WEP Warwick Evans Publishing

Transactions on Environment, Energy and Earth Sciences
ISSN: 2960-1924, eISSN: 2960-2300 | Volume 3
GRET 2024

Optimizing Water Level Regulation in the Great Lakes: An Integrated Approach Using Genetic Algorithms and Network Flow Models for Sustainable Resource Management

Yuxuan Tang*, Yuke Zhang, Yixi Wu

School of International Economics and Trade, Central University of Finance and Economics, Beijing, China, 102206

* Corresponding Author Email: tyx1424319372@foxmail.com

Abstract. The Great Lakes, as a vital freshwater resource, are significantly influenced by climate change and anthropogenic activities, leading to acute water level fluctuations with profound impacts on various sectors including agriculture, industry, and ecology. This study addresses the need for scientifically sound water level regulation by developing an integrated approach that employs a genetic algorithm and a network flow model. The model is designed to balance the interests of multiple stakeholders and optimize water levels across different periods, thereby ensuring the sustainable use of water resources and ecological balance. Through comprehensive data collection and processing, the study constructs a robust prediction model that accounts for economic activities and ecological protection. The application of genetic algorithms enhances the efficiency and accuracy of the model, providing a practical tool for decision-makers in water resource management. The findings underscore the model's reliability and its potential to support policy-making for sustainable water resource management in the Great Lakes region.

Keywords: Great Lakes; Water Level Regulation; Genetic Algorithm; Network Flow Model; Sustainability.

1. Introduction

The Great Lakes—Superior, Michigan, Huron, Erie, and Ontario—represent the largest group of freshwater lakes globally, storing approximately 21% of the world's surface freshwater. They are rich in biodiversity and significantly influence agricultural irrigation, industrial water usage, shipping, tourism, and ecosystem maintenance in the United States and Canada. However, recent climate change and human activities have led to pronounced fluctuations in the water levels of the Great Lakes, impacting shipping, fishing, tourism, and increasing the risk of natural disasters such as floods and droughts, which poses a threat to the safety and property of lakeside residents. Therefore, scientifically regulating the water levels of the Great Lakes is essential for sustainable water resource utilization, ecological balance, and economic stability.

According to the data from the NWS, the impact of global climate change on Great Lakes water levels is complex. The IPCC reports indicate that global climate change has escalated dramatically in the last 60 years, with the average surface temperature rising approximately 0.85 °C since 1880, notably affecting the Great Lakes climate. Water level fluctuations directly influence agriculture, industry, shipping, and tourism in the region, necessitating control strategies that consider economic activities and ecological protection.

Various mathematical models have been proposed for water level regulation. Wang Hong (2009) introduced stochastic differential equations to more accurately describe reservoir water level changes, demonstrating their practical utility in hydraulic engineering. Lin Hao (2021) estimated water level and discharge relationships using neural networks, integrating genetic algorithms and information entropy. Yu Xinpeng et al. (2022) applied machine learning to predict water levels and discharge in Jinan City through various regression models, ultimately identifying an optimal regression equation. Chen Huaniang et al. (2022) utilized a BP neural network optimized by genetic algorithms to predict karst spring water levels in Northern China, enhancing prediction stability and reducing computational costs. Zhao Lubin and Song Xiaodan (2024) developed a dynamic flood limit water level control model using genetic algorithms for real-time regulation, enhancing reservoir benefits. Lv Haifeng et al. (2024) adopted a hybrid CNN-LSTM-Attention model for water level predictions, significantly improving accuracy and robustness. Zhang Jian (2024) proposed a SARIMA-based monitoring system, effectively addressing non-stationarity and seasonal variability in hydrological data.

Building upon the existing body of research, our study presents an innovative approach to water level regulation in the Great Lakes by integrating genetic algorithms with network flow models, marking a significant advancement in sustainable water resource management. This integration provides a nuanced and holistic method that addresses the complex interplay between ecological needs and economic activities, offering a robust framework for decision-makers to balance multiple stakeholder interests.

A critical contribution is the development of a comprehensive income index, which quantifies stakeholder preferences and sensitivities to water level changes. This index is integral to the genetic algorithm's fitness function, ensuring optimization aligns with the broader economic and ecological objectives of the Great Lakes region. Advanced data processing techniques and a dynamic, real-time regulation strategy are applied to enhance the predictive accuracy and efficiency of the model. Monte Carlo simulation analysis further validates the model's reliability, demonstrating its low sensitivity to variations in control dam outflow and environmental factors.

In conclusion, the research introduces a scientifically sound, data-driven model for sustainable water resource management in the Great Lakes with practical implications for policy-making. The findings contribute to the literature and provide a foundation for future studies focused on enhancing the resilience and sustainability of freshwater ecosystems in response to climate variability and anthropogenic pressures.

2. Research methods

2.1. Data collection and processing

As an important freshwater resource in North America, the fluctuation of the water level of the Great Lakes has a profound impact on the regional economy, ecological environment and social welfare. Accurate water level prediction is not only related to the rational allocation of water resources, but also the key to flood control and disaster reduction, shipping safety and ecosystem protection. Therefore, comprehensive and accurate data collection is the cornerstone of building effective water level prediction models.

2.1.1. Data source and selection basis

This study selected multiple authoritative data sources to ensure the reliability and representativeness of the data:

Water level data: The United States Geological Survey (USGS) and Environment Canada provide exhaustive historical water level records, which are the basis for analyzing water level trends and patterns.

Socioeconomic data: Federal Economic Data (FRED), the U.S. Energy Information Administration (EIA), and Statistics Canada provide data on population, economic, and industrial activity related to the Great Lakes region, which help us assess the potential impact of water level changes on human activities.

2.1.2. Data pre-processing

Data preprocessing is a key link to improve the accuracy of model prediction. In this study, data is processed and integrated through the following three steps:

Data cleansing: Using a combination of automated scripts and manual audits, missing values and outliers are identified and processed to ensure the integrity and accuracy of the data set.

Data standardization: Z-score standardization method is adopted to convert data of different dimensions to the same scale to improve the convergence speed and prediction performance of the model.

Data integration: By aligning data from different sources and types according to time series, a multi-dimensional and comprehensive database is built to provide rich input features for the model.

2.1.3. Reliability and Limitations of the data

In the process of selecting and processing data, the research pays attention especially to the reliability of data. All data are sourced from recognized authorities and are subject to strict quality control. However, the limitations of the data should not be ignored. For example, the spatio-temporal resolution of some meteorological data may not be enough to fully capture small changes in local areas. In future studies, higher resolution data sources to improve the refined prediction ability of the model will be further explored.

2.1.4. Data visualization and analysis

Data visualization is an effective means to understand data characteristics and find potential rules. Figure 1 uses charts and graphs to show seasonal changes in Great Lakes water levels, long-term trends, and relationships to other environmental and socioeconomic factors, the changes in water levels in the Great Lakes over the last two decades are shown through a line chart, clearly revealing seasonal fluctuations and long-term trends in water levels.

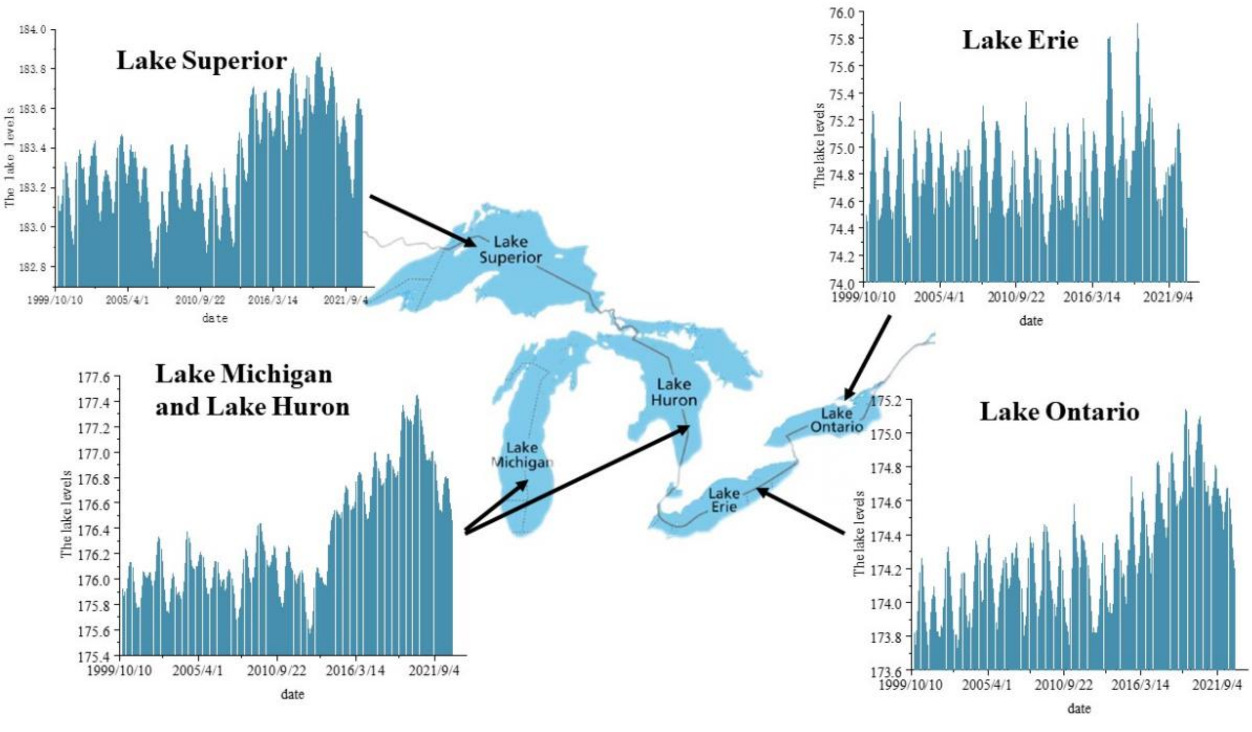

Figure 1. The trend of water level change in the Great Lakes over the last two decades

2.2. Model construction -- optimal water level prediction model based on genetic algorithm

The water level regulation of the Great Lakes is a complex multi-objective optimization problem, involving the needs of ecological protection, economic development, social welfare and other aspects. To build a scientific and reasonable model, it is not only necessary to accurately predict water level changes, but also to seek a balance between the interests of multiple parties. This study aims to provide decision support for the optimal management of water level in the Great Lakes through advanced mathematical models and algorithms.

2.2.1. Stakeholder analysis and preference quantification

In the Great Lakes region, water level changes affect different stakeholders differently. The study begins by identifying key stakeholders including agriculture, industry, shipping, ecological conservation and tourism. Through various data collection channels, quantitative data on the impact of water level change on the output value of various industries were obtained and further combined with relevant mathematical models to quantify the preference and sensitivity of various stakeholders to water level change.

2.2.2. Comprehensive income index

The comprehensive income indicator is the key established by this study to measure the overall impact of the Great Lakes water level on the various stakeholders along the lakes. The metric is constructed based on the following formula:

$$H(t) = \sum_{i=1}^{n} \omega_i \bullet \rho_i(t) \tag{1}$$

Where, ω_i is the weight of the fourth stakeholder, reflecting its importance in overall economic activity; $\rho_i(t)$ is the difference between the stakeholder's output value and the theoretical optimal output value under time, calculated by the following formula:

$$\rho_i(t) = V_i(t) - \bar{V}_i \tag{2}$$

Where, $V_i(t)$ is the actual output value of the fourth stakeholder in time, and $\bar{V}_i$ is its long-term average output value.

2.2.3. Design and implementation of genetic algorithm

The genetic algorithm, recognized for its efficiency in solving complex optimization problems, is pivotal in this study for regulating water levels in the Great Lakes. The process begins with the random generation of an initial population, where each individual represents a potential water level regulation scheme for the lakes. The fitness function evaluates each individual's performance based on a comprehensive income index, ensuring that individuals yielding higher potential benefits are favored. Utilizing the roulette selection method, individuals with higher fitness are more likely to be chosen as parents for the next generation.

Subsequent operations involve crossover and mutation to enhance the solution's diversity and quality. The crossover operation combines advantageous genes from parent individuals through single-point and multi-point techniques, producing offspring with potentially superior characteristics. Mutation introduces random variations in some genes to further diversify the population and steer clear of local optima. This iterative cycle of selection, crossover, and mutation continues until specific termination conditions are satisfied, such as reaching a predetermined number of iterations or achieving an acceptable solution quality. Finally, the output results of the genetic algorithm are meticulously analyzed to identify the optimal or near-optimal water level regulation scheme, contributing to sustainable water resource management in the region.

2.3. Optimization strategy of genetic algorithm

In order to improve the efficiency and quality of genetic algorithm, the following optimization strategies are adopted in this study:

Adaptive crossover and mutation rate: The crossover and mutation rate are dynamically adjusted according to the iterative process of the algorithm to balance the ability of global search and local search.

Introduction of elite strategy: Retain a certain percentage of optimal individuals in each iteration directly into the next generation to ensure the inheritance of excellent solutions.

Diversity preservation mechanism: By implementing a NicheMethod based approach, the coexistence of different solutions in the population is encouraged to avoid premature convergence.

2.3.1. Differential equation of water level change

The research have developed differential equations that describe changes in water levels in each of the Great Lakes, taking into account lake inflow, outflow, precipitation, evaporation, and other human impacts. Here's the equation:

$$\frac{dH_i}{dt} = \frac{1}{A_i}(N_{in,i} - N_{out,i} + P_i - E_i + F_i + I_i) \tag{3}$$

Among them, i represents the water level of the fourth lake, A_i is the area of the lake, $N_{in,i}$ and $N_{out,i}$ is respectively the total inflow and outflow of the lake, P_i is the precipitation, E_i is the evaporation, F_i represents other natural factors, and I_i represents the impact of human activities.

2.3.2. Modeling of inter-lake discharge relationship

In order to describe the interaction between lakes, the research built a network flow model, taking the average water level of Lake Superior S_0 as the dependent variable, establishing a linear regression model for the average water level of the other four lakes as the independent variable, and the corresponding weighted parameters as the regression coefficients. The following model was obtained:

$$S_0 = \alpha_1 S_1 + \alpha_2 S_2 + \alpha_3 S_3 \tag{4}$$

The correlation coefficient $\rho(i,0)$ is taken as the estimated value of each, where the correlation coefficient α_i representing the difference between the water level of the fourth lake and Lake Superior, Cov represents the covariance, Var represents the variance:

$$\rho(i,0) = \frac{Cov(S_i - S_0, S_0)}{\sqrt{Var(S_i - S_0)Var(S_0)}} \tag{5}$$

From the above two formulas, the model can be simplified as:

$$S_0 = \rho(1,0)S_1 + \rho(2,0)S_2 + \rho(3,0)S_3 \tag{6}$$

3. Results

3.1. Results of optimal water level prediction

In the model, genetic algorithm is used to predict the optimal water level of each lake in the Great Lakes. The algorithm takes into account the needs of different stakeholders and aims to maximize the comprehensive benefits. The forecast results show that the optimal water level for each lake as shown in Table 1:

Table 1. Optimal water level prediction results

Lake names	Lake Superior	Lake Michigan and Lake Huron	Lake Erie	Lake Ontario
Best water level (m)	183.415	176.628	174.475	74.776

The above water level data is the optimal solution obtained after considering the preference weights of various stakeholders through algorithm iteration, that is, the sum of the overall interests of all the stakeholders of the lakeside collective reaches the maximum lake water level in each time period.

3.2. Water level regulation results

3.2.1. Construction of regulation network model

According to the above model analysis, the network flow model of the Great Lakes and their connected rivers is constructed to describe the water flow relationship between the lakes. The model uses differential equation to describe the water level change, and adopts multi-objective programming method to determine the optimal water level.

3.2.2. Water level control strategy

The flow relationship between each lake calculated by the model is as follows Fig.2:

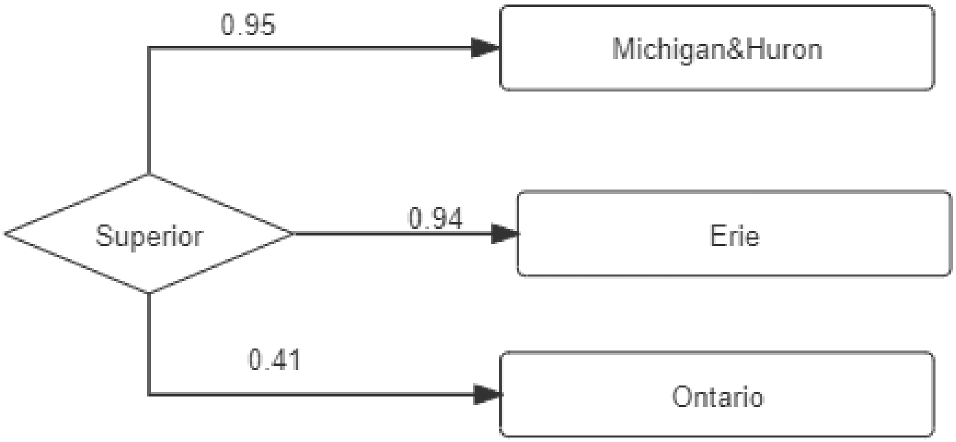

Figure 2. Calculation results of inter-lake discharge relationship

It can be seen from the above results that the coefficients of the influence of the outflow of water from Lake Superior on the other three lakes are 0.95, 0.94 and 0.41, respectively. By bringing the above coefficients into the model of lake discharge relationship, it can be concluded that by controlling the discharge of two DAMS, the strategy of achieving the optimal water level of each lake can be achieved.

3.3. Monte Carlo simulation analysis

Through Monte Carlo simulation, the model has performed sensitivity analysis on the outflow changes of the two control DAMS. The simulation generated a large number of possible effluent combinations and observed the effects of these changes on water levels and stakeholder satisfaction. Here are the steps for analysis using Monte Carlo simulation:

(1) Define input variables: Identify the outflow from the two control DAMS as input variables.

(2) Set a probability distribution: Assume that the input variables follow a uniform distribution, and define the range of variation of the variables.

(3) Generate a random sample: Generate a large number of random samples from the probability distribution of the input variables.

(4) Run the control algorithm: Run the control algorithm on each sample to get the output indicator.

Measure the sensitivity of the simulation by the difference between the pi value obtained from the simulation and the real PI value. If the resulting pi value from the simulation is closer to the true PI value, the sensitivity of the simulation will be reduced, and vice versa. We run the Monte Carlo simulation of the model in this problem and get pi=3.178620112107486, which shows that, on the whole, the results of the model are relatively stable and reliable. To make it easier to see, here plotted histograms representing the probability distribution of outflows from the two dams.

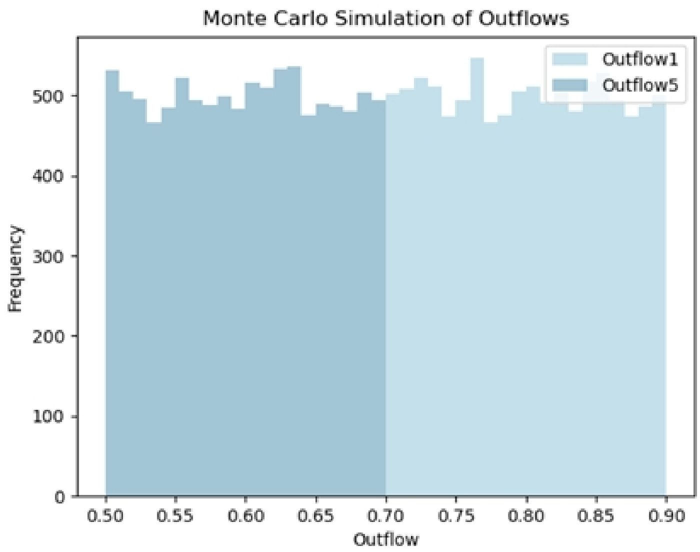

Figure 3. Histogram of the probability distribution of outflows from two dams

4. Analysis and discussion

4.1. In-depth analysis and discussion of results

Through in-depth analysis of the findings, the complexity of water level regulation in the Great Lakes and its importance to regional water resources management will be able to be better understood .

Sensitivity analysis also reveals the model's response to changes in the control dam outflow and environmental factors. The results of Monte Carlo simulations show that the model has a low sensitivity to these changes, which means that the model can maintain relatively stable predictions even when the dam outflow changes or environmental conditions fluctuate. This finding is crucial for the development of adaptive water management strategies, as it shows that the model can still provide reliable predictions under uncertain conditions, thereby supporting decision makers to make more robust water management decisions. This low sensitivity not only reflects the robustness of the model, but also highlights its value in practical applications, especially in the face of climate change and changing hydrological conditions. These findings enhance the applicability of the model in the field of water resource management and lay a solid foundation for future research and practice.

4.2. Practical application and policy recommendations of the model

The application results of the model show that a balance can be found between protecting the ecosystem and meeting the needs of economic development through reasonable regulation of water level. Therefore, based on the results of the model, the following policy recommendations can be made:

(1) Water resources management: It is recommended that relevant departments adopt the model proposed in this study to guide the regulation of water levels in the Great Lakes. By integrating genetic algorithm and network flow model, the model can effectively predict the trend of water level change and evaluate the effect of different regulation strategies, which is helpful to achieve the balance of multi-interests. Given the low sensitivity shown by the model, it is also recommended to adjust the water level regulation strategies according to seasonal changes and hydrological conditions to adapt to the needs of different time periods while ensuring the health of the ecosystem.

(2) Environmental monitoring: Strengthening monitoring of environmental factors: It is recommended to strengthen monitoring of environmental factors, including but not limited to changes in meteorological parameters such as precipitation, evaporation and temperature, as well as hydrological variables such as upstream inflow of water. This will help to better understand and predict the impact of these factors on lake water levels.

Data Integration and Analysis: A comprehensive data collection and analysis system is recommended for integrating environmental data from different sources to support more accurate forecasting and decision making.

(3) Risk assessment: Periodic risk assessment: It is recommended that relevant authorities conduct regular risk assessments to identify and mitigate potential water management risks, such as the likelihood of extreme events such as floods and droughts.

Scenario simulation: Using models to simulate different scenarios to assess the impact of various management measures on water level regulation, to ensure that the most appropriate response strategy can be adopted in the face of uncertainty.

5. Conclusion

In conclusion, this study presents an integrated approach that combines genetic algorithms and network flow models to optimize water level regulation in the Great Lakes, addressing the challenges posed by climate change and anthropogenic activities. By effectively balancing the interests of diverse stakeholders—including agriculture, industry, and ecological conservation—this model provides a scientific basis for sustainable water resource management. The robust data collection and preprocessing, coupled with advanced modeling techniques, enable accurate predictions of water level changes and comprehensive income assessments for stakeholders.

The findings indicate that the optimal water levels identified through the genetic algorithm align with the predictions from the network flow model, reinforcing the credibility of the results. Furthermore, the low sensitivity evident in the Monte Carlo simulations suggests that the model can reliably adapt to varying hydrological conditions and management strategies. This adaptability underscores the model's potential as a practical tool for policymakers, facilitating informed decision-making in response to changing environmental conditions. Ultimately, this research not only contributes to the understanding of water level dynamics in the Great Lakes but also lays a foundation for future studies aimed at enhancing the resilience and sustainability of freshwater resource management amidst climate variability.

References

[1] Lü Haifeng, et al. A Hybrid Water Level Prediction Model Based on CNN-LSTM-Attention and Autoregressive Techniques[J]. Water Resources and Hydropower Engineering, 2024, 55(06): 16-31.

[2] Long Yuan-nan, et al. Research on a Water Level Prediction Scheme Using the IPSO-EGA-LSTM Model for Dongting Lake[J]. Resources and Environment in the Yangtze Basin, 2024, 33(06): 1262-1272.

[3] Tan Kun, Huang Qianzi. Water Level Prediction Technology Based on Multi-Feature Spatiotemporal Fusion Networks[J]. Water Transport Engineering, 2024, 05: 151-155.

[4] Tian Songyi, et al. Adaptive Robust Prediction of Groundwater Levels Based on an LSTM Network with Attention Mechanism Fusion[J]. Engineering Science and Technology, 2024, 56(01): 54-64.

[5] Huang Ying, et al. Research and Implementation of an Efficient Long Sequence Water Level Prediction Model[J]. Journal of University of Electronic Science and Technology, 2023, 52(04): 595-601.

[6] Zhang Sen, et al. Research on Water Consumption Rate Calculation of Hydropower Stations and Its Impact on Water Level Prediction[J]. Hydropower Generation, 2023, 49(12): 94-100.

[7] Chen Huanliang, et al. Research on Water Level Prediction of Karst Springs Based on Genetic Algorithm Optimized BP Neural Network[J]. Journal of Water Resources and Water Engineering, 2022, 33(04): 50-57+63.

[8] Yu X P, Qiu D S, Cao M. Application of machine learning algorithms in predicting water levels and flow rates in Jinan Spring Group[J]. Shandong Water Conservancy, 2022, (11): 7-9.

[9] Chen Juan, Zhong Ping'an, Xu Bin. Risk Analysis of Reservoir Flood Control Scheduling Based on Stochastic Differential Equations[J]. Journal of Hohai University (Natural Sciences), 2013, 41(05): 400-404.

[10] Tao Xidong. Water Quality Management System in the Great Lakes Region of the United States and Canada: Experience and Insights[J]. Social Sciences, 2009, 06: 25-32+187.