of these two technologies establishes a complete intelligent control closed-loop system. Artificial intelligence optimization algorithms are employed to optimize the hyperparameters, network architecture, weight thresholds, and other elements of machine learning models, thereby improving the models' prediction accuracy and generalization performance. At the same time, the water quality prediction results obtained from machine learning models are used as the objective functions and constraint conditions for the optimization algorithms. This approach enables the algorithms to guide the dynamic adjustment of water treatment process parameters in a reverse direction. Ultimately, it achieves the synergistic optimization of water quality compliance, reduced energy consumption, and cost savings [6].

### **4.2 Typical Coupled Models and Their Applications**

In engineering practice and academic research, commonly used coupled models include PSO-SVM, GA-SVM, HHO-DNN, and SCA-RF, all of which have achieved favorable application results.

For example, the PSO-SVM model employs Particle Swarm Optimization to tune the kernel function parameters and penalty coefficients of the Support Vector Machine. This optimization significantly improves the prediction accuracy of effluent water quality. Based on the improved predictions, chemical dosage can be optimized, resulting in a reduction of 15%–20% in reagent consumption compared with traditional methods.

The HHO-DNN model utilizes the Harris Hawks Optimization algorithm to optimize the architecture and initial weights of deep neural networks. This approach enhances the prediction accuracy of effluent quality in membrane treatment systems and provides reliable support for the regulation of key parameters such as aeration intensity and transmembrane pressure.

In addition, methods that combine deep reinforcement learning with surrogate models enable autonomous learning of optimal strategies through online interaction. These approaches achieve adaptive control in dynamic water quality environments and offer a new technical solution for emergency response to sudden water pollution events [7].

## **5. Challenges and Future Prospects**

### **5.1 Main Challenges**

Although coupled models have achieved remarkable progress in the field of water treatment, their practical application still faces multiple challenges. First, water quality monitoring data often suffer from low sampling frequency, high noise levels, and a large number of missing values, which directly impair model training effectiveness and optimization accuracy. Second, the generalization ability of most models remains insufficient; the majority are trained for specific water sources and treatment processes, resulting in a significant decline in performance when transferred to different scenarios. Third, the combination of deep learning with multi-objective optimization leads to substantial computational demands and high hardware requirements, making it difficult to meet the needs of real-time regulation in water treatment plants. Fourth, the poor interpretability of models, characterized by their black-box nature, makes it difficult for engineering personnel to understand the logic behind decisions, thereby hindering technology promotion and on-site implementation.

## 5.2 Future Development Directions

To address the existing challenges, future research can advance in five key directions. First, methods such as incremental learning and transfer learning should be introduced to enable online model updating and improve cross-scenario adaptability. Second, high-performance hybrid intelligent algorithms should be designed to balance optimization accuracy with computational efficiency and reduce computing power requirements. Third, a full-process monitoring system should be established based on the Internet of Things and big data technologies, combined with data cleaning and fusion techniques to enhance data quality. Fourth, research on explainable artificial intelligence should be conducted, using approaches such as feature analysis and visualization to reveal the decision-making logic of models. Fifth, customized coupled models should be developed for different water sources and treatment processes, forming standardized engineering solutions to accelerate the practical implementation and application of these technologies [8].

## 6. Conclusion

The coupled application of Artificial Intelligence optimization algorithms and machine learning provides a highly efficient and feasible technical framework for water treatment process optimization and water quality prediction. While optimization algorithms—such as Particle Swarm Optimization, Genetic Algorithms, and Harris Hawks Optimization—enable the effective global search for optimal process parameters, models including Deep Neural Networks, Support Vector Machines, and Random Forests achieve the precise prediction of water quality indicators.

The integration of these methodologies into a closed-loop “prediction-optimization-control” system significantly enhances the operational efficiency and effluent stability of water treatment facilities. Furthermore, this synergy effectively reduces energy consumption and operational costs, serving as a core pillar for the intelligent evolution of the water treatment sector.

Despite current limitations regarding data quality, model generalizability, and computational complexity, the continuous convergence of AI, the Internet of Things, and Big Data will drive the refinement of theoretical frameworks and the expansion of engineering applications. Moving forward, this coupling technology is poised to become a critical instrument for water resource management and environmental remediation. It will provide a robust technical foundation for ensuring national water security and advancing the green and low-carbon development of the water treatment industry.

## References

- [1] Pereira, J. L. J., Oliver, G. A., Francisco, M. B., Cunha, S. S., Jr., & Gomes, G. F. (2022). A review of multi-objective optimization: Methods and algorithms in mechanical engineering problems. *Archives of Computational Methods in Engineering*, 29(4), 2285–2308. <https://doi.org/10.1007/s11831-021-09663-x>
- [2] Geng, K., Pan, Q., & Huang, J. (2022). A many-objective evolutionary algorithm with Pareto-adaptive scalarizing functions. *Neurocomputing*, 474, 141–151. <https://doi.org/10.1016/j.neucom.2021.12.023>
- [3] Abualigah, L., Zitar, R. A., Al-Abbas, A., & Alshari, H. (2025). A comprehensive survey of automated surrogate-assisted evolutionary algorithms. *Swarm and Evolutionary Computation*, 92, Article 101817. <https://doi.org/10.1016/j.swevo.2024.101817>
- [4] Abualigah, L., Zitar, R. A., Al-Abbas, A., & Alshari, H. (2025). Hybridization of meta-heuristic optimization algorithms: A comprehensive review. *Archives of Computational Methods in Engineering*, 32(2), 735–784. <https://doi.org/10.1007/s11831-024-10152-0>
- [5] Geng, K., Pan, Q., & Huang, J. (2022). A many-objective evolutionary algorithm with Pareto-adaptive scalarizing functions. *Neurocomputing*, 474, 141–151. <https://doi.org/10.1016/j.neucom.2021.12.023>
- [6] Li, W., Chen, Q., Zhang, J., Lin, J., & Wang, Y. (2023). A survey of multi-objective evolutionary algorithms based on decomposition. *IEEE Transactions on Evolutionary Computation*, 27(2), 402–421. <https://doi.org/10.1109/TEVC.2022.3172012>

- [7] Wang, X., Zhang, B., Wang, J., Zhang, K., & Jin, Y. (2022). A cluster-based competitive particle swarm optimizer with a sparse truncation operator for multi-objective optimization. *Swarm and Evolutionary Computation*, 71, Article 101083. <https://doi.org/10.1016/j.swevo.2022.101083>
- [8] Lv, Z., & Song, H. (2022). Real-time digital twins: Visions and applications. *Data Science and Engineering*, 7(4), 402–427. <https://doi.org/10.1007/s41019-022-00193-5>

### **Funding**

This research received no external funding.

### **Conflicts of Interest**

The authors declare no conflict of interest.

### **Acknowledgment**

This paper is an output of the science project.

### **Open Access**

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

