Research On Optimisation and Regulation of Water Level in The Great Lakes Based on Simulated Annealing-Dynamic Programming Algorithm

Fuzhi Zhan *

School of Electric Power, South China University of Technology, Guangzhou, China, 510641

* Corresponding Author Email: epcopy@mail.scut.edu.cn

Abstract. Shipping, fishing, power generation... Lakes are crucial to ecosystems and increasingly impact human life. Rising conflicts over lake resource use highlight water management issues. This paper simulates Great Lakes water level changes to maximize benefits for all stakeholders. Two models are established: Model I: Optimal Water Level Estimation Model; Model II: Water Level Dynamic Regulation Model. This paper implemented a simulated annealing strategy to develop an optimal water level estimation model; by incorporating the historical water levels of the Great Lakes basin and the needs of various stakeholders, closely approximated and determined the optimal monthly water levels for each lake; the model's validity was confirmed with a Pearson correlation coefficient of 0.828. Based on the optimal water level estimation model, This paper utilized the concept of dynamic programming to analyze data on precipitation, evaporation, river flow, and total water usage. By incorporating theories of regulatory time delay and supply water volume lag factors, we established a dynamic water level regulation model, thus ensuring that the water levels of the Great Lakes remain within optimal limits.

Keywords: Dynamic Water Level Regulation Model, Simulated Annealing Algorithm, Dynamic Programming Model, Water Level Control.

1. Introduction

The Great Lakes of the United States and Canada constitute the world's largest fresh-water lake system, providing crucial water resources for numerous cities. Their usage is diverse, encompassing fishing, recreation, power generation, drinking water, and shipping. Given the involvement of two federal governments, eight states, and over 200 sovereign tribes [1], resource management of the Great Lakes involves numerous stakeholders from various sectors. Among these, regulating lake water levels is a crucial task influenced by multiple factors, including temperature, wind, precipitation, evaporation, and seasonal climate variations [2]. Despite the presence of certain control structures such as the Soo Locks and the Moses-Saunders Power Dam, predicting the impact of rainfall, evaporation, and other natural factors remains challenging. This complex and dynamic environment makes addressing lake management challenges challenging due to the intertwined demands and uncertainties among various stakeholders [3].

There are many scholars who have studied lake level problems. Among them, the future lake levels predicted by Seglenieks Frank et al. were calculated using data from the North American portion of the Coordinated Regional Downscaling Experiment (CRDE). The final lake level results were correlated with changes in global average temperatures of 1.5 °C, 2.0 °C, 2.5 °C and 3.0 °C [4]. Wei Yanling et al. used Qilu Lake as the study area and constructed BMA (VC, BP, SVR) and BMA (BP, SVR) models based on Bayesian statistical theory by utilizing meteorological and hydrological factors and the water level data in the previous January with a view to improving the accuracy of lake level prediction in long time series [5].

Scholars for the seasonal pattern of change of lake water level, water level prediction and factors affecting the change of lake water level and other aspects of a large number of studies, and research results applied to the specific practice, for the lake ecosystem function and regional socio-economic development is of great significance. At present, scholars have done less research on the dynamic regulation of lake water level, based on which this paper takes into account the expectations of different stakeholders to estimate the optimal water level of the lake at various times of the year. According to the inflow and outflow data of the lake, the algorithm to maintain the optimal water level of the lake is modeled to understand the sensitivity of the algorithm. Based on the data, we determine the level of satisfaction of the stakeholders with the new methodology and analyse the sensitivity of the algorithm to environmental conditions (precipitation, winter snowfall, hailstorms, etc.) to improve the study of water levels in lakes.

2. Data Description

2.1. Data Collection

Data on precipitation, lake levels, runoff, evapotranspiration, and water consumption of the U.S. Great Lakes were collected using the U.S. Mathematical Modeling website, Joint Commission International, NOAA - Great Lakes Environmental Research Laboratory, Great Lakes Coordinating Committee and Great Lakes Commission. The data links are shown in Table.1.

Table.1. Data source website

	Web Link
Data Source l	https://www.contest.comap.com/undergraduate/contests/mcm/login.php
Data Source 2	https://ijc.org/en/what/water-levels
Data Source 3	https://www.glerl.noaa.gov/ahps/mnth-hydro.html
Data Source 4	https://www.greatlakescc.org/en/coordinating-committee-products-and-datasets/
Data Source 5	https://waterusedata.glc.org/index.php

2.2. Data Pre-processing

The lake level and river flow data were screened for outliers and the normality test was performed on each monthly lake level sample. The results showed that the data conformed to a normal distribution with a 95% confidence level. Anomalous data with lake level fluctuations exceeding 0.9 meters were cleaned as required. Subsequently, the monthly average water levels of the lakes were calculated and analyzed by line graphs, which showed that there was a trend of continuous changes in water levels. Data on precipitation, evapotranspiration, runoff and water consumption were processed for outliers, and unreasonable data were removed, resulting in a final analysis of the composite data from 2012 to 2020.

3. Optimal Water Level Estimation Model

3.1. Introduction to the Simulated Annealing Model:

The simulated annealing algorithm is a general-purpose stochastic search algorithm inspired by the annealing process of solids[6]. This probability-based algorithm mimics the process of solids being heated to high temperatures and then slowly cooled. During the heating phase, the disorder within the solid's particles increases, raising the internal energy; whereas during slow cooling, the particles gradually become ordered. At each temperature stage, the system reaches equilibrium, ultimately reaching the ground state at room temperature, with the internal energy reduced to its minimum value. This algorithm, based on the similarity between optimization problems and the physical annealing process, achieves global optimization by appropriately controlling the temperature reduction process.

3.2. Objective Optimization and Constraint Setting

3.2.1 Constraints

This paper models the impact of Great Lakes water level data on various stakeholders from six different perspectives: shipping companies, fishery workers, agricultural and livestock workers, urban residents, environmentalists, and hydropower centers.

First, this paper classified the demands of various stakeholders regarding the water level conditions of the Great Lakes into four categories, summarizing the following requests, each providing upper and lower water level limits and satisfaction function requirements:

(1) Requirements for high water levels only: Agricultural and livestock workers in areas near the Great Lakes, and residents of urban communities along the lakes.

(2) Requirements for low water levels, water quality, and stable flow: Fishery workers in aquaculture.

(3) Requirements for water levels to be within a satisfactory range, neither too high nor too low: Shipping companies, which have minimum water level requirements for vessel navigation and face risk crises due to the highest water levels affecting infrastructure construction such as ports near the lakes.

(4) Requirements for water level differences during unique time periods: Hydropower centers require differences in water levels between the sides of dams during peak electricity usage periods or after the rainy season.

$$y_t = \sum_{i=1}^{n=4}\left(1 - \frac{1}{e^{\alpha|X_g - X_r|}}\right) + \beta e^{(\gamma_1|X_{4r} - X_{4g}| + \gamma_2|X_{1r} - X_{1g}|)}\ln\left(\prod|t - t_{rain}| + 1\right) \tag{1}$$

Where y_t represents the sum of satisfaction of all parties of the Great Lakes in the month, X_{ig} represents the target satisfaction level of Lake i, X_{ir} represents the actual water level of Lake i, α_i represents the satisfaction coefficient of the shipping stakeholders (i.e., the third category) of Lake i, β represents the satisfaction coefficient of the power generation stakeholders (i.e., the fourth category) of Lake Ontario, γ represents the power generation capacity of the two dams. According to statistics, the average annual power generation of the Moses-Saunders Dam is 1,050 megawatt-hours, and the average annual power generation of the St. Lawrence Seaway compensating works is 1,400 megawatt-hours, with γ_1 =0.4286, γ_2 =0.5714, t_{rain} represents the months of the rainy season. The matrix of correlation coefficients is shown in Table.2.

Table.2. Correlation coefficient matrix

Pearson Correlation Coefficient	$\lvert X_g - X_r\rvert$	y_t
$\lvert X_g - X_r\rvert$	1.000	0.828
y_t	0.828	1.000

Through the Pearson correlation coefficient analysis conducted on the Great Lakes water levels and the related satisfaction derived from formula (1), the correlation coefficient $\rho_{|X_g - X_r|, y_t}$ of 0.828 was obtained, indicating a significant correlation [7].

3.2.2 Calculation and Selection of Model Parameters:

X_g Determination: Through the small-sample Shapiro-Wilk test, a p-value between 0.1 and 0.2 was obtained, which is above the significance level. It was determined that the monthly water levels of the Great Lakes in different years follow a normal distribution. That is, the median water level of each month in different years reflects the most common phenomenon, which can be used to determine the target water level of satisfaction X_g.

After consulting various related literature, it confirmed that shipping and ensuring that the water levels of the Great Lakes remain within safe limits are of significant importance. Additionally, the requirements for ecological protection and the development and utilization of fishery resources need to be considered, hence the following parameters were determined $\beta = 0.3$.

Furthermore, based on the differences in the utilization of shipping, ecological protection, and resource use among the Great Lakes, with particular focus on Lake Michigan, which is connected to the Mississippi River, and Lake Ontario, which is connected to the Erie Canal and the Saint Lawrence River, parameters were determined $\alpha_1 = \alpha_3 = 100$, $\alpha_2 = 200$, $\alpha_4 = 240$.

3.3. Solving for the Optimal Water Level of Lakes

By using Matlab and the Simulated Annealing algorithm for programming solutions, and then substituting control parameters and key parameters, the aforementioned optimization algorithm was run 10 times to obtain its best result. The convergence curve of the objective function values is shown in Figure 1.

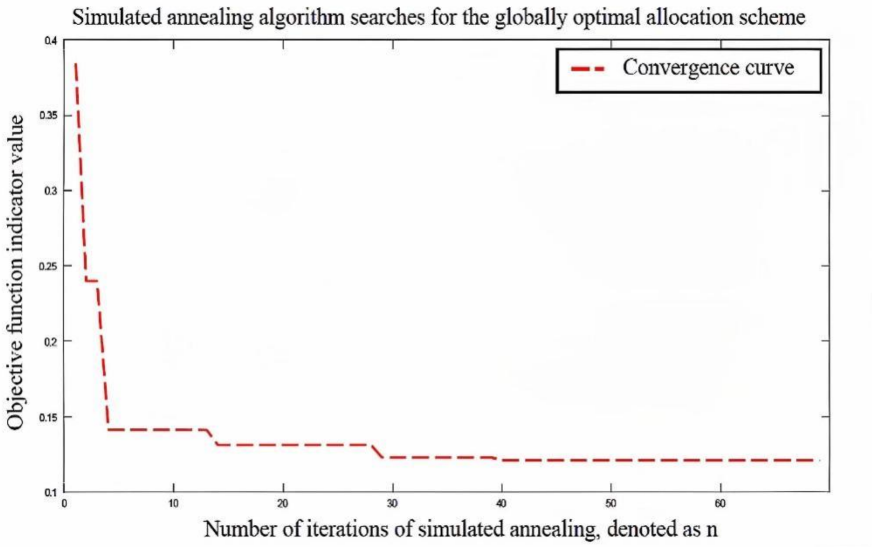


Figure 1. Simulated Annealing algorithm execution diagram

Based on the analysis, the optimal water levels for each lake for every month were determined, as presented in the Table.3 and Table.4 below.

Table.3. Optimal monthly water levels of the Great Lakes from January to June

lake	1	2	3	4	5	6
Lake Superior	183.492	183.499	183.452	183.545	183.657	183.594
Lake Michigan and Lake Huron	176.177	176.050	176.245	176.272	176.328	176.295
Lake St. Clair	174.872	174.905	174.932	174.930	174.967	175.092
Lake Erie	174.032	174.160	174.168	174.155	174.359	174.321
Lake Ontario	74.742	74.775	74.912	74.873	74.937	74.982

Table.4. Optimal monthly water levels of the Great Lakes from July to December

	7	8	9	10	11	12
Lake Superior	183.544	183.596	183.432	183.444	183.478	183.490
Lake Michigan and Lake Huron	176.351	176.256	176.237	176.124	176.138	176.005
Lake St. Clair	175.066	174.994	174.960	174.895	174.911	174.835
Lake Erie	174.322	174.269	174.246	174.170	174.159	174.177
Lake Ontario	75.011	74.877	74.968	74.842	74.883	74.714

4. Model II: Water Level Dynamic Regulation Model

4.1. Model Development with Dynamic Programming Algorithm

4.1.1 Introduction to Dynamic Programming Algorithm

Dynamic Programming (DP) is an algorithmic approach used for solving problems that exhibit overlapping subproblems and optimal substructure properties. It decomposes the original problem into several subproblems, solves these subproblems recursively, and stores their results, eventually obtaining the solution to the original problem. By establishing a network model of the Great Lakes, describing the water flow relationships helps the dynamic regulation model for the Great Lakes' water levels to more accurately capture the interactions and influences between water flows, enhancing the model's reliability and predictive capability. The network model is shown in Figure 2.

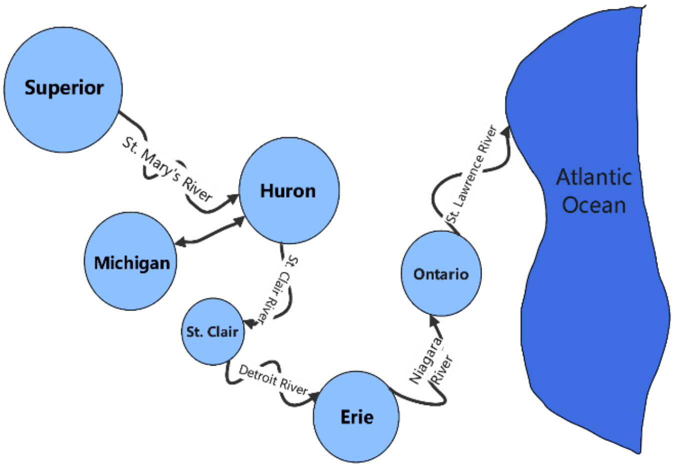


Figure 2. Network model diagram of the Great Lakes

4.1.2 Scenario Analysis and Model Development

Due to the flat terrain, the slowing of water flow, and the resulting delay in the adjustment of water quantity and water level by water regulation facilities in upstream areas, it takes longer for significant effects to occur. The lag time can range from several hours to several days, depending on factors such as the scale, performance, and size of the watershed. The climate conditions in the Great Lakes region also affect the lag time, with the mild and humid climate helping to reduce the lag time[8]. The seasonal variation rate of water levels in the Great Lakes is shown in Figure 3.

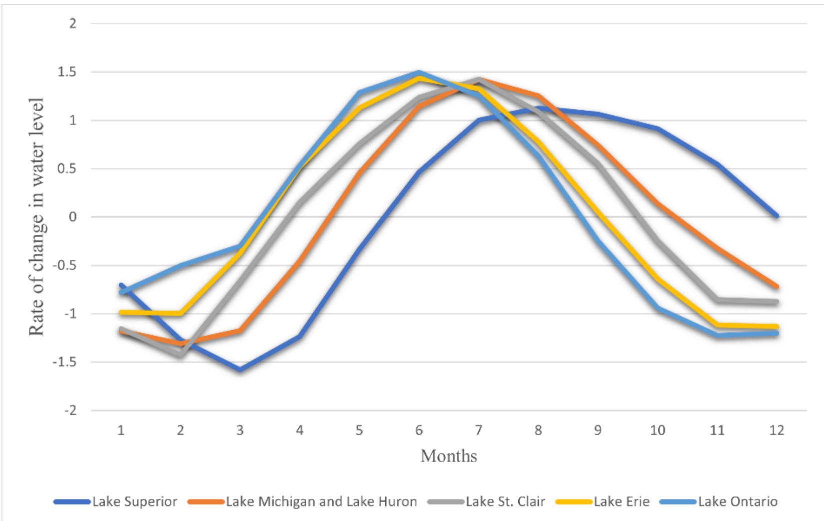


Figure 3. Seasonal variation rate of water levels in the Great Lakes

When considering the lag effect of water transfer, it can be incorporated into the hydrological model. A common method is to introduce a time delay parameter τ to describe the time required for water to transfer from upstream to downstream in the hydrological model. When establishing a hydrological model, the following factors can be considered:

River Transport Delay:

$$\frac{dV}{dt} = [R(t-\tau_1) - E(t)] \times S_{basin} - Q(t) - U(t) \tag{2}$$

Lake Water Accumulation Delay:

$$\frac{dV}{dt} = [R(t) - E(t)] \times S_{lake} + P(t-\tau_2) - Q(t) - U(t) \tag{3}$$

For the lake level on i day:

$$W_i = W_{i-1} + \left(R(i) - E(i) + \frac{P(i-\tau) - Q(i) - U(i)}{S}\right) \times \delta \tag{4}$$

Where the symbols are defined as shown in Table.5.

Table.5. Symbol explanation for Model II

Symbol	Description	Unit
$R(t)$	Precipitation on t day	m
$U(t)$	Water Use Totals on t day	m^3
$E(t)$	Evaporation on t day	m
$Q(t)$	Total Runoff Outflow Volume on t Day	m^3
$P(t)$	Total Runoff Recharge Volume on t Day	m^3
$V(t)$	Volume of Water in the Lake or River on t Day	m^3
τ	Delay Time of River Transport or Accumulation Time of Lake Water Volume	day
W_i	Lake Water Level on t Day	m
δ	Loss Factor Set Based on Leakage, Overflow, etc.	/

4.2. Model Solving: Finding Strategies

Analyzing the maps and cross-sectional diagrams of the Great Lakes, it is observed that there are four main lakes between two dams, each connected by rivers of varying lengths. Therefore, the dynamic regulation of water quantity between lakes by controlling the dams needs to consider issues of lag in water resource transfer and variations in river flow velocity. To address these, dynamic programming variables are introduced for an appropriate solution. ν represents the scheduling scale factor of the dam, u represents the threshold coefficient of the lake based on the optimal water level, Max_1 represents the maximum drainage capacity of the Compensating Works of the Soo Locks, Max_2 represents the maximum drainage capacity of the Moses-Saunders Dam, The water level needs to be dynamically adjusted between the upper and lower thresholds to maximize benefits[9]. Based on references, "There is more recent concern for the management of the water level for Lake Ontario." The two dams divide the Great Lakes and their downstream areas into three regions: the Superior Lake region, the central three-lake region including Lake Michigan, the Ontario Lake region, and its downstream area of the St. Lawrence River. Based on the different importance of each region, w is introduced as a risk coefficient to determine the sequence of decisions for water level changes in the three regions. For the Superior Lake region and the central three-lake region: $w_1 = w_2 = 0.3$, For the Ontario Lake region and its downstream area of the St. Lawrence River: $w_3 = 0.4$, so that $w_1 + w_2 + w_3 = 1$.

From the above formulas, the water level for the current day W_i can be obtained. The daily water level difference can be calculated from the optimal water level, and the decision priority of the three regions can be determined by the water level differences of each region separately.

$$\triangle W_i = |W_i - W_{best}| \tag{5}$$

$$Sort(\Delta W_{1i} \times w_1, \Delta W_{2i} \times w_2, \Delta W_{3i} \times w_3) \tag{6}$$

According to the priority, for the most important region, which is the one with the highest product of water level error and risk coefficient, its water situation is the most critical, and its decision should be given the highest priority. The water level differences under different conditions will lead to changes in priority, thus achieving dynamic adjustment.

For lakes where the water level exceeds the optimal threshold, the opening or closure of dam should be implemented, with the corresponding actions υ, determined based on the magnitude of the water level difference ΔW_i.

$$\nu_i = \frac{1}{2} - \frac{1}{2\pi} \arctan[50\pi(\Delta W_i - 0.01)] \tag{7}$$

Due to the time lag in water transfer and variations in river flow velocity, the dynamic regulation of dam's experiences temporal and spatial delays. Lakes situated closer to the dam receive adjustments more rapidly and in greater volumes. Based on a review of relevant literature, a matrix illustrating the flow rate lag is provided:

$$C_1 = \begin{bmatrix} 1.00 & 0.60 & 0.25 & 0.15 \\ 1.00 & 0.50 & 0.30 & 0.20 \\ 1.00 & 0.33 & 0.33 & 0.33 \end{bmatrix} \qquad C_2 = \begin{bmatrix} 0.04 & 0.06 & 0.10 & 0.80 \\ 0.10 & 0.10 & 0.20 & 0.50 \end{bmatrix} \tag{8}$$

The state transition equations for Lake Superior, Lake Michigan, Lake Huron, and Lake Erie are as follows:

$$W_i' = \begin{cases} W_i + \dfrac{\nu \times Max_1 \times C_{1(j,t-\tau)}}{S_j} & W_i \leqslant W_{best}(1+u) \\ W_i - \dfrac{\nu \times Max_1 \times C_{1(j,t-\tau)}}{S_j} & W_i \geqslant W_{best}(1+u) \\ W_i & W_{best}(1-u) < W_i < W_{best}(1+u) \end{cases} \tag{9}$$

The state transition equation for Lake Ontario is as follows:

$$W_i' = \begin{cases} W_i + \dfrac{\nu \times Max_2 \times C_{2(j,t-\tau)}}{S_j} & W_i \leqslant W_{best}(1+u) \\ W_i - \dfrac{\nu \times Max_2 \times C_{2(j,t-\tau)}}{S_j} & W_i \geqslant W_{best}(1+u) \\ W_i & W_{best}(1-u) < W_i < W_{best}(1+u) \end{cases} \tag{10}$$

Utilizing Matlab software, set variable parameters. u =0.0003, $Max_1 = 1.339 \times 10^9$, $Max_2 = 1.227 \times 10^9$, $S = [8.21 \times 10^{10} \ \ 1.176 \times 10^{11} \ \ 2.57 \times 10^{10} \ \ 1.9 \times 10^{10}]$, Calculate for the period 2012-2020, obtaining the following dynamic water level adjustment images shown in Figure 4.

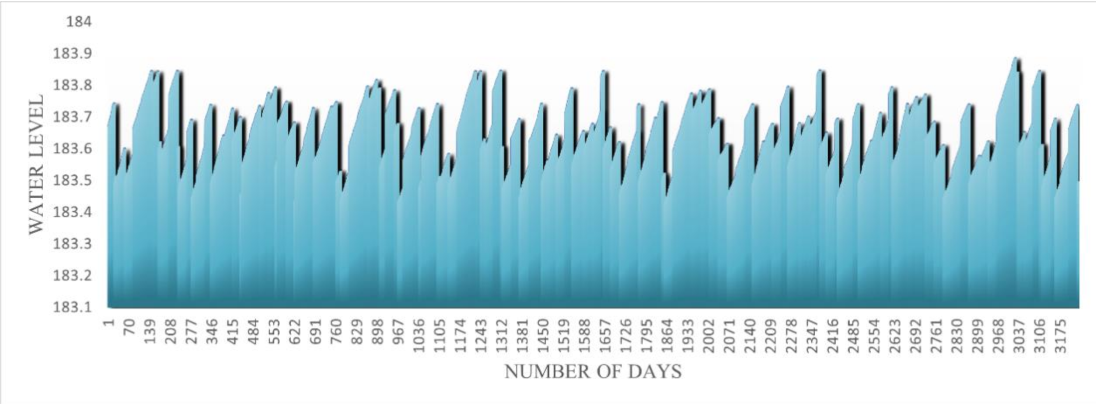


Figure 4. Dynamic water level adjustment graph of Lake Superior

As evident from the graph, after dynamic adjustment, the fluctuation in water levels significantly decreases, with variations within [a certain range], rendering it more stable.

4.3. Resolution of Stakeholder Satisfaction in 2017

4.3.1 Perspective 1: Satisfaction Fitting Formula

Calculated from Formula (1): The satisfaction fitting formula determines the satisfaction of stakeholders based on the actual water level in 2017 and the water level adjusted dynamically. Plotting satisfaction on the line graph shown in Figure 5 for comparison reveals that dynamically adjusted satisfaction is higher than actual satisfaction.

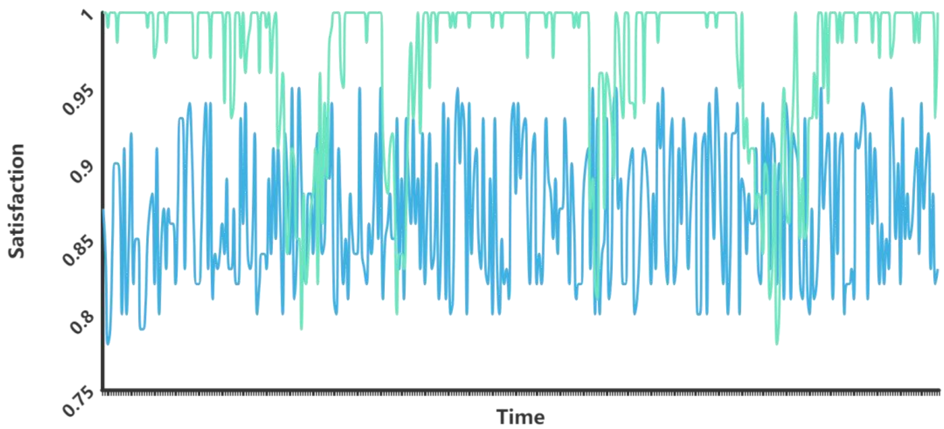


Figure 5. Comparison of actual satisfaction and regulated satisfaction in 2017

4.3.2 Perspective 2: Mahalanobis Distance

Using Mahalanobis distance to calculate the similarity between the actual water level and the dynamically regulated water level in 2017 with the optimal water level[10]. The results are shown in Table.6.

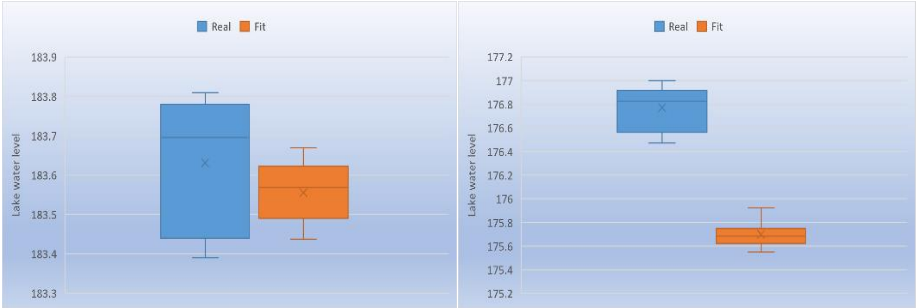
Table.6. Mahalanobis Distance to the optimal water level

lake	D_{real}	D_{fit}
Lake Superior	4.904	2.296
Lake Michigan and Lake Huron	4.189	2.970
Lake Erie	5.158	2.725
Lake Ontario	4.640	2.915

The Mahalanobis distance triad table indicates that the dynamically regulated water level is closer to the optimal water level and, hence more similar.

4.3.3 Perspective 3: Box Plot

The box plot is a chart that describes the shape, central tendency, and dispersion of data by displaying five statistical measures (minimum, lower quartile, median, upper quartile, and maximum). The distribution of Great Lakes water levels in 2017 is shown in Figure 6.



(a)Lake Superior (b)Lake Michigan and Lake Huron

(c)Lake Erie (d)Lake Ontario

Figure 6. 2017 Distribution of water levels in the Great Lakes

Compare the actual and dynamically regulated water level data of the Great Lakes in 2017 by plotting box plots. From the graph, it is evident that the water levels after dynamic adjustment are more concentrated and stable compared to the actual levels.

It can be concluded that the water level data after dynamic adjustment in 2017 demonstrates greater stability compared to the actual data, resulting in higher stakeholder satisfaction. This outcome also indirectly indicates the effectiveness of the formulation of Formula (3).

5. Conclusions

In this paper, a set of water level dynamic management models for the Great Lakes region is designed to realize the estimation of optimal water level and the formulation of dam control strategies by introducing time lag and recharge lag factors and assessing the regional importance based on the correlation of interests. The model was verified by actual data and showed good prediction accuracy and practical applicability. Meanwhile, the model design was continuously optimized through literature research and sensitivity analysis to ensure its robustness.

Future research will focus on optimizing the model parameter settings, improving the ability to predict future changes, expanding the multi-objective optimization function, and exploring the cross-basin collaborative management mechanism. Through the introduction of machine learning techniques, long-term data analysis, multi-interest considerations, and international cooperation, we aim to further improve the model performance to better serve practical water resources management and decision-making.

References

[1] Selzer M, Jonas L J, Pebbles V, et al.Crossing boundaries between science and policy: Two case studies illustrate the importance of boundary organizations in the Great Lakes Basin[J].Journal of Great Lakes Research,2020,46(4):933-946.

[2] Dogan M. How does the North Atlantic Oscillation affect the water levels of the Great Lakes with regard to hydro-climatic indicators?[J]. Theoretical and Applied Climatology,2016,126(3-4):597-609.

[3] Dang Guanchao. Ecological Consequences of St. Lawrence Seaway Development and U.S. Responses (1954-2008) [D]. Shandong Normal University,2023.

[4] Frank S, André T. Future water levels of the Great Lakes under 1.5 °C to 3 °C warmer climates[J]. Journal of Great Lakes Research,2022,48(4):865-875.

[5] WEI Yanling, WANG Fei, CHEN Jing, et al. Simulated prediction of water level in highland lakes based on Bayesian stochastic method[J]. Water Resources Planning and Design,2024, (02):103-107.

[6] Mehrizi A A, Ghanbari R, Sadeghi S, et al.Solving continuous quadratic programming for the discrete balanced graph partitioning problem using simulated annealing algorithm hybridized local search and multi-start algorithms[J].Applied Soft Computing,2024,165:112109-112109.

[7] LIU Lirui, LI Peiqiang, CHEN Jiayu, et al. Hybrid energy storage capacity allocation based on Pearson correlation analysis and T-tFD under VMD decomposition[J]. China Electric Power,2024,57(07):82-97.

[8] YU Zhou, JIANG Tao, FAN Penghui, et al. multi-timescale prediction of lake water level based on EMD-DELM-LSTM combined model[J]. Journal of Changjiang Academy of Sciences,2024,41(06):28-35.

[9] ZHU Xi, LI Shiqiong, WANG Zhenghua, et al. Study on the dynamic control strategy of water level in joint scheduling operation during flood season in the middle and lower reaches of Wujiang River [J/OL]. China Rural Water Conservancy and Hydropower,1-13[2024-09-24].

[10] Astivia O L O. A method to simulate multivariate outliers with known Mahalanobis distances for normal and non-normal data[J]. Methods in Psychology,2024,11:100157-100157.