

Fuzzy logic and multi-objective optimization-based water level regulation model for the Great Lakes of North America

Xinyi Huang^{1,*}, Hugen Lv^{2,#}, Junlei Zhang^{2,#}

¹ Department of Electronic Information Engineering, Beijing Jiaotong University, Weihai, China, 264200

² Department of Computer Science and Technology, Beijing Jiaotong University, Weihai, China, 264200

* Corresponding author: 21721049@bjtu.edu.cn

#These authors contributed equally.

Abstract. In this study, a series of models was developed to ascertain and maintain the optimal water levels of the Great Lakes, aiming to reconcile the diverse requirements of stakeholders. Initially, a fuzzy logic algorithm was applied to calculate the weighting coefficients of five stakeholder categories, facilitating the creation of a multi-objective optimization framework to harmonize all demands and thereby determine the optimal water level for each lake. Subsequently, an analysis of the geographical and meteorological data of the Great Lakes informed the construction of a dynamic equilibrium model for water levels, grounded in the calculation of influx and efflux rates to sustain the optimal water level. Following this, simulated tests were conducted using historical data, with Principal Component Analysis (PCA) utilized to assess the model's pre-optimization and post-optimization, providing a clear comparison of outcomes. This article aims to develop and apply optimization models to balance the diverse needs of stakeholders in the Great Lakes, manage water levels scientifically, enhance water resource management efficiency, and promote regional sustainable development.

Keywords: Multi-objective Programming; Analytic Hierarchy Process; Dynamic System Simulation.

1. Introduction

The Great Lakes are the largest group of freshwater lakes in the world, and changes in their water levels can significantly impact the lives of surrounding residents. G. A. Meadows et al. (1997) noted that fluctuations in Great Lakes water levels alter wave energy, which subsequently changes the impact on shorelines, affecting the lives of coastal residents. As temperatures change, the water levels of the Great Lakes also undergo corresponding variations, significantly impacting numerous ecosystems, industries, and coastal processes (Kayastha et al., 2022). Managing the water levels of North America's Great Lakes is a complex challenge that has been addressed through various methods over the years. Recent studies, such as those by Kayastha et al. (2022) and Zhu et al. (2020), have employed climate models and machine learning to predict water level fluctuations, emphasizing the need for innovative solutions in light of changing environmental conditions [1-2]. Building on this foundation, this paper introduces a cutting-edge model that utilizes fuzzy logic and multi-objective optimization to regulate water levels more effectively. The approach uniquely integrates stakeholder preferences with comprehensive geographic and meteorological analyses to establish a dynamic equilibrium model, which is then refined through simulations using historical data. (Data sources: <https://ijc.org/en/loslrb/watershed/water-levels>).

2. Establish the model of the optimal water level

2.1. Data analysis

Due to the ambiguous nature of the stakeholder demands for water levels, categorized as "high," "medium," and "low," the fuzzy logic method is employed to calculate the weight of each interested party. According to the data from Fagherazzi, given that stakeholders of each lake have varying interest needs, the weights of the corresponding stakeholders for each lake are initially calculated [3]. By minimizing the absolute value of the difference between the actual water level and the requirements of each interested party, multi-objective planning is achieved, resulting in the determination of the optimal water level for each lake. The result is shown in the Table 1.

Table 1. Data Description

Lake	Maximun	Minimum	25%	75%
Ontario	1745.1	173.75	174.09	174.67
Superior	177.45	175.57	176.06	176.77
Michigan	75.36	74.29	74.56	75.13
Erie	175.98	174.48	174.25	175.54
Clair	183.81	182.79	183.11	183.58

2.2. Fuzzy logic algorithm determines the weights

Define input and output variables: The input variables include the demand of environmentalists, fishermen, residents, hydroelectric companies, shipping companies, and the output variables are the weight coefficients ω_1 to ω_5 .

Fuzzy sets for input and output variables are established as follows: Three fuzzy sets—"low," "medium," and "high"—are added for the "water level" input variable. Similarly, three fuzzy sets—"low," "medium," and "high"—are designated for each output variable, which corresponds to the weight of stakeholders.

Make fuzzy rules: For example, for Lake Ontario, fuzzy rules are very important for environmentalists and fishermen, and less important for hydro power companies. Shipping companies and residents are less important [4]. Other lakes are similar.

Design a Fuzzy Logic system: Use MATLAB's Fuzzy Logic Toolbox. The corresponding weights for environmental protectors, fishermen and property owners, hydroelectric companies, shipping companies, and surrounding residents are obtained for different lakes respectively:

The respective weights for various stakeholders—including environmental protectors, fishermen and property owners, hydroelectric companies, shipping companies, and surrounding residents—are derived for each lake under consideration in the Table 2.

Table 2. _weight of the various stakeholders

Lake	environmentalist	Fishermen and owners	Hydroelectric power company	Shipping company	Manages shipping terminals and people living near ports
Ontario	0.30	0.30	0.20	0.10	0.10
Superior	0.18	0.18	0.27	0.18	0.18
Michigan	0.18	0.18	0.27	0.18	0.18
Erie	0.18	0.18	0.27	0.18	0.18
Clair	0.18	0.18	0.27	0.18	0.18

2.3. Multi-objective programming model

Define the objective function, one objective function for each stakeholder, f_1, f_2, \dots, f_n is the objective function, where each f represents the water level demand satisfaction of one stakeholder [5].

Environmentalists aim for Lake Ontario's seasonal high and low water levels to help maintain habitat for species and clean up static bays and tributaries. An analysis of the data found that high water levels occur in the summer (July, August and September) and low water levels in the winter (January, February and March). So

$$f_1: M_{summer} < M_{75}, M_{winter} < M_{25} \quad (1)$$

The goal of fishermen and property owners, who use facilities such as docks and boat launch ramps, want the water level to be moderate and stable, so

$$f_2: M_{25} < M_{fish} < M_{75} \quad (2)$$

The goal of hydro power companies, which want more control over water levels, is to use high water levels as storage systems to maximize flow during periods of high energy consumption.

The goal of shipping companies, who want the river to be high and still (no flow), and ships to be able to carry more cargo, raising the volume of transportation, and thus increasing economic returns. So

$$f_3, f_4: M_{hydroelectric} > M_{75} \quad (3)$$

People who manage shipping terminals or live near the Port of Montreal aim for lakes that are low and stable. So

$$f_5: M_{people} < M_{25} \quad (4)$$

$$Minimize = w_1f_1 + w_2f_2 + w_3f_3 + w_4f_4 + w_5f_5 \quad (5)$$

Table 3. Optimum water level and Objective function value

Lake	Optimum water level	Objective function value
Ontario	183.58	0.122536
Superior	175.54	0.283768
Michigan	174.67	0.275219
Erie	176.77	0.275725
Clair	75.13	0.472609

All the objective function values in the Table 3 are less than 0.5, which is a relatively small value, indicating that the water level found is a good balance between the needs of the various stakeholders.

3. Water level dynamic balance model

Firstly, leveraging the optimal water levels of each lake determined in the initial query, the inflow and outflow data for the lakes are obtained based on the geographical positioning and natural conditions of the Great Lakes. Subsequently, a dynamic water regulation algorithm is established.

The core of this algorithm is to adjust the outflow volume so that the water level is close to the optimal water level of the lake.

At the same time, the relationship between the average monthly discharge of the river adjacent to the lake and the water level of the adjacent lake is considered to avoid the negative impact caused by excessive dam storage and excessive human pumping.

3.1. A water level balance model is established

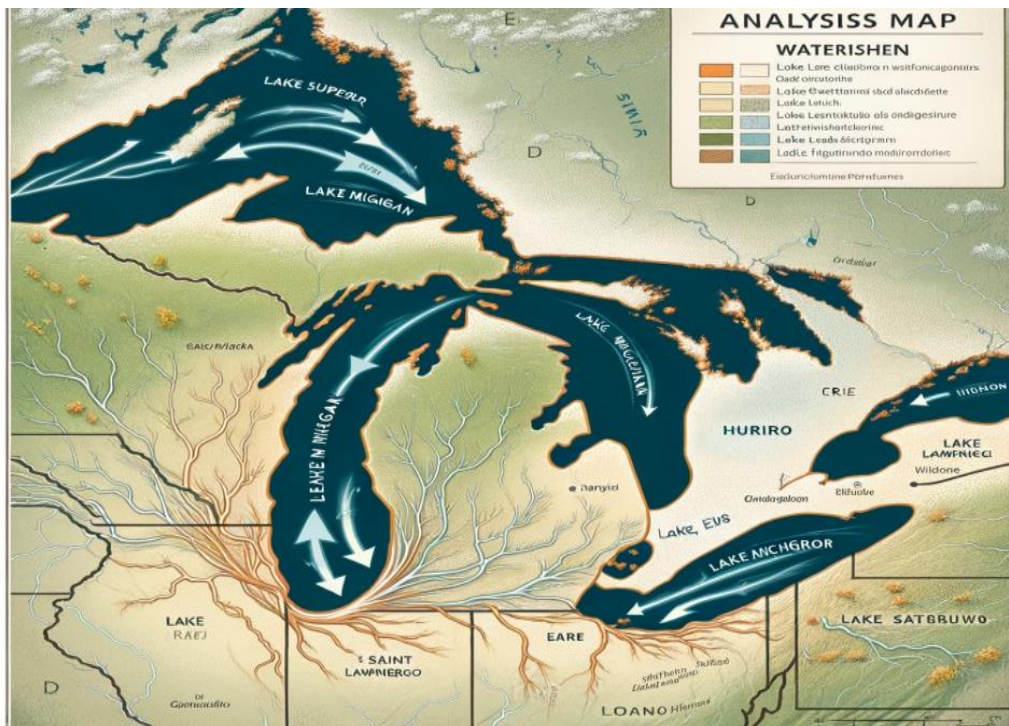


Figure 1. Flowchart of the Great Lakes

First, according to the fig.1, Lake Superior and St. Marys River flow to Lake Michigan-Huron, Lake Michigan-Huron and St. Clair River flow to Lake St. Clair, Lake St. Clair and Detroit River flow to Lake Erie, and Lake Erie and Niagara River flow to Lake Ontario. Eventually Lake Ontario flows to the St. Lawrence River.

In the subsequent phase of the research, the requisite parameters for the model, specifically the surface area of each lake, are ascertained. These calculations are conducted utilizing the area formula, based on the five system overview images provided by the Integrated Coastal Management (ICM)^[6] as shown in Table 4.

Table 4. The Great lakes Surface Areas

Name	Superior	Michigan-Huron	St. Clar	Erie	Ontario
Surface area(km2)	82100	117250	11400	25657	19000

The core of the model is based on the following difference equation, which calculates the change in the water level of the lake at any given time, accounting for precipitation, evaporation, and outflow. Let the amount of the lake level above the optimal level be L.

$$\frac{dV}{dt} = A * \frac{dh}{dt} \tag{6}$$

$$\frac{dh}{dt} = \text{inflow} - \text{outflow} + \text{precipitation} - \text{evaporation} \quad (7)$$

For Lake Superior, controlling its optimal water level only requires controlling the dam of the Soviet-style lock compensation project.

For Lake St. Clair, its water level is equal to the average monthly discharge of the incoming St. Clair River minus the average monthly discharge of the Detroit River plus rainfall - evaporation.

For Lake Erie, its water level is equal to the average monthly discharge of the incoming Detroit and Niagara rivers + the excess (L) by which Lake St. Clair exceeds its optimal level.

For Lake Ontario, its water level can be controlled by the dam on the right side of the figure, which has an inflow equal to Lake Erie's above-optimum flow (L) and the average monthly flow of the Niagara River.

Because water flows lower, add altitude constraints to the model: Lake Michigon-Huron cannot be lower than the monthly flows of the St.Clair and Detroit Rivers and the level of Lake Erie, because the former is at a higher elevation than the latter [7-8].

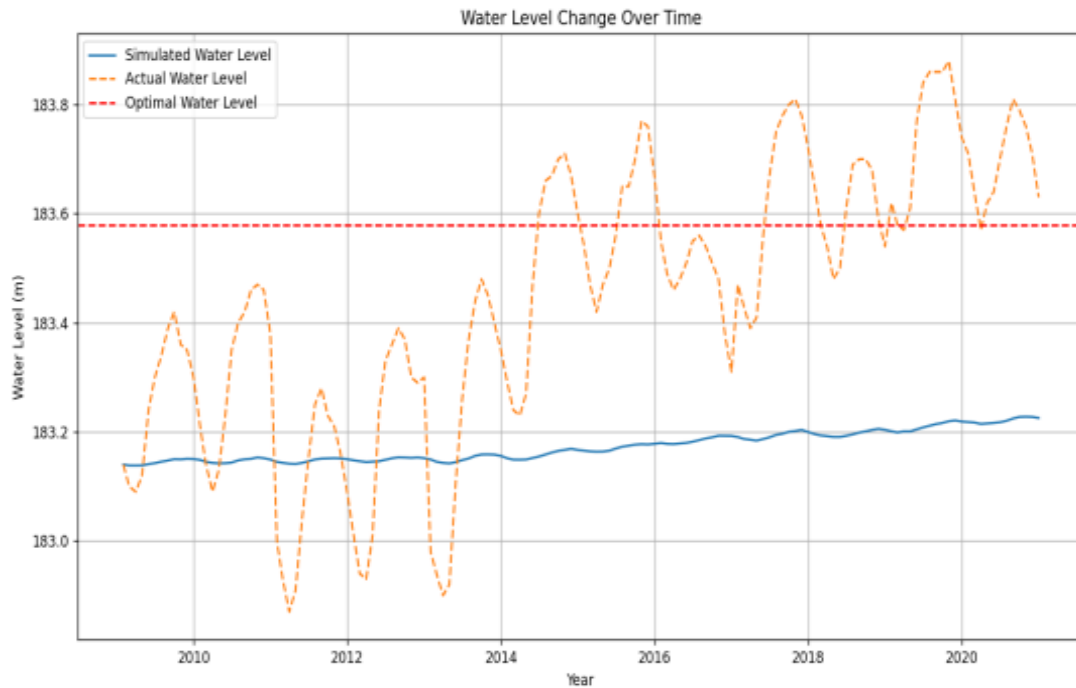


Figure 2. Simulation of Lake Superior

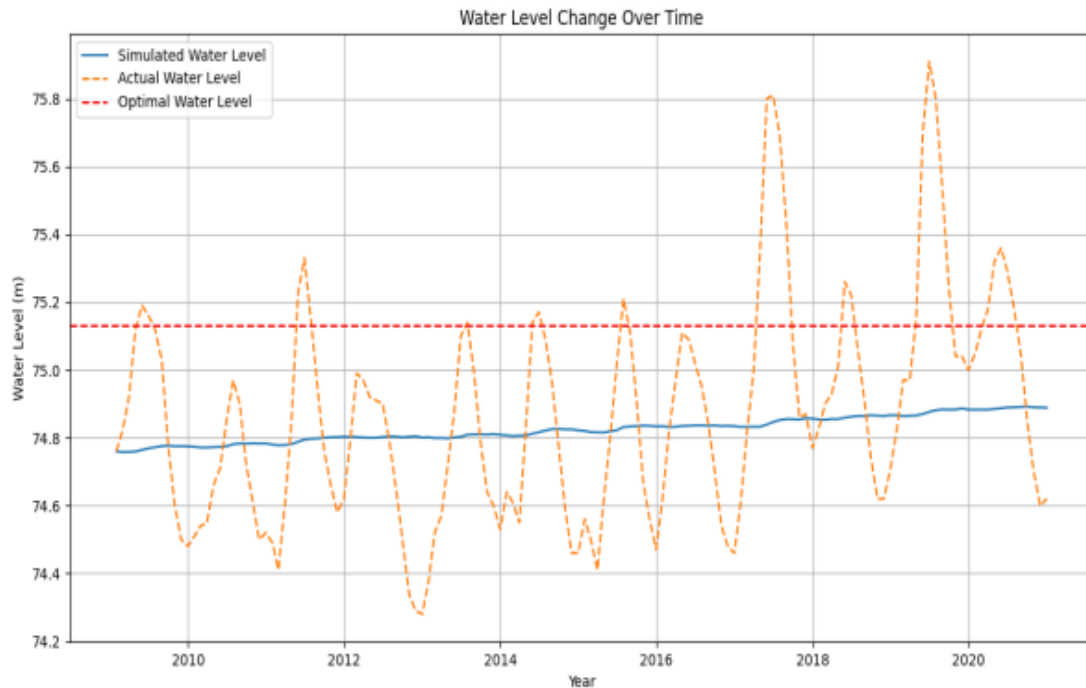


Figure 3. Simulation of Lake Ontario

"Actual Water Level" represents the observed Water Level data of the lake, and "Optimal Water Level" (dotted red line) represents the ideal water level of the lake. "Simulated Water Level" represents the lake water level calculated according to the dynamic equilibrium model of water level. This is a simulated value derived by simulating the input and loss of lake water (such as rainfall, evaporation and possible outflow). It can be seen from figs.2 and 3 that the trend of simulated values is the same as that of real values.

4. Principal component Analysis model

Principal Component Analysis (PCA) is a commonly used data dimensionality reduction technique that transforms the original data by a linear transformation into a set of new variables that are orthogonal to each other. These new variables are called principal components.^[9]

Principal components are arranged in descending order of variance, capturing the main variations in the original data. Principal Component Analysis (PCA) reduces the dimensionality of the original dataset by extracting the most significant information. This process not only diminishes storage and computational demands but also simplifies data understanding and visualization.^[10] To support the assertion that the simulated data would surpass the actual data in effectiveness, principal component analysis was chosen.

4.1. Sensitivity analysis

Through sorting out the parameters of precipitation, evaporation and water level change, a reasonable variation range for each parameter were defined. The sensitivity of the control algorithm to the key parameters (water level stability, water level height, seasonal variation), and the impact of these parameters on the adjustment strategy of dam outflow also have been analyzed.

4.2. Visualization of the PCA model

Principal component analysis visualization of the Great Lakes water level:

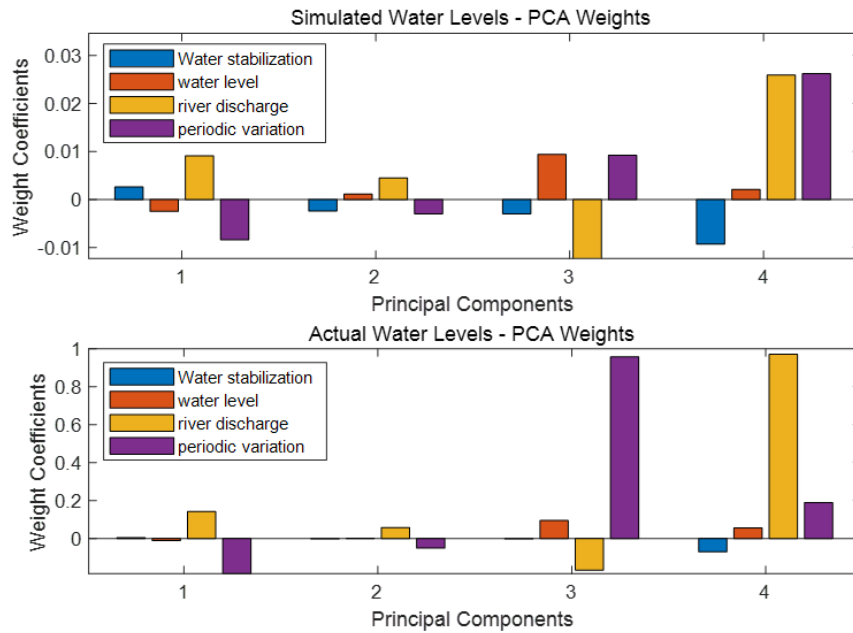


Figure 4. PCA weight coefficient diagram

This fig.4 shows the weight coefficient of four indicators (water level stability, water level height, river discharge, cycle variation) for different lakes (Superior, Michigan, Erie, Ontario, St. Clair) by principal component analysis (PCA). The numbers 1 to 4 on the X-axis represent the different principal components considered in the analysis. In PCA, these variables are the features in the original data set.

The weight coefficients (also called loadings) shown on the Y-axis indicate the importance of each metric in each principal component. A larger absolute value of the load signifies a greater contribution of that metric in defining the principal components. Positive loads indicate that the indicator is positively correlated with the principal components, whereas negative loads indicate a negative correlation. Different colors represent various metrics, and the PCA model generates distinct weights for each lake through these different metrics. The height and direction of the bars reflect the contribution of each variable to each lake after the PCA transformation. For example, if Lake Michigan has a high positive value on variable 1, this means that variable 1 has a strong positive effect in the first principal component in the Lake Michigan data.

Clearly, it is known from the graph that the simulated water level explains the overall variance better, which indicates that our simulated water level is better than the actual recorded water level.

5. Conclusion

This study presents a novel model for determining optimal water levels in the Great Lakes region, taking into account stakeholder aspirations and environmental dynamics. Utilizing a fuzzy logic algorithm, weight coefficients were assigned to five types of stakeholders, facilitating the development of a multi-objective programming model. Through this approach, an optimization function was established to balance competing demands and identify the most suitable water level for each lake within the system.

Additionally, a dynamic equilibrium model was constructed by analyzing geographical location and weather data, allowing for the calculation of inflow and outflow rates. By maintaining optimal water levels across all five lakes, this model ensures consistency between actual and desired conditions.

Simulation tests, spanning over 2107 years of historical data, evaluated the performance of the dynamic balance model under different hydrological conditions. Based on the support of a large number of data, the established model can be approximated to the real data without limitation, which greatly improves the prediction effect. Principal component analysis facilitated intuitive comparison

of results, highlighting the effectiveness of the optimization process. Furthermore, Lake Ontario was treated separately within the modeling framework, considering the unique interests of stakeholders and examining the advantages and disadvantages of the overall approach.

In conclusion, the integration of multi-objective programming, analytic hierarchy process, and dynamic system simulation offers a comprehensive strategy for optimizing water levels in the Great Lakes region. By prioritizing stakeholder aspirations and accounting for environmental variability, this model provides a foundation for sustainable water resource management in the area.

References

- [1] Kayastha Miraj B., Ye Xinyu, Huang Chenfu, Xue Pengfei. Future Rise of the Great Lakes Water Levels under Climate Change [J]. *Journal of Hydrology*, 2022, 612 (B): 128205.
- [2] Zhu, S., Lu, H., Ptak, M. et al. Lake Water-Level Fluctuation Forecasting Using Machine Learning Models: A Systematic Review [J]. *Environ Sci Pollut Res*, 2020, 27, 44807 – 44819.
- [3] Fagherazzi, L., Fay, D. and Salas, J. ‘Synthetic Hydrology and Climate Change Scenarios to Improve Multi-Purpose Complex Water Resource Systems Management. The Lake Ontario: St Lawrence River Study of the International Canada and US Joint Commission’ [J] *Water Resources Management IV*, 2007, 163 – 177.
- [4] Fry, L.M., Apps, D. and Gronewold, A.D. ‘Operational Seasonal Water Supply and Water Level Forecasting for the Laurentian Great Lakes’ [J] *Journal of Water Resources Planning and Management*, 2020, 146 (9).
- [5] Gronewold, A.D. et al. ‘Coasts, Water Levels, and Climate Change: A Great Lakes Perspective’, *Climatic Change*[J], 2019, 120 (4), 697 - 711.
- [6] Murty, P. S., Pandey, A., & Suryavanshi, S. Application of Semi-Distributed Hydrological Model for Basin Level Water Balance of the Ken Basin of Central India. [J] *Hydrological Processes*, 2014 28 (13), 4119 – 4129.
- [7] Sun, Z. Research on PID Control of Water Level in Pile Legs Based on Improved PSO. *IOP [J] Conference Series. Earth and Environmental Science*, 2020, 558 (2), 22019.
- [8] Moeini, M., Zahraie, B., & Sadeghi, F. Developing a New Lumped Monthly Water Balance Model for Estimating Groundwater Level and Runoff Volume. [J] *Sustainable Water Resources Management*, 2024, 10 (3), 113.
- [9] Lever, J., Krzywinski, M. & Altman, N. ‘Principal Component analysis’, [J] *Nat Methods*, 2017,14, 641 – 642.
- [10] Abdi, H., & Williams, L. J. Principal component analysis. [J] *Wiley Interdisciplinary Reviews: Computational Statistics*, 2020, 2 (4), 433 - 459.