Research on Management of Lake Water Levels Based on Multi-Objective Genetic and PID Control Algorithm

Zeyuan Hou *

School of Control Science and Engineering, Shandong University, JiNan, China, 250002

* Corresponding Author Email: 202222171161@mail.sdu.edu.cn

Abstract. A comprehensive optimization approach utilizing a multi-objective genetic algorithm and PID control algorithm is proposed for addressing water level management in lake systems. Taking the Great Lakes of North America as a case study, this research investigates dynamic regulation methods for lake water levels. Through analysis of monthly water level data from the lakes, an optimal water level calculation method specific to each lake is determined. Subsequently, a water level optimization model is developed with primary objectives of ensuring drinking water supply, supporting navigation, and preserving the ecological environment. The non-dominated sorting genetic algorithm (NSGA-II) is employed to solve the model and determine the optimal water levels for the Great Lakes. Recognizing the challenge of simultaneously achieving optimal water levels in the actual Great Lakes system, a PID control algorithm is utilized for water level adjustments. By establishing a dynamic water level change model, a PID controller is designed and its parameters are calibrated through numerous experiments to effectively maintain water levels within the optimal range. This research offers a scientific foundation and practical methodology for enhancing water level management in the Great Lakes, thereby contributing significantly to the sustainable utilization of lake water resources.

Keywords: Great Lakes, Multi-objective Genetic Algorithm, Water Level Management, PID Control Algorithm.

1. Introduction

The Great Lakes of North America are the world's largest group of freshwater lakes, and their water resources hold significant value in fishing, recreation, power generation, drinking water, navigation, and wildlife habitats. However, with climate change and increased human activity, fluctuations in lake levels have become increasingly pronounced, posing major challenges to water resource management[1]. How to scientifically and reasonably regulate lake levels and balance the needs of all stakeholders has become a critical issue in current water resource management.In recent years, many scholars have conducted in-depth research on the water level management of lake systems[2]. Early studies focused on optimization for a single objective, such as predicting water level changes through hydrological models or analyzing historical water level data through statistical methods. With the development of multi-objective optimization techniques, researchers began to try to apply multi-objective genetic algorithms[3-4] to water resources management to balance the needs of different stakeholders. In addition, PID control algorithm[5] is widely used in the field of water level regulation because of its simple and efficient characteristics. However, most of the existing researches focus on single lake or single target, and lack of comprehensive optimization and dynamic regulation of the overall water level of the Great Lakes.

This study introduces a novel optimization strategy that integrates a multi-objective genetic algorithm with a PID control algorithm to improve water level management in the Great Lakes. It constructs a comprehensive optimization framework that considers potable water supply, navigation, and ecological conservation. The non-dominated sorting genetic algorithm (NSGA-II) is employed to identify optimal water levels for the Great Lakes throughout the year. To tackle the challenges associated with achieving optimal water levels for the entire system, a water level adjustment model based on the PID control algorithm is proposed for dynamic regulation. The paper is organized as follows: the second section establishes the water level optimization model and elucidates the NSGA-II algorithm; the third section introduces the water level regulation technique utilizing the PID control algorithm and validates it through experimental analysis; the fourth section synthesizes the results and delineates future research avenues.

2. Water Level Optimization Model Based on Multi-Objective Genetic Algorithm

This section classifies stakeholders in the water levels of the Great Lakes and establishes a target optimization model to maximize the sum of the benefits obtained by various stakeholders, thereby obtaining the optimal water level of the Great Lakes.

2.1. Data source and preprocessing

The study utilizes data extracted from monthly water level records of the Great North American Lakes (Lake Superior, Lake Michigan, Lake Huron, Lake Erie, and Lake Ontario) spanning from 2000 to 2022,sourced from the 2024 ICM Problem D dataset. Prior to formulating a target optimization model, preprocessing of the collected Great Lakes water level data is essential. Initially, the data undergoes cleaning procedures where outliers and missing values are eliminated, and standardization into meters is conducted. Subsequently, the data is categorized, organized by lakes, years, and months to construct a structured dataset. Considering the seasonal fluctuations in water levels, an analysis of the seasonal rate of change in lake water levels enables a more precise determination of optimal water levels for the Great Lakes across different periods of the year. To furnish dependable raw input data for subsequent multi-objective optimization models, this section computes the average water level across the five lakes.The formula for calculating the average water level $\bar{H}_i$ of the **i**-th lake is:

$$\bar{H}_i = \frac{1}{N}\sum_{y=2000}^{2022}\sum_{m=1}^{12} H_{i,y,m} \tag{1}$$

Among them, $N = 23\times 12$ is the total number of months, $H_{i,y,m}$ is the water level of the i-th lake in the m-th month of the y-th year,and the average water level of each lake is shown in Figure 1. Further, calculate the average water level $\bar{H}_{i,m}$ for each lake every month, and the formula is:

$$\bar{H}_{i,m} = \frac{1}{23}\sum_{y=2000}^{2022} H_{i,y,m} \tag{2}$$

where $m$is a specific month. In order to more intuitively reflect the changes of water levels with seasons, this study selected the average water levels of January, March, May, July, September and November to draw the line chart as shown in Figure 2.

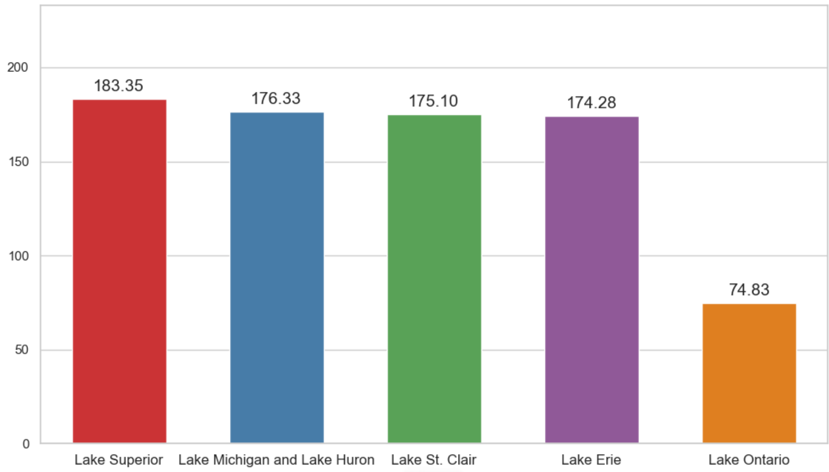

Figure 1. Average water level in the Great Lakes

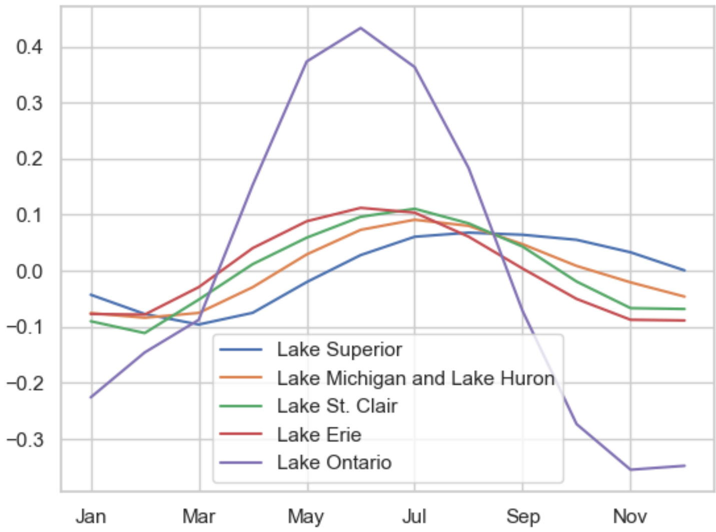

Figure 2. Seasonal changes in the average water level of the Great Lakes

The mean water levels of Lakes Superior, Michigan, Huron and Erie can be found to fluctuate less with the seasons, indicating that water levels in these lakes are relatively stable throughout the year. Therefore, for these four lakes, the average water level throughout the year can be used as the best water level at any time. In contrast, the average water level of Lake Ontario shows significant seasonal fluctuations and is greatly affected by the seasons, so it is necessary to use the seasonal average water level as the best water level in different seasons. The average water level obtained is shown in Table 1.

2.2. Formulation of a Water Level Optimization Model

The water resources of the Great Lakes are used for a wide variety of purposes, including fishing, recreation, power generation, drinking water, shipping, habitat for plants and animals, and so many stakeholders are involved. Too high or too low a water level can affect the interests of some stakeholders. The key to water level management is to balance three core interests: ensuring drinking water supply, maintaining shipping, and protecting the ecological environment. In order to meet the needs of different stakeholders, the ideal water level of each stakeholder is determined by combining the average water level value of different seasons obtained from historical water level data. At the same time, different weights are assigned to these three objectives according to the statistical analysis of relevant reports, as shown in Table 1. Then, by analyzing the difference between the ideal water level and the optimal water level, the multi-objective optimization model between the interests of stakeholders and the optimal water level is established. The objective function is defined as follows:

$$\text{Max} \sum_{i=1}^{5}\sum_{s=1}^{4}\sum_{n=1}^{3} R_{i,s,n}(L_{i,s}) \times \frac{\prod_{n=1}^{3}\tanh(R_{i,s,n}(L_{i,s}))}{\sum_{n=1}^{3}\tanh(R_{i,s,n}(L_{i,s}))}, \tag{3}$$

where $L_{i,s}$ represents the water level of the i-th in the s-th season, $R_{i,s,n}(L)$ is the return function[6] of the n-th stakeholder, and its formula is:

$$R_{i,s,n}(L_{i,s}) = \frac{w_n}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{(L_{i,s}-L^*_{i,s,n})^2}{2\times\sigma^2}} \quad (\sigma = 4), \tag{4}$$

It is indicated that the benefit is highest at the optimal water level, and the benefit decreases gradually as the water level deviates from the optimal value, w_n represents the weight of the n-th stakeholder and $L^*_{i,s,n}$ is the ideal water level value of the n-th stakeholder in the i-th in the s-th season.

Remark1: A 2 to 3 foot deviation from the normal water level can have a significant impact on the interests of certain stakeholders. Therefore, the water level should be strictly limited to 0.9 m of the initial optimal water level, i.e., $|L_{i,s} - L^{*}_{i,s,n}| < 0.9$.

Table 1. Ideal water levels for stakeholders

Lake	Ideal water level [water supply($w_1 = 4$)、shipping($w_2 = 2$)、environment($w_3 = 1$)]/m			
	Spring(s = 1)	Summer(s = 2)	Autumn(s = 3)	Winter(s = 4)
Lake Superior (i = 1)	[183.6,183.85,183.35]			
Lake Michigan (i = 2)	[176.58,176.83,176.33]			
Lake Huron (i = 3)	[176.58,176.83,176.33]			
Lake Erie (i = 4)	[174.53,174.78,174.28]			
Lake Ontario (i = 5)	[75.19,75.44,74.94]	[75.32,75.57,75.07]	[74.90,75.15,74.65]	[74.90,75.15,74.65]

2.3. Non-Dominating Ranking Genetic Algorithm(NSGA-II)

In the process of solving the optimal water level in the Great Lakes for four seasons, we use the Non-Dominating Ranking Genetic Algorithm for optimization. NSGA-II is a multi-objective optimization algorithm that can find a set of Pareto optimal solutions between multiple conflicting targets[7-8]. In the optimization process of NSGA-II, we first randomly generate a set of initial water level values $L_{i,s}$ as the initial population, each water level value represents a possible solution, and the fitness function is defined to be the same as the objective function Eq. (3). In the process of population evolution, in order to ensure the diversity of the population and the dispersion of the solution, we used two methods: non-dominant ranking and crowding distance ranking to evaluate the advantages and disadvantages of individuals. The core idea of the non-dominant ranking algorithm is to stratify the population according to the dominance relationship between individuals, and the solution closest to the front of Pareto has the highest rank, which is the first layer, followed by the second layer, and so on. For individuals in the same layer, the crowding distance is evaluated, and the larger the crowding distance, the more sparse the individual is in the solution space, which is more conducive to maintaining the diversity of the population.The evaluation systems employ a binary tournament selection strategy. Two individuals are randomly chosen from the population and compared based on their non-dominant ranking. The individual with the lower ranking is then selected. In cases where both individuals have the same rank, the one with a larger crowding distance is chosen to preserve solution diversity.

The individual that wins using the binary tournament method is used as Parent 1, and the same operation is performed to obtain Parent 2. To avoid premature convergence in genetic algorithms, additional checks are added to ensure that Parent 1 and Parent 2 are different. Then, the two parents are crossed to produce two offspring. To prevent premature convergence in genetic algorithms, this paper employs the "Simulated Binary Crossover (SBX)" operator during the crossover process. For the two offspring produced, one undergoes mutation while the other remains unchanged. The mutation operator chosen in this paper is "Polynomial Mutation":

$$Z_{j+1} = Z_j + \Delta_j \, , \tag{5}$$

$$\Delta_j=\begin{cases}(2\mu_i)^{\frac{1}{\eta+1}}, & \mu_i \le 0.5\\ 1-[2(1-\mu_i)]^{\frac{1}{\eta+1}}, & \mu_i \ge 0.5\end{cases}, \tag{6}$$

In the formula, Z_{j+1} is the offspring individual, Z_j is the parent individual, Δ_j is the variation quantity, μ_i is the random number uniformly distributed in the interval $(0,1)$, and η is the variation distribution parameter. A larger value of η can reduce the variation amplitude and ensure the stable evolution of the offspring individual in the solution space.

In the population selection process, this section adopts an elite retention strategy[9] to maintain the optimal solution set. This involves merging the parent population P_t and the offspring population Q_t into a larger population R_t, then performing non-dominated sorting and crowding distance calculation on this population.Subsequently,the parents P_{t+1} of the next generation are sorted and selected based on the above methods through ranking and choosing,and then crossed to generate the offspring Q_{t+1}using a general method.Through 100 iterations, NSGA-II progressively converges the individuals in the population towards a set of solutions, specifically the optimal water level value $L^*_{i,s}$ as presented in Table 2. The iterative progress is illustrated in Figure 3.

Table 2. Optimal water level $L^*_{i,s}$/m

i S	1	2	3	4
1	183.64			
2	176.62			
3	176.62			
4	174.57			
5	75.23	75.36	74.94	74.94

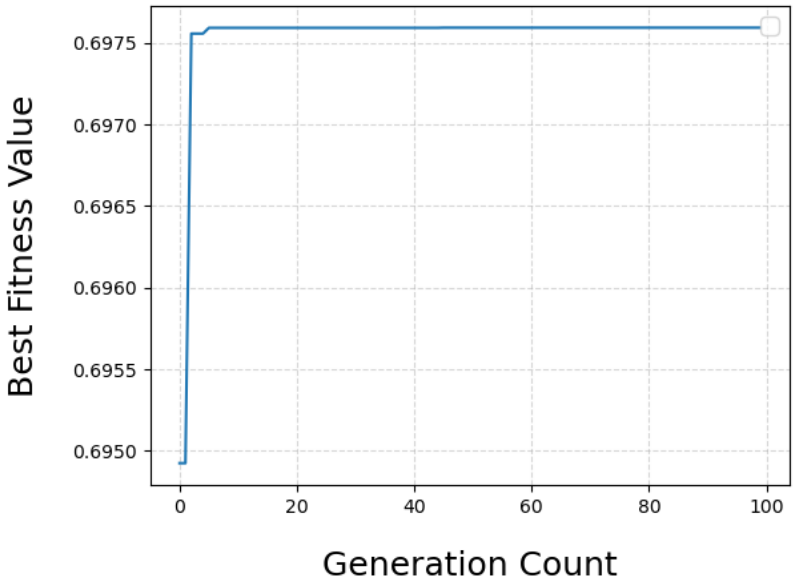

Figure 3. Genetic iteration diagram

Remark2: The analysis in section 1 shows that the average water level of Lakes Superior, Michigan, Huron and Erie fluctuated less with the seasons, so the optimal water level in Table 2 can be regarded as the optimal water level at any time of the year. For Lake Ontario, the optimal water level in a season is used as the optimal water level at any time in the season.

3. Water Level Regulation Method Based on PID Control Algorithm

In view of the fact that the five lakes system may not reach the optimal water level simultaneously, this section proposes a water level adjustment method based on PID control algorithm, which dynamically adjusts the lake water volume according to the data of lake water inflow and outflow (temporarily ignoring the influence of natural factors) and artificial control factors (dam adjustment), and maintains the water level in the optimal range while avoiding global adjustment at the expense of the interests of a few lakes. PID control algorithm can adjust the system output more accurately by combining proportional (P), integral (I) and derivative (D) control components, so as to achieve effective control of water level.

3.1. Dynamic Model for Water Level Fluctuations

In this section, the regulated target water level of each lake is determined based on the optimal water level in Table 2 obtained in Section 2.Based on the hydrographic data provided and the Great Lakes System Profile as shown in Figure 4, the average discharge of the rivers connecting the lakes can be calculated, and the inflow and outflow data of the Great Lakes can be obtained in combination with the surface area of the lakes. Through these data, the current water level can be derived after the water level changes, and the specific dynamic change formula of water level is as follows:

$$\begin{cases} H_{1,s} = H_{1,s} - \dfrac{Q_{1,s}}{S_1} \\ H_{2,s} = H_{2,s} + \dfrac{Q_{1,s}}{S_2+S_3} - \dfrac{Q_{2,s}}{S_2+S_3} \\ H_{3,s} = H_{3,s} + \dfrac{Q_{1,s}}{S_3+S_2} - \dfrac{Q_{2,s}}{S_3+S_2} \\ H_{4,s} = H_{4,s} + \dfrac{Q_{3,s}}{S_4} - \dfrac{Q_{4,s}}{S_4} \\ H_{5,s} = H_{5,s} + \dfrac{Q_{4,s}}{S_5} - \dfrac{Q_{5,s}}{S_5} \end{cases}, \tag{7}$$

where, $H_{i,s}$ is the water level of the i-th lake in the s-th season, S_i、$Q_{i,s}$ is the surface area of the i-th lake and the average discharge of the i-th river in the s-th season, respectively, which are given in Table 3 and Table 4. (Download URL for Figure 4: https://vividmaps.com/great-lakes-profile/)

Table 3. Surface area of lake

Lake(i)	S/m^2
1	82103
2	57757
3	59570
4	25667
5	19554

Table 4. Average river discharge $Q_{i,s}/m^3$

i S	1	2	3	4
1	1969.62	2414.68	2302.34	1993.53
2	5676.65	5803.47	5803.96	5220.14
3	5992.52	5934.67	5966.7	5556.78
4	6109.37	6185.87	5857.62	5867.94
5	7727.07	8368.91	7857.5	7379.97

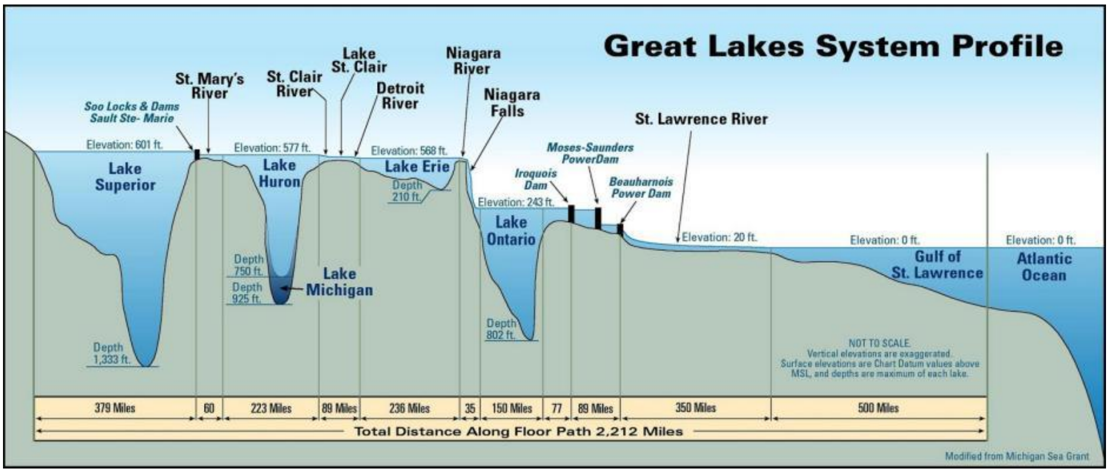

Figure 4. Great Lakes System Profile

3.2. PID controller design

If the sampling period is set to $T_s = 1$ day, the output of the PID controller is expressed in discrete form, and then the output control quantity of the i-th lake is obtained:

$$u_i(k) = K_p \cdot e(k) + K_i \cdot T_s \cdot \sum_{t=0}^{k} e(t) + K_d \cdot \frac{e(k) - e(k-1)}{T_s}, \tag{8}$$

where K_p,K_i and K_d respectively denote the proportional , integral and derivative gains of the controller and $e(k)$ represents the deviation of the water level on the k-th day from the optimal water level. According to the characteristics of low-frequency sampling once a day, proportional gain K_p should not be too large, integral gain K_i and differential gain K_d should be appropriately reduced to avoid excessive accumulation error and excessive sensitivity to noise. Therefore, the water level optimization model constructed by formula (7) in this section is used to input the water level $H_{i,s}$ before regulation and the target water level $L^*_{i,s,n}$, calculate the output of PID controller by formula (8), and denote $H_{i,s} + u_i(k)$ as the water level value H'_i after artificial regulation.By simulating daily water level fluctuations throughout a year, we construct a curve depicting the temporal changes in water levels post-adjustment. This enables us to observe the dynamic response characteristics of the system and assess the stability of the controller's adjustment.After multiple tests[10], the controller parameters are obtained as: $K_p = 0.7$、$K_i = 0.001$、$K_d = 0.1$. This set of controller parameters can better adjust the water level and stabilize the water level near the optimal water level as shown in Figure 5.

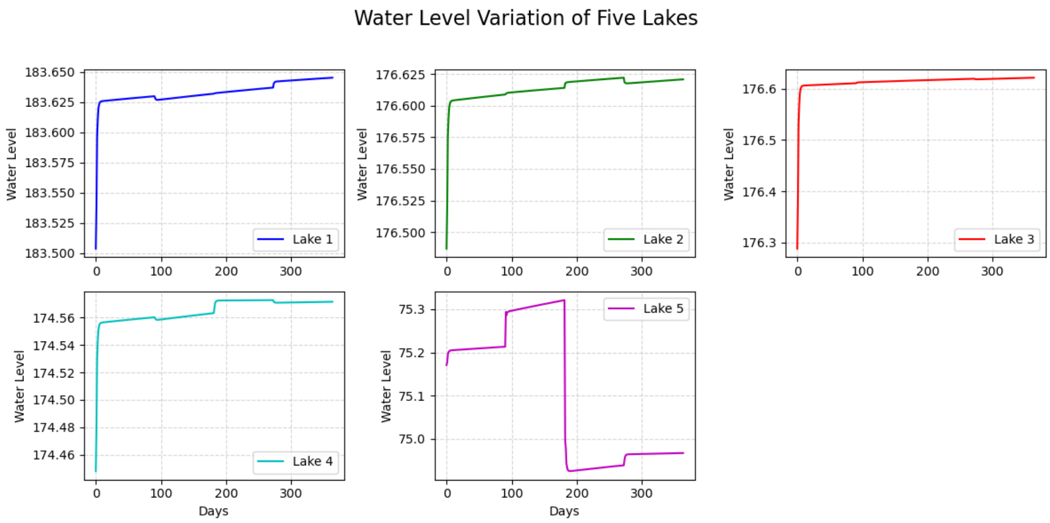

Figure 5. Lake water level after artificial control

4. Conclusions

This paper successfully constructed a Great Lakes water level optimization model based on multi-objective genetic algorithm, determined the ideal water level of various stakeholders with the help of historical water level data, and obtained the optimal water level of the Great Lakes at any time by using NSGA-II algorithm, effectively coordinating the core interests of drinking water supply, shipping and ecological protection. At the same time, the proposed water level regulation method based on PID control algorithm, through the construction of dynamic water level change model and the design of a reasonable PID controller, optimized the controller parameters after many tests, realized the dynamic adjustment of lake water volume according to the actual situation, accurately maintained the water level in the optimal range, and avoided the damage to the interests of some lakes in the process of global adjustment. This study offers a scientific and systematic approach to managing water levels in the Great Lakes, enhancing resource utilization efficiency and fostering sustainable ecological development. Moreover, it holds valuable implications for water level management research in analogous lakes, with the potential to advance the field of lake water resources management.Future research should broaden the scope of model applications by integrating long-term influencing factors like climate change, population growth, and industrial development to enhance model adaptability. Additionally, there is a need to enhance the computational efficiency and accuracy of the non-dominated ranking genetic algorithm II and PID control algorithms under complex conditions. Long-term monitoring of the implementation effects of the Great Lakes water level management strategy should also be conducted to continuously evaluate and adjust the model.

References

[1] Ethan J. Theuerkauf, Katherine N. Braun. Rapid water level rise drives unprecedented coastal habitat loss along the Great Lakes of North America[J]. Journal of Great Lakes Research, 2021, 47(4): 945-954.

[2] Ruan Shi. Development of rule curves for reservoir sluice operation and water level management [J]. Water Resources Science and Cold Zone Engineering, 2023, 6(9): 51-54.

[3] DAI Haishu, GUO Zhimin, ZHAI Jiang, et al.Multi-objective parameter optimization of piston pump damping groove based on NSGA-II[J].Chinese Hydraulics & Pneumatics, 2023, 47(8): 26-33.

[4] LIU D, HUANG Q, YANG Y, et al. Bi-objective algorithm based on NSGA-II framework to optimize reservoirs operation[J]. Journal of Hydrology, 2020, 585(3): 124830.

[5] Hu Qiuyan. Automatic gate control of irrigation water diversion system based on variable structure PID algorithm[J]. Electromechanical Technology of Hydropower Station, 2024,47(11):71-73.

[6] ZHANG Pingdong,MA Jun. Supply chain network equilibrium model based on newsboy model under normal distribution of demand[J].Logistics Science and Technology, 2024, 47(21): 126-130.

[7] HUANG Xiaomin,LEI Xiaohui,et al. Study of multi-objective optimization of parameters of HYMOD model using NSGA-II algorithm[J]. Yangtze River, 2012, 43(2): 16-21.

[8] HUANG Suning, HUANG Xianfeng, et al. Research on water and soil resources optimal allocation based on multi-objective genetic algorithm[J]. China Rural Water Conservancy and Hydropower, 2013(5): 33-36,41.

[9] Liu Xiaoli. Green multimodal transport path optimization based on elite retention strategy genetic algorithm [J]. China Logistics and Purchasing, 2025, (01): 94-97.

[10] Cai Bingyu, Lu Xujin, Chen Chaoran. Time domain and frequency domain analysis of PID parameter tuning method [J]. Automation Applications, 2024, 65 (18):93-95+98.