Problem Chosen	2024 MCM/ICM Summary Sheet	Team Control Number
D		2411075

MPC: A Dynamic Water Ballet of the Great Lakes

Summary

The Great Lakes of the United States and Canada constitute the largest group of freshwater lakes in the world, and their water level management is of great importance to multiple stakeholders. The regulation of lake levels faces many challenges, including changing environmental conditions, conflicting needs of stakeholders, and the interaction of different lake levels. This paper established the Profit Consumption Status(PCS) optimal water level model and the Model Predictive Control(MPC) optimal water level control model, and conducted multi-dimensional experiments on the 2017 data to verify the effectiveness and stability of the models.

For the PCS optimal water level Model (Model I), we collected the profits and water consumption of the vast industries in each lake, and constructed the PCS weight (eq.(4)) according to the two factors and the social status of each industry, which can objectively reflect the important situation of each industry during the construction of the optimal water level. By analyzing the wishes of stakeholders, the value functions of growth type (eq.(2)) and intermediate type (eq.(3)) are constructed, and the value function curves of the five Great Lakes about their respective water levels are obtained by linear combination (figure (5)). And the optimal water level interval of the Great Lakes is obtained, like $[183.38, 183.68](m)$ for LAKE SUPERIOR.

For the MPC optimal water level control Model (Model II), we first establish the discrete state-space expression about the water level of the Great Lakes, and then construct the natural state-transition matrix and the decision state-transition matrix which change with time by means of stepwise linear regression. Then, the loss function is constructed by the decision variable and the difference between the state water level and the optimal water level. After simplifying the loss function into a quadratic form, the decision variable that minimizes the loss function is obtained by quadratic programming, which constitutes the MPC dynamic programming process. We use this model to regulate the water level of the Great Lakes in 2017, and obtain the function curve of MSE over time, as shown in the figure (10).

In the sensitivity analysis, we optimized the MPC optimal water level control model, added the abnormal weather water level variation Δ into the discrete state-space expression, and built a predictor-Corrector MPC system, so that the model can be adjusted according to abnormal weather conditions. In addition, considering the surge of precipitation in May, June and July in Lake Ontario in 2017, the water level was re-regulated by actual precipitation data and the optimized model, and the function curve of MSE over time was obtained (figure (13)). We use different controllable flow rates v to regulate the water level of the Great Lakes in 2017, and obtain the function curve of MSE with time (figure (12)). Through these function curves, the effect of MPC optimal water level control algorithm on water level regulation can be clearly observed, and the effectiveness and stability of the model are verified.

Finally, we discussed the advantages and disadvantages of the model, as well as how to optimize and promote it, and provided a one-page memo to the IJC leadership describing the data, features, usage methods, and optimization methods of the model.

Keywords: PCS weights, State-transition matrix, MPC , Dynamic Programming, Predictor-Corrector MPC system, Sensitivity analysis

Contents

1 Introduction

1.1 Background

- **Water problems in the Great Lakes**

  The Great Lakes of the United States and Canada are the largest group of freshwater lakes in the world, including Lake Superior, Lake Michigan, Lake Huron, Lake Erie and Lake Ontario. These five lakes and connected waterways make up a huge drainage basin that contains many of the major cities of both countries. The climate and local weather conditions of these lakes vary, forming a rich and diverse ecological environment.

  Water from the Great Lakes has many uses, and a wide variety of stakeholders have an interest in the water management of the lakes. The water level of each lake is determined by the amount of water coming in and out of the lake and is the result of many complex interactions. There are two main control mechanisms in the flow of water in the Great Lakes system: the Sioux River Lock Compensation Project, the Mary (three hydroelectric plants, five sailing locks, and a gate dam at the top of the rapids), and the Moses-Sanders Dam in Cornwall.

  While the two control DAMS, many channels and canals, and watershed reservoirs can be controlled by humans, the rate of rainfall, evaporation, erosion, ice jams, and other water flow phenomena cannot be controlled by humans. Local government policies may have different impacts than expected, as may seasonal and environmental changes in watersheds. These changes, in turn, affect the ecosystem of the region, which in turn affects the health of plants and animals inside and outside the lake, as well as the inhabitants living in the water basin.

- **Great Lakes 2014 Plan**

  The Great Lakes' 2014 Plan is the Great Lakes Recovery Action Plan published by the U.S. Environmental Protection Agency (EPA) in 2014. The plan is designed to accelerate the conservation and restoration of the world's largest surface freshwater system, achieve long-term sustainability of the Great Lakes ecosystem, and provide additional resources to achieve the most critical long-term goals for critical ecosystems.

  The 2014 plan's algorithms are based on trigger points and thresholds for Lake Ontario levels, Ottawa River flows, and Port of Montreal levels, but may not adequately account for other factors. The plan was controversial when record lake levels were recorded in 2017 and 2019, with some stakeholders expressing skepticism about the plan.

- **MPC control algorithm[1]**

  Model Predictive Control (MPC) is an advanced process control method, which has been used in process industries such as chemical industry and oil refining since the 1980s,[2] and has been applied in the economic field. MPC is a multivariable control strategy, which involves the dynamic model of the loop in the process, the historical value of the control quantity, and an optimal value equation in the prediction interval.

  The advantage of MPC is its ability to handle multiple constraints, such as the output range of the controller, the state limits of the system, and so on. It consists of three steps: predicting the future dynamics of the system, solving the loop optimization problem, applying the first element of the optimization solution to the system, which is a feedback control strategy.

1.2 Restatement of the Problem

With this background, our team will build a network model for the Great Lakes and address the following questions:

1. Consider the wishes of stakeholders to construct an optimal water level for the Great Lakes and bring the optimal water level close to the long-term average water level.

2.According to the optimal water level, a control algorithm is designed to make the lake water level as close as possible to the optimal water level.

3.The sensitivity of the model control algorithm to the outflow of the two dams was analyzed. And consider whether the new control method will make the actual recorded water level in 2017 more satisfying to multiple interests.

4.Analyze the sensitivity of the model control algorithm to changes in environmental conditions (e.g., precipitation, winter snow cover, ice jams).

5.The extensively analyzed model algorithm focuses on the factors affecting Lake Ontario and verifies its effect.

6.Provide IJC leadership with a one-page memo explaining the key features of the model to convince them to choose your model. This should include the use of historical data to explain the model and its parameters, and to compare the new control strategy with the previous model.

1.3 Our work

Our work includes model assumptions, data processing, model building, model testing, sensitivity analysis, model optimization and generalization, and writing memos to IJC leadership, Refer to the work flow chart (1).

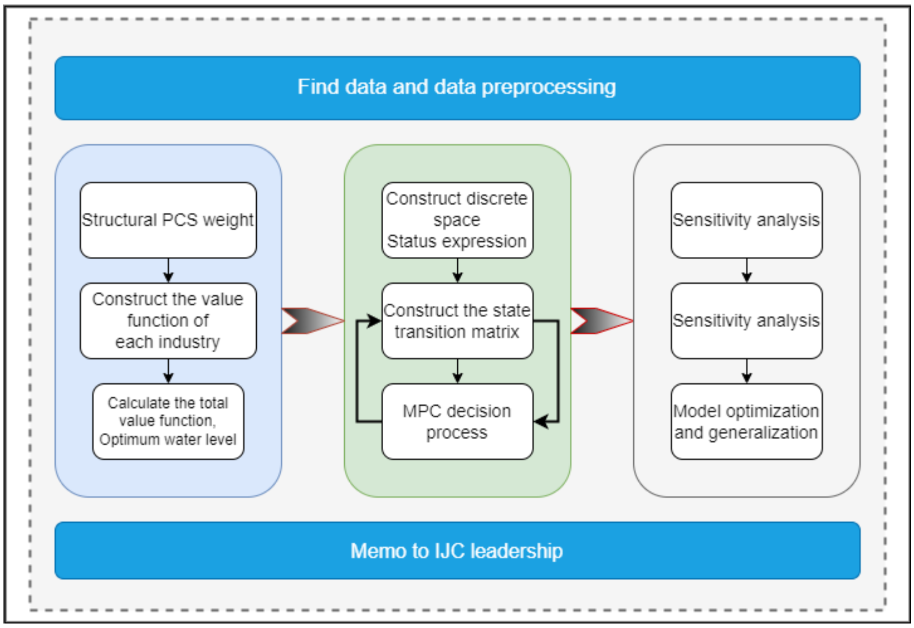

Figure 1: Work flow chart

2 Assumptions and Justification

To simplify the problem and make it convenient for us to simulate real-life conditions, we make the following basic assumptions, each of which is properly justified.

- **It is believed that the data obtained and used by us are true and effective, and the multi-party interests taken into account are comprehensively representative**. The data used include water level data, precipitation, evaporation, snow and ice thickness, river flow, dam controllable flow, etc. Considering the interests of all parties, including shipping, ports, environmental protection, lakeside owners and fishermen fishing vessels, hydropower, irrigation, etc., their interests can reflect the water level demand.

- **Assuming that there will be no extreme weather conditions in recent years and no significant changes in water level, the optimal water level should be in line with the average value in recent years and in a reasonable small range**. In order to reduce the difficulty of regulation and control, assuming that there will be no extreme weather, in order to reduce the cost of regulation and control and to maximize the interests of all parties and environmental protection, the optimal water level should be close to the long-term average water level and have seasonal changes.

- **The water level of the dam can be controlled only by the adjacent upstream lake and all the downstream lakes, and the upstream lakes can not be directly controlled by the dam except for the lakes directly connected to the dam**. In order to reduce the complexity of the model, this assumption is made to obtain the state transition equation with clear relationship.

- **The available data can be approximated as the natural variation of lake level, which is rarely affected by the current dam control strategy**. When building the prediction model, we need to build a model of the natural change of lake water level, which needs to ignore the influence of existing dam control strategies on water level as much as possible.

- **Does not consider the impact of other tributaries on the water level of the Great Lakes**. There are many tributaries in the Great Lakes region, which will also affect the water level change of the Great Lakes, and it is difficult to calculate the flow rate of tributaries in the Great Lakes region, so we make this assumption and only consider the influence factors of major rivers on the Great Lakes to build the model.

3 Model preparation

3.1 Symbol specification

Symbols	Description	Unit
$J_i(x)$	The value function of the i industry	-
w_i	Indicates the PCS weight of the i industry	-
$x_i(t)$	t time Great Lakes water level	m
$x_i^*(t)$	Optimal water level of the Great Lakes	m
$H_i(t)$	Lake i level at t time	m
$u_i(t)$	Decision open time of the i dam at time t	h
A	Natural state-transition matrix	-
B	Decision state-transition matrix	-
P_i	The rate of change in the water level of i lake that can be controlled by dams	m/s
$L(x)$	Loss function in MPC procedure	-
$e_i(t)$	The difference between the i lake level and the optimal water level at t time	m
Δ	The amount of water level change beyond the predicted value	m

Where we define the main parameters while specific value of those parameters will be given later.

3.2 The establishment of Great Lakes network model

By intercepting the major lakes and major rivers in the Great Lakes region, as well as the locations of key nodes and dams, we constructed a network model for the Great Lakes region, as shown in figure(2).

In order to better describe the dynamic change of water quantity and water level in the future, we conducted differential equation modeling and constructed a dynamic flow model of differential equation of lake water storage capacity, which simplified the freezing caused by seasonal cold waves. The change of the water volume (V_i) due to inflow (I_i), outflow (O_i), precipitation (R_i), evaporation (E_i) is a combination of four factors, such as formula(1) is constructed dynamic model of differential equations.

$$\frac{dV_i}{dt} = I_i(t) + R_i(t) - E_i(t) - O_i(t) \tag{1}$$

4 Model I:PCS optimal water level model

4.1 Data preprocessing

The optimal water level should be close to the mean value of the long-term water level. We cleaned the data by drawing a box plot (see the figure (3)), eliminated the abnormal water level data, and filled the missing value by means of mean value filling. The data processing table is shown in the table (1).

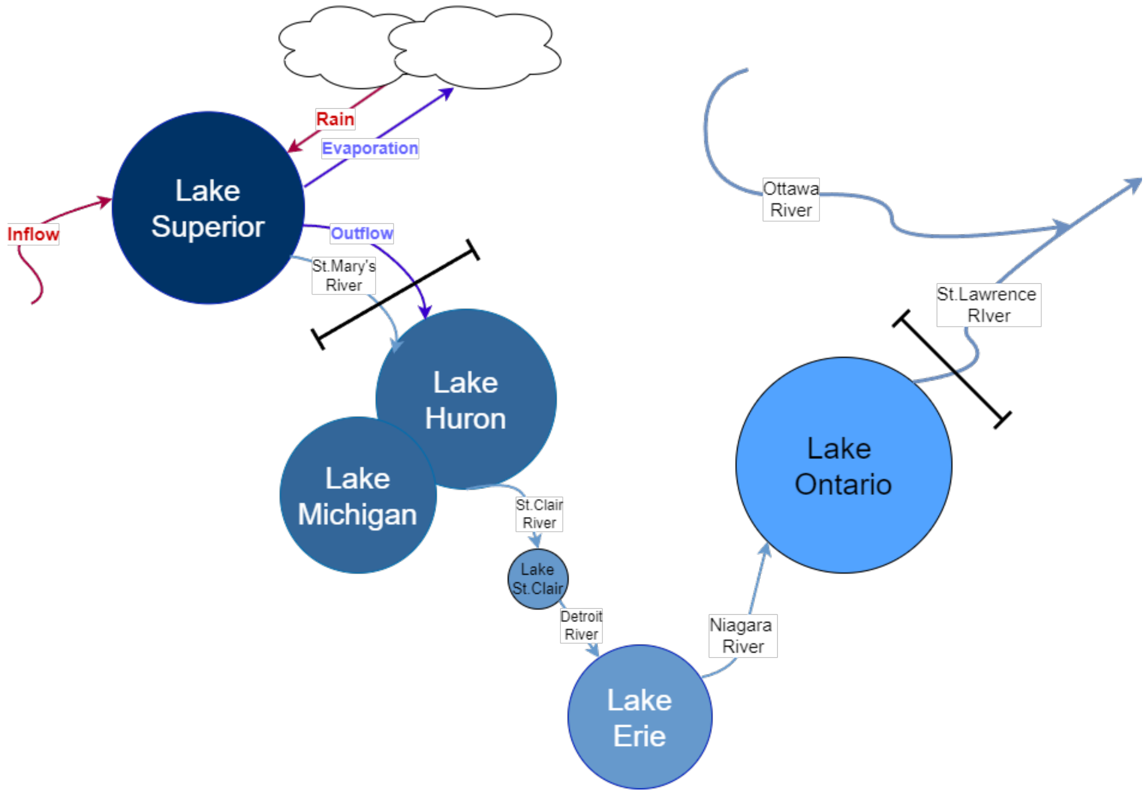

Figure 2: Great Lakes network model

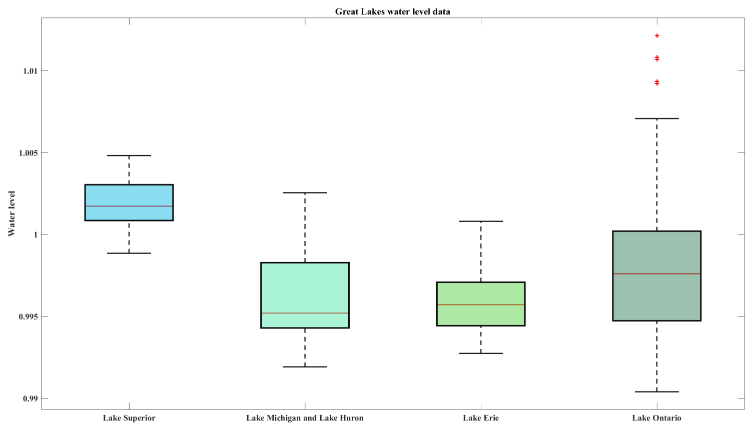

Figure 3: Box plot

4.2 Value function and weight construction

Because of the geographical differences in the five Lakes, the industrial weights and water levels are also different, so we conducted an isolated analysis on the setting of the best water level of the five Lakes, mainly based on the interests of several stakeholders or activities such as shipping, ports, environmental protection (i.e. habitat for animals and plants), lakeside owners and fishermen and fishing boats, hydropower generation, and irrigation, and determined the best water level of the five Lakes respectively.

We construct the total value function by constructing a single industry value function and

Table 1: Data cleaning and filling

Abnormal data	Lake	Time	Filled value
75.8	Ontario	2017/May	74.98
75.81	Ontario	2017/Jun	74.98
75.69	Ontario	2017/Jul	74.98
75.7	Ontario	2019/May	75.09
75.91	Ontario	2019/Jun	75.09
75.8	Ontario	2019/Jul	75.09

weighting linear combination, and maximize the total value function to determine the optimal water level of the lake. First, the requirements of each industry on water level are analyzed, a single value function is assigned, and the personal will value function of a single industry is obtained. There are two methods for constructing the personal will value function:

1. **Growth type**

   For example, for industries such as hydropower, which require a high water level and cannot generate electricity when the water level is below a certain value, the following value function is constructed for such industries:

$$J(x) = \begin{cases} 0, & x \leqslant h \\ \frac{1}{1+\alpha(x-I)^{-2}}, & x > h \end{cases} \tag{2}$$

   Where, $J(x)$ is the value function, the range is between 0 and 1, x is the water level of the lake, h refers to the height of the lowest generating water level, that is, the standing water level, α is the function hyperparameter, the function image is shown in the figure (4).

2. **Intermediate type**

   For example, fishing is an industry that neither requires a high water level nor too low water level, but maintains a specific water level range to obtain the maximum benefits. In order to get a continuous value function, we deformed the Gaussian function and constructed the following value function for such industries:

$$J(x) = \begin{cases} ae^{-\frac{(x-h_1)^2}{c}}, & x < h_1 \\ a, & h_1 \leq x \leq h_2 \\ ae^{-\frac{(x-h_2)^2}{c}}, & h_2 < x \end{cases} \tag{3}$$

   In the above formula, a is the highest value, which is consistent with the growth type, $a = 1$, so that $J(x)$ ranges from 0 to 1, x is the water level of the lake, h_1 is the lowest water level to obtain the maximum value, h_2 is the highest water level to obtain the maximum value, and $c$is the hyperparameter, which depends on a. The determination of h_1 and h_2 and the distance between them is determined by the adaptability and sensitivity of the specific industry to the water level. The function image is shown in the figure (4).

In this way, we have obtained the corresponding value function of each industry, as shown in table (2).

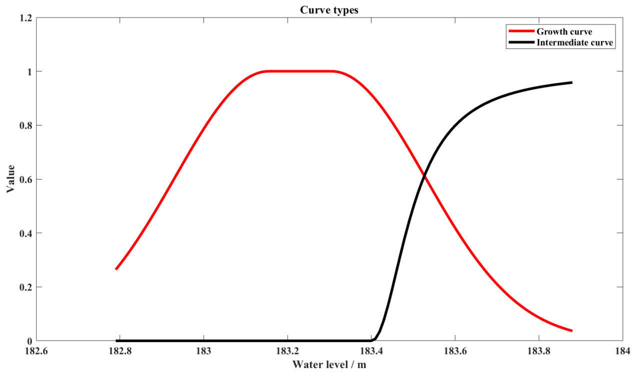

Figure 4: Function image

Then we constructed the value weights of each industry, and constructed the total value function in a linear combination. In order to construct a more objective weight, we investigated the industry-related data of the Great Lakes, including the contribution proportion of local GDP[3] (P) and water consumption[4] (C), and determined the PCS weights according to their industrial status (S).[5] To make the industry obtain the highest possible profit while using the least amount of resources as a weight evaluation criterion, the expression is shown as eq.(4).

$$w_i = \frac{P_i}{log_{10}C_i} \times S_i \tag{4}$$

For industry i, P_i is the contribution proportion of local GDP,[6] and C_i is the water consumption. Due to the huge difference in water consumption, the log function with base 10 is used to scale, and a more reasonable weight value S_i is the industry status, and w_i is the P-C-S weight of the industry. P_i and C_i can be obtained through survey data. S_i is the weight assigned according to the status of the industry. We divide the industry into three levels according to the degree of importance it attaches to the people, and assign weights of $1, 2, 3$ respectively. Taking Lake Superior as an example, we have collected and averaged the profit and water consumption data of various industries in the past ten years. The data for Tab(2) is obtained.

Table 2: Benefit data

Profession	$C_i(MGD)$	$P_i(\%)$	S_i	Value function	Range (ft)
Tourism and Recreation	57	17 .0	2	Intermediate type	0.5
Transportation,Warehousing	-	11.0	2	Growth type	-
Agriculture,Fishing,Food	29	10.0	3	Intermediate type	0.5
Science and Engineering	210	2.0	2	Intermediate type	0.6
hydraulic electrogenerating	30238	59.0	1	Growth type	-
eco-environment protection	5	1.0	2	Intermediate type	0.3

Through these data, the PCS weights of each industry can be obtained. After normalization, weighted linear combination operations are performed on the value functions of each industry through weights, i.e. eq.(5).

$$J(x) = \sum_i w_i J_i(x) \tag{5}$$

$J(x)$is the value function of a lake, which is the same latitude mapping between water level and value, and its global maximum position can be obtained directly through the image.

4.3 Solving the optimal water level

As described in the material, the water level change of 2-3 feet will have a great impact on the interests of the industry, so we carry out the $[\mu - 0.5, \mu + 0.5]$ (feet) limit on the optional water level interval, respectively, to obtain the four value function curves of the Great Lakes (two of which have roughly the same water level). Then, an appropriate range is selected to reduce the cost of seasonal regulation and meet the demand for environmental protection, and the water level within the range that makes its value function highest is their best water level. This model determines the best water level of the Great Lakes, as shown in the figure (5). Finally,

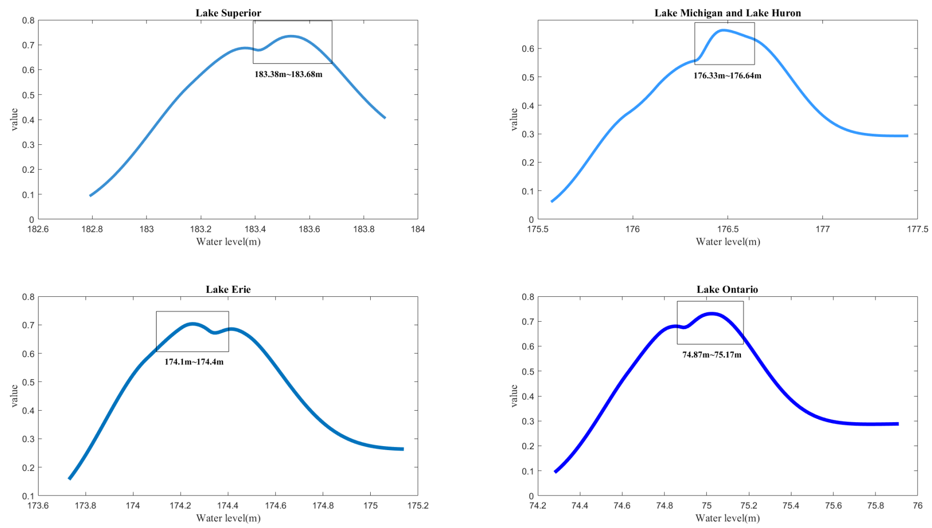

Figure 5: Value function

the optimal water level interval of LAKE SUPERIOR is $[183.38, 183.68](m)$, and the optimal water level interval of LAKE MICHIGAN and LAKE HURON is $[176.33, 176.64](m)$. The best water level range in LAKE ERIE is $[174.1, 174.4](m)$, and the best water level range in LAKE Ontario is $[74.87, 75.17](m)$.

5 Model II:MPC optimal water level control model

5.1 Establishment of MPC linear state transition equation

We need to help manage and control the two dams, which are the most important human controllable factors in controlling the water level of the Great Lakes, so that the water level of the lake is as close as possible to the optimal level that we set and meet the wishes of a wide range of stakeholders. However, by observing the water level of the Great Lakes, there are some obvious regular changes. It can be seen that the natural shaping effect is the most important factor in the change of the water level of the lake, so it is an essential step to predict the dynamic change relationship of the lake water through natural conditions.

MPC, full name Model Predictive Control[7],[8] is a control optimization algorithm for multiple objectives, the value function of this algorithm is generally a 2-norm of error and cost, and the predictive description of this algorithm is linear. We apply it to water level control eq(6). That is, discrete state-space expression.

$$x(t+1) = Ax(t) + Bu(t) \tag{6}$$

Where $x(t)$ is the water level of the Great Lakes at time t, and $x(t+1)$ is the water level of the Great Lakes at time t in the next month. The meaning of the A matrix is how the water level at time t is transferred to the water level at time t (t+1), which is named as the natural state-transition matrix, and $u(t)$ is the decision variable. The B matrix describes how the decision variables at the time of t affect the water level at the time of $t+1$, named as the decision state-transition matrix. The specific mathematical form of the water level and decision variables is as follows:

$$x(t) = \begin{bmatrix} H_S(t) \\ H_m(t) \\ H_E(t) \\ H_0(t) \end{bmatrix} \quad u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \tag{7}$$

In the above formula, $H_S(t), H_m(t), H_E(t), andH_O(t)$ are the water levels of the five Great Lakes respectively (among them, two lakes have the same water levels, so they are synthesized into one $H_m(t)$ variable). $u_1(t)$ and $u_2(t)$ are the opening times of the first upstream and downstream dams in t months, respectively.

In this way, as long as the natural state-transition matrix and decision state-transition matrix are obtained, the MPC method can be used to construct the objective value function and enter the dynamic multi-objective decision iteration process.

5.2 Construction of natural state-transition matrix

The natural state-transition matrix describes the change of the water level of the Great Lakes without human intervention (decision variables). Therefore, when constructing the natural state-transition matrix, we ignore the decision state-transition matrix, that is, the second term of the linear combination in eq(6). At this time, the water level predictive control model is eq(8).

$$x(t+1) = Ax(t) \tag{8}$$

To expand the equation, i.e

$$\begin{aligned} x(t+1) &= \begin{bmatrix} H_S(t+1) \\ H_M(t+1) \\ H_E(t+1) \\ H_O(t+1) \end{bmatrix} = \begin{bmatrix} \overrightarrow{a_1} \\ \overrightarrow{a_2} \\ \overrightarrow{a_3} \\ \overrightarrow{a_4} \end{bmatrix} \begin{bmatrix} H_S(t) \\ H_M(t) \\ H_E(t) \\ H_O(t) \end{bmatrix} = \begin{bmatrix} \overrightarrow{a_1} \cdot \vec{x}(t) \\ \overrightarrow{a_2} \cdot \vec{x}(t) \\ \overrightarrow{a_3} \cdot \vec{x}(t) \\ \overrightarrow{a_4} \cdot \vec{x}(t) \end{bmatrix} \\ &= \begin{bmatrix} a_{11}H_S(t) + a_{12}H_M(t) + a_{13}H_E(t) + a_{14}H_O(t) \\ a_{24}H_S(t) + a_{22}H_M(t) + a_{23}H_E(t) + a_{24}H_O(t) \\ a_{31}H_S(t) + a_{32}H_M(t) + a_{33}H_E(t) + a_{34}H_O(t) \\ a_{41}H_S(t) + a_{42}H_M(t) + a_{43}H_E(t) + a_{44}H_O(t) \end{bmatrix} \end{aligned} \tag{9}$$

Analyze only H_S, get

$$H_S(t+1) = a_{11}H_S(t) + a_{12}H_M(t) + a_{13}H_E(t) + a_{14}H_O(t) \tag{10}$$

With the water level of each great lake at the time of t as the independent variable and $H_S(t+1)$ as the dependent variable, the coefficient vector in eq(10) can be obtained by means of multi-variable linear regression, that is, the values of $a_{11}, a_{12}, a_{13}, a_{14}$ can be obtained. Other elements of the natural state-transition matrix can also be obtained by using this method, that is, the natural state-transition matrix can be obtained by linear regression of multiple variables.

The variation of the water level of the Great Lakes is full of random factors, and it is inevitable that the linear regression method cannot fit the change of water level in a long period. We select different time periods for linear regression and calculate their R^2 values to obtain the functional relationship between the time period length (month) and R^2, as shown in the figure (6).

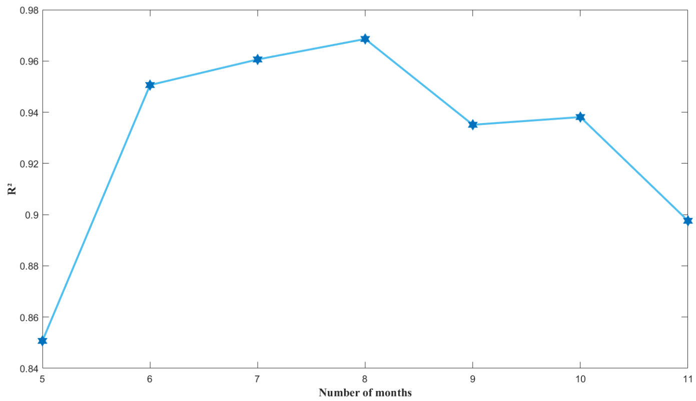

Figure 6: Multivariable linear regression

When the water level of the first eight months of the current month is selected as the basis of linear regression, R^2 reaches the highest value, about 0.97. We believe that the linear regression at this time can correctly represent the water level change in a period of time, and the water level value of the last eight months is uniformly used for linear regression when constructing the natural state transition equation.

5.3 Construction of decision state-transition matrix

Now reconsider the decision factor and transform eq(8) :

$$\begin{aligned} X_{\text{True}}(t+1) &= X_{\text{Natural}}(t+1) + \Delta x(t). \\ &= Ax(t) + Bu(t) \\ \Rightarrow X_{\text{True}}(t+1) - X_{\text{Natural}}(t+1) &= \Delta x(t) = Bu(t) \end{aligned} \tag{11}$$

i.e:

$$\Delta x(t) = Bu(t) \tag{12}$$

Where, $X_{\text{Natural}}(t+1)$ refers to the water level of the lake at the time of $t+1$ when only the natural state-transition matrix is considered, and $X_{\text{True}}(t+1)$ refers to the condition when both the natural state-transition matrix and the decision state-transition matrix are considered. The water level of the lake at $t+1$, where $\Delta x(t)$ is the difference between the two.

First, we set up a hypothesis to simplify the solution of the decision state-transition matrix. This hypothesis is that the water level of the upstream lake directly connected to the dam can be affected only by the opening and closing of the dam, and the water level of the upstream lake is not affected by the downstream dam. For example,SOO LOCKS can affect the water level of all lakes. The MOSES-SAUNDERS POWER DAM only affects the water level in LAKE Ontario and downstream water levels.

With the above assumptions, we can simplify the decision state-transition matrix, i.e:

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & 0 \\ b_{31} & 0 \\ b_{41} & b_{42} \end{bmatrix} \tag{13}$$

Then use the simplified B matrix to expand eq(12), i.e:

$$\Delta x(t) = \begin{bmatrix} \Delta H_S(t) \\ \Delta H_M(t) \\ \Delta H_E(t) \\ \Delta H_O(t) \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & 0 \\ b_{31} & 0 \\ b_{41} & b_{42} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} b_{11}u_1 \\ b_{21}u_1 \\ b_{31}u_1 \\ b_{41}u_1 + b_{42}u_2 \end{bmatrix} \tag{14}$$

The linear equations satisfied by the matrix elements can be obtained:

$$\begin{cases} \Delta H_S(t) = b_{11}u_1 \\ \Delta H_M(t) = b_{21}u_1 \\ \Delta H_E(t) = b_{31}u_1 \\ \Delta H_O(t) = b_{41}u_1 + b_{42}u_2 \end{cases} \tag{15}$$

Now consider the equation $\Delta H_S(t) = b_{11}u_1$ of the first line of linear equations, we assume that the flow rate of the dam is constant, and the controlled flow rate of the two dams is v_1, v_2, then there is

$$\begin{aligned} b_{11} &= Cv_1 \\ v_1 u_1 &= \Delta V_s \end{aligned} \tag{16}$$

i.e:

$$\Delta H_s(t) = Cv_1u_1(t) = C\Delta V_s \tag{17}$$

In the above formula, C is a constant and ΔV_s is the amount of water in LAKE SUPERIOR that changes due to the decision variable, that is, the volume of water. Considering that the water level change of 2-3 feet has a great impact, and the whole water body has a depth of $406m$, the change of water level is very small for the water depth, we can equivalent the change volume of the lake water volume to an irregular column, as shown in the figure (7). S is the surface area of

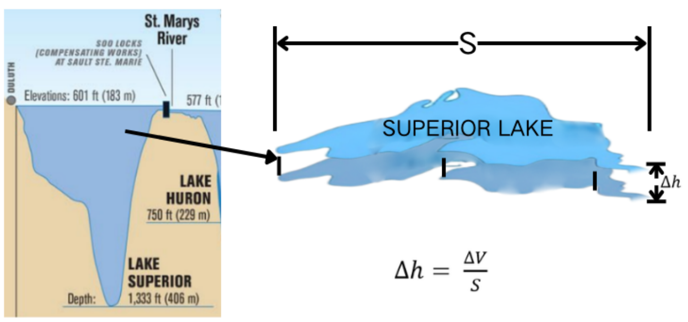

Figure 7: Column cross-section

the lake, ΔV_s is the volume change of the lake, and Δh is the change of the water level, through which we can further deduce the above formula, as follows:

$$\Delta H_S(t) = C_0 \Delta V_S = -\frac{\Delta V_S}{S} = -\frac{v_1}{S} u_1 = -P_S u_1 \tag{18}$$

In eq18, because $\frac{v_1}{S}$ is the dam flow velocity and the ratio of the surface of the lake, both as a constant, so we make $\frac{v_1}{S} = P_S$, so you can get $b_{11} = -P_S$.

Next, consider the equations(15) of the second and third lines of the linear equations eq.(15), which have the same form. Taking the second equation as an example, if it is difficult to judge the increase of the downstream water level directly from the dam flow, it may be better to settle for the second place and find the relationship between different water level changes. We can already represent the change of LAKE SUPERIOR water level through u_1. If there is an implicit linear correlation between the change of water level between LAKE SUPERIOR and other lakes, we can determine its linear correlation coefficient by means of unary linear regression, and then deduce it as follows:

$$\Delta H_M(t) = k_M \Delta H_S(t) = -k_M P_S u_1(t) = b_{21} u_1(t) \tag{19}$$

Obtain $b_{21} = -k_M P_S$, similarly, $b_{31} = -k_E P_S$, where k_M and k_E are obtained by unary linear regression. By choosing different time lengths (months) for linear regression, we can obtain the functional relationship between time length and R_2. As shown in the figure (8). When the

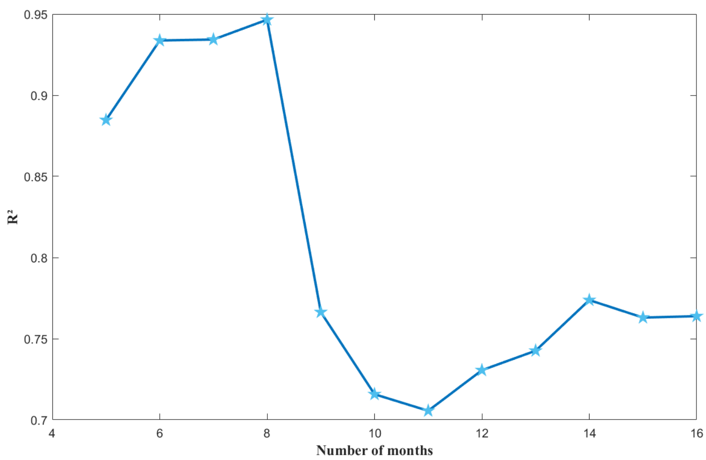

Figure 8: Unary linear regression

selection time length is 8 months, the maximum value of R_2 is obtained, which is about 0.95. We also use the values of the last 8 months for linear regression, and obtain k_M and k_E.

Considering the last line of the linear equation equation eq.(15), the water level regulation of H_O is jointly controlled by two dams. Since the model we built belongs to the linear model, it must satisfy the superposition principle. At this time, the equation is consistent with the equations in the second and third rows of the linear equations eq.(15), and $b_{41} = -k_O P_S$can be obtained in the same way. K_O is also obtained using unvariable linear regression, and then it is assumed that only the downstream dam is used for regulation. This is the same as the equation in the first row of the linear equation system eq.(15). By constructing a similar bar graph,$b_{42} = -P_O$ can be obtained, where $P_O = \frac{v_2}{S}$ and S are the surface area of Lake Ontario.

Through the above process, we solve the linear equations eq.(15), whose solution is the decision state-transition matrix B, i.e:

$$B = \begin{bmatrix} b_{11} & 0 \\ b_{21} & 0 \\ b_{31} & 0 \\ b_{41} & b_{42} \end{bmatrix} = \begin{bmatrix} -P_S & 0 \\ -k_M P_S & 0 \\ -k_E P_S & 0 \\ -k_O P_S & -P_O \end{bmatrix} \tag{20}$$

5.4 MPC optimization control

Through the above series of operations, we have completed all the preparatory operations of MPC, according to the MPC principle,[9] we construct its loss function, which is the target value function $L(x)$.

$$L(x) = \sum_{i=0}^{i+1} \left(e(t+i \mid t)^\top Q e(t+i \mid t) + u(t+i \mid t)^\top R u(t+i \mid t) + e(t+N)^\top F e(t+N) \right) \tag{21}$$

In the above formula, Q, R, F is the custom weight, and $e(t)$ is the difference between the Great Lakes water level $x(t)$ and the best water level $x^*(t)$ at time t, that is, $e(t) = x(t) - x^*(t)$. $e(t+i \mid t)^\top Q e(t+i \mid t)$ is the 2-norm of the error matrix at each time within the selected step, $u(t+i \mid t)^\top R u(t+i \mid t)$ is the 2-norm of the decision control time, $e(t+N)^\top F e(t+N)$ is the 2-norm of the final error of the prediction space, the purpose of which is to determine the decision $x(t)$ so that the value of the loss function $L(x)$ is minimized, i.e

$$\text{Object: MIN}\, L(x) \tag{22}$$

The meaning of this goal is to minimize the error between the current water level and the optimal water level by minimizing the number of controls (the shortest control time). Through some mathematical derivation and simplification of the formula,[10] the simplified form of the loss function can be obtained:

$$L(x) = e(t)^\top G e(t) + 2e(t)^\top E U(t) + U(t)^\top H U(t) \tag{23}$$

Where, $U(t), G, E, H$ satisfies the following form:

$$U(t) = \begin{bmatrix} u(t \mid t) \\ u(t+1|t) \\ \vdots \\ u(t+N-2 \mid t) \\ u(t+N-1 \mid t) \end{bmatrix} \begin{cases} G = M^\top \bar{Q} M \\ E = C^\top \bar{Q} M \\ H = C^\top \bar{Q} C + \bar{R} \end{cases} \tag{24}$$

There are also some variable constructs, namely variables $M, C, \bar{Q}, \bar{R}$ as follows:

$$\bar{Q} = \begin{bmatrix} Q & & & \\ & \ddots & & \\ & & \ddots & \\ & & & Q \\ & & & & F \end{bmatrix} \quad \bar{R} = \begin{bmatrix} R & & & \\ & \ddots & & \\ & & \ddots & \\ & & & R \end{bmatrix}$$

$$M = \begin{bmatrix} I \\ A \\ A^2 \\ \vdots \\ A^N \end{bmatrix} \quad C = \begin{bmatrix} 0 & \cdots & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & & & & \vdots \\ B & 0 & & & \vdots \\ AB & B & \ddots & \ddots & \\ \vdots & & & \ddots & 0 \\ A^{N-1}B & A^{N-2}B & & \cdots & B \end{bmatrix}$$

Through such simplification, all variables can be obtained through the state-transition matrix and the custom weight matrix, the loss function can be simplified, and the quadratic programming[11] can be used to optimize the solution.

The hyperparameter N is the prediction interval, that is, the length of the future decision time period. It is an important hyperparameter, and we will adjust the parameter in the simulation.

In the above process, we construct the state-transition matrix and define the MPC process of gradual update. Our MPC process is based on the latest data, automatically constructs the state-transition matrix and obtains the new state quantity, which is an algorithm of real-time update data planning. The pseudo-code of the algorithm is shown in Algorithm.(1), and the flow is the NO path in the figure (9).

Algorithm 1: Model Predictive Control Simulation

Data: State matrix A, Input matrix B, Weight matrices Q, R, F, Prediction horizon N, Demand samples
Result: Optimal control input U_K

Define state-space matrices A, B, Q, R, F;
Define multivariate linear fitting data;
Initialize state variables x_t and control inputs U_K;
Compute matrices E, H, G using function MPC_Matrices;
for $t = 1$ *to* t_{steps} **do**
    Update A, B, Q, R, F if $t > 1$;
    Compute optimal control input $U_t(:, t)$ using quadratic programming;
    Compute next state $x_t(:, t+1)$;
end
Plot state variables and control inputs;

5.5 The MPC model was used to optimize the control of water level data in 2017

In order to intuitively represent the control effect of the MPC optimal water level control model on the water level of the Great Lakes, MSE is used to establish the water level control index. The MSE water level control index is as follows:

$$MSE = \frac{1}{n} \sum_i \left(x_i(t) - x_i^*(t)\right)^2 \tag{25}$$

In the above equation, n is the number of the Great Lakes, $x_i(t)$ is the water level of the Great Lakes at time t, $x_i^*(t)$ is the best water level of the Great Lakes at time t, and MSE depicts the

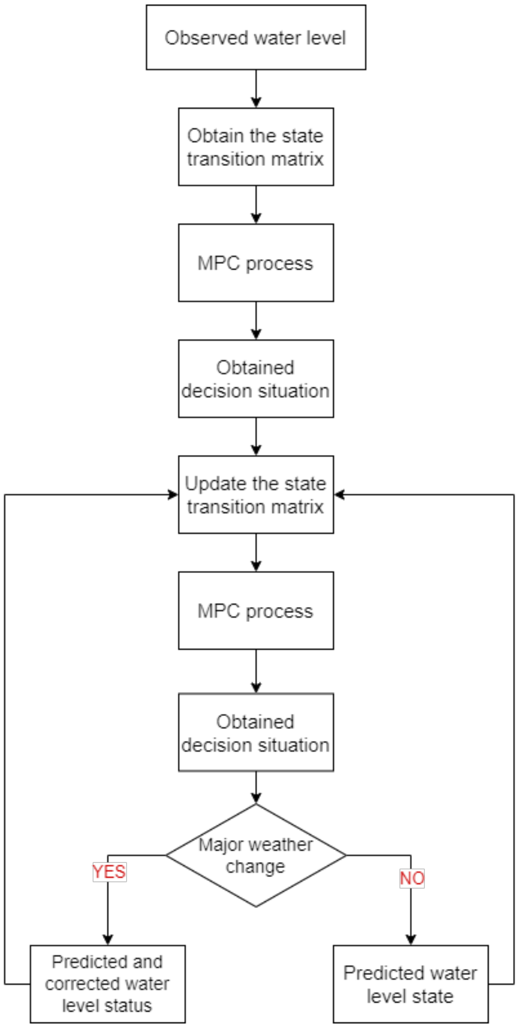

Figure 9: Algorithm flow chart

gap between the water level at time t and the best water level. The smaller MSE is, the closer the water level of the Great Lakes is to their respective best water levels.

We used the MPC optimal water level control model to regulate the water level of the Great Lakes in 2017, and obtained the relationship between MSE of the real regulation data and MSE of the water level controlled by the MPC algorithm over time, as shown in the figure (10). In addition, we also make a profile of the relationship between the absolute value of the water level error of the Great Lakes and the time ($|e(t)| - t$), as shown in the figure (11).

It can be observed that, assuming natural conditions (no major weather fluctuations), the MPC can reduce the MSE indicator to near 0 in just a few months, which is a good way to complete the water level control. However, through our observation and search of data, there was a large amount of precipitation in the Great Lakes region in May, June and July of 2017, which far exceeded the normal precipitation. Considering the abnormal precipitation in 2017, we will consider and simulate it in the sensitivity analysis below.

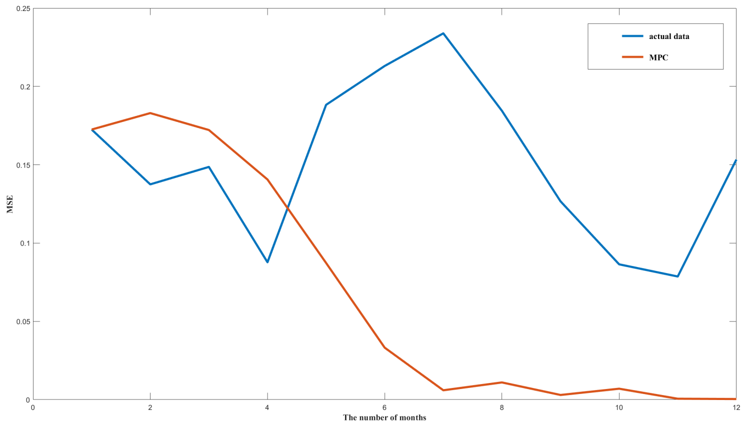

Figure 10: $MSE - t$

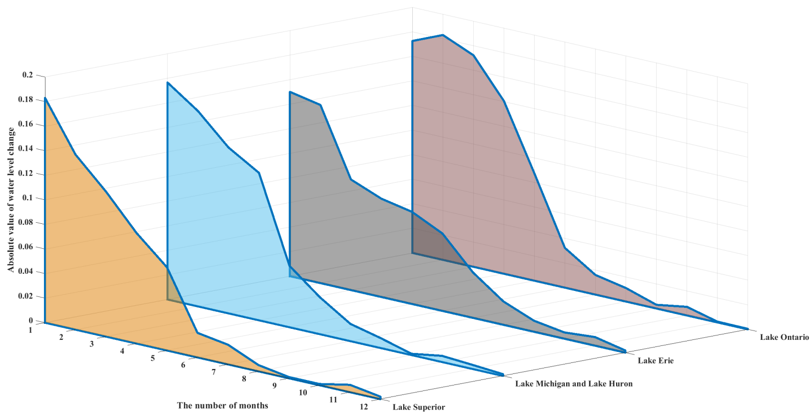

Figure 11: $|e(t)| - t$

6 Sensitivity Analysis

6.1 The sensitivity of the model to dam flow

The dam flow rate used in the above MPC model is a constant, but it is impossible for the dam to operate only with constant displacement during operation. The changes in the drainage velocity of the dam are considered here to analyze the sensitivity of the MPC model to the flow velocity controlled by the dam.

In order to facilitate the comparison, we still choose the data of 2017 for comparison. When the dam control flow rate v_1, v_2 change, the natural state-transition matrix will not change, but the decision state-transition matrix will change. It is recalled that the values of eq.(20), k_M, k_E, k_O are still determined by unvariable linear regression, but the values of P_S, P_O will change due to the change of flow velocity. Therefore, the constants of the decision state matrix need to be reassigned. The flow rate that can be controlled by the dam is appropriately increased and reduced, and the MSE curve under three states of fast, normal and slow flow rate is drawn, as shown in the figure (12).

It can be seen that the MPC optimal water level control model can still make the water level

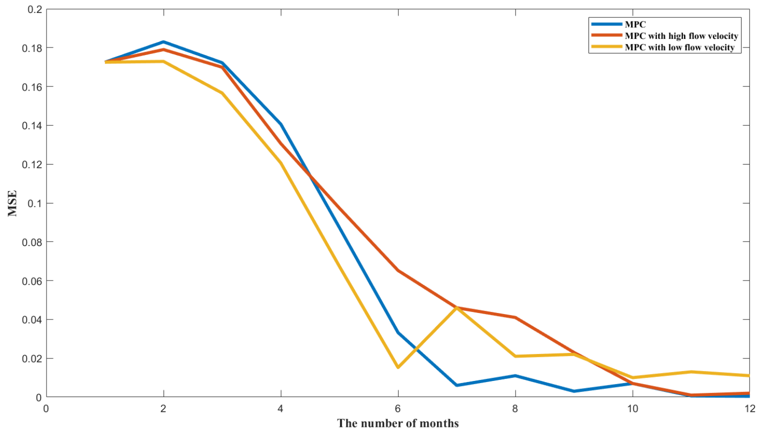

Figure 12: $MSE - t$

of the Great Lakes approach the optimal water level relatively quickly under the condition of dam velocity variation, which shows the stability of the model.

6.2 How sensitive the model is to environmental conditions

When we established the model, we did not consider the major changes in natural factors such as rainfall, winter snow cover and ice jam. However, in actual circumstances, these environmental conditions may fluctuate violently, resulting in major changes in water level. The major fluctuations of these environmental factors will not affect the decision-making state-transition matrix, but will have a greater impact on the natural state-transition matrix.

Considering that the impact of environmental factors is unknowable, it is impossible for us to update the natural state matrix adaptively before significant changes in environmental conditions are observed. Therefore, the natural state-transition matrix used in dynamic optimization is still constructed according to past water levels, and the decision-making method is the same as when there are no major fluctuations in environmental conditions. It is only after major fluctuations in environmental conditions have occurred that additional effects on water level changes can be known. The effect of major fluctuations in natural conditions, such as rainfall, winter snow cover, and ice jams, is manifested in the water level of the Great Lakes as an increase or decrease in water level that is not predicted by the natural state-transition matrix. For the MPC process, the normal prediction is first made, and then the increase or decrease of the water level outside the discrete state-space expression is added after the decision and observation, i.e. eq.(26).

$$x(t+1) = Ax(t) + Bu(t) + \Delta \tag{26}$$

Δ represents the change of water level outside the discrete state-space expression, which takes into account the change of water level under natural circumstances. Therefore, the determination of Δ needs to subtract the abnormal change of water level from the normal change of water level, i.e. eq.(27).

$$\Delta = \Delta_u - \Delta_n \tag{27}$$

In the above formula, Δ_u represents the abnormal change of water level due to weather factors, and Δ_n represents the normal change of water level under natural state that can be described by natural state-transition matrix.

If a huge rainfall occurs in t month of a certain year, resulting in an abnormal surge in lake water level, we can use the original natural state-transition matrix construction method to linearly predict the precipitation of t month Δ_n by using the precipitation of the first 8 months. Then the MPC process is carried out normally to obtain the decision variables and the state of the next moment, and the situation of the next state is obtained by using the discrete space state expression, and then the value of Δ is calculated. The value of Δ_u is determined by the actual precipitation change. By putting eq.(27) into eq.26, the water level of the next state under the influence of major rainfall can be obtained, and the situation after prediction and decision can be updated using this situation (the state can be obtained by the discrete space state expression), and then the algorithm process can be carried out normally. The flow of the algorithm is the YES path in the figure (9).

The optimized algorithm is very similar to the predictor-corrector method in numerical analysis. Firstly, the algorithm with low accuracy but low computation amount is used for one-step prediction, and then the algorithm with high accuracy but large computation amount is used for correction, so that the algorithm has both high accuracy and low computational complexity. Based on this mathematical model, We named the optimized MPC optimal water level control model as the predictor-corrector MPC system.

In the above article, we mentioned that there was abnormal precipitation in the Great Lakes region in May, June and July 2017, and the natural state-transition matrix did not well describe the water level change in the Great Lakes region. Therefore, we used eq(26) method to dynamically optimize the data of 2017 again, obtained the relationship between the change of the new MSE over time under regulation, and added it to the figure (10), as shown in the figure (13). It can be seen that after we recorded the precipitation situation and improved the discrete

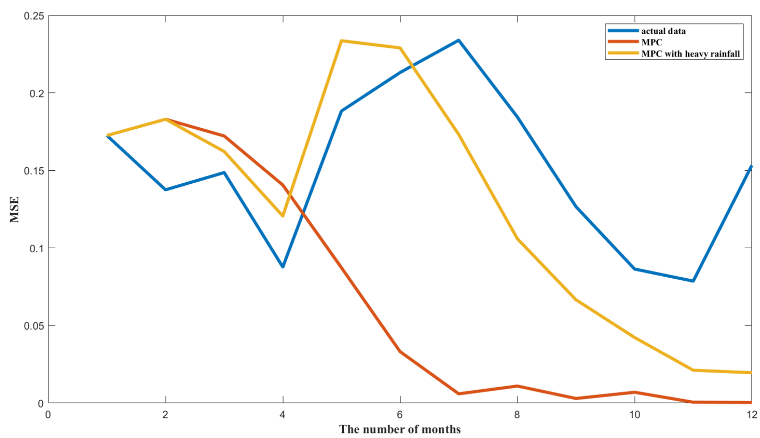

Figure 13: $MSE - t$

state-space expression, the water level curve under the control of MPC increased significantly. However, due to the feedback, MPC made a correct decision and quickly lowered the MSE index to a relatively low value. This shows that our model can better cope with the irregular water level increase or decrease brought by abnormal environmental conditions, and has good stability.

7 Further Discussion

7.1 The model focuses on Lake Ontario

Given the recent increased focus on Lake Ontario's water level management, the idea of applying the model to Lake Ontario has permeated our mathematical modeling process.

Our control algorithm can control the opening time of the two dams and regulate the water level of Lake Ontario. The water level and velocity of the Port of Montreal are jointly affected by the St.Lawrence River and Ottawa River. It is impossible to regulate only the two dams on the five Lakes. In addition, through the cooperation of the dam and hydropower station on the Ottawa River can the water level and velocity of the middle and lower reaches of the St.Lawrence River be regulated. Assuming that they are all in normal operation, our model can meet the water level demand of most stakeholders in Lake Ontario.

In specifying the optimal water level, we focused on the wishes of all stakeholders in Lake Ontario, determined the optimal water level according to their related industries, and constructed the relevant value function through eq.(2) and eq.(3). Based on the industrial data and eq.(4) of Lake Ontario, the PCS weight of the industry is obtained, and then the optimal water level interval of Lake Ontario $[74.87, 75.17](m)$ is determined by solving the extreme value of the value function.

When controlling the water level, we also increase the consideration of Lake Ontario, which is reflected in the setting of the weights of Q, R and F in the MPC control process. In the MPC model of controlling the water level, we strengthen the setting of the weights of Lake Ontario, which can achieve the optimal overall water level under the premise of ensuring that under the influence of various factors, the water level of Lake Ontario can be preferentially reached the optimal water level range.

7.2 Model optimization:Nonlinear and predictor-corrector MPC systems

For a system with strong randomness, our prediction model may not be able to capture the characteristics of the state change of the system. In this case, the constructed state-transition matrix may be quite different from the real state-transition matrix, resulting in the failure to accurately predict the state of the system at the next moment.

To solve this problem, we can optimize the construction mode of the state-transition matrix in the MPC optimal water level control model, and use polynomial regression, cubic spline interpolation, SVR, NNs and other methods to replace the linear regression prediction used in the construction of A, B matrix, so as to better capture the characteristics of system state changes. Then the Jacobian determinant is used to map the nonlinear relation to the linear relation to get the state-transition matrix which can better reflect the characteristics of the system state change, and is applied to the solution of MPC process.

In addition, it is more suitable to adopt the discrete state-space form of eq.(26) in practical application, add Δ to the end of the discrete state-space expression, and correct the prediction of the model by observing the actual data, so as to complete the construction of the predictor-corrector MPC system.

7.3 The model is extended to other lake systems

The MPC optimal water level control model can be applied not only to the Great Lakes region, but also to develop dam control strategies for other lakes, as follows:

1. The weight of PCS is constructed to obtain the optimal water level interval of the target water area.
2. According to the observed data, the natural state-transition matrix and the decision state-transition matrix are constructed by adopting appropriate prediction methods.
3. According to the MPC flow, the dynamic optimization satisfies the figure (14) flow.

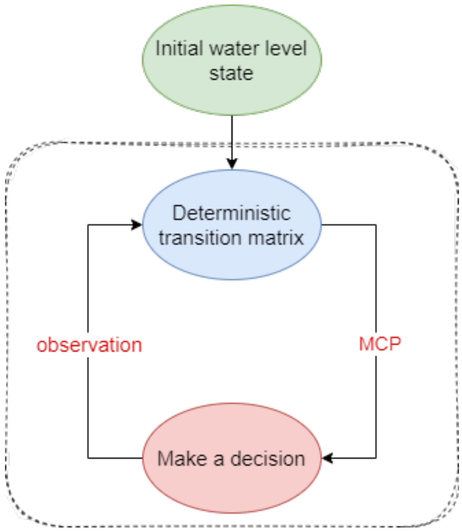

Figure 14: Extended algorithm flow chart

8 Model Evaluation:Strength and Weakness

8.1 Strength

- The effect of controlling water level is remarkable: by constructing the state-transition matrix, the change of water level under natural state can be captured and applied to the MPC solution process to control the change of water level.
- Good robustness and stability: the strategy can be adaptive adjusted according to dam flow, weather data, etc., and can control the change of water level in complex situations.
- High flexibility: By adjusting the weight matrix Q, R and F, the nature of the control strategy can be changed to make it more in line with the actual demand.
- Strong portability: It is easy to transplant the PCS optimal water level model and MPC optimal water level control model to the control of other lake systems.

8.2 Weakness

- Large computational complexity: Every decision has to recalculate all parameters to solve a quadratic programming problem, which can require a lot of computational resources, especially when the system has a high dimension or a large number of prediction steps.
- Prediction accuracy in the face of a system with strong randomness: For a system with large randomness, the construction of the state-transition matrix may need to adopt a more complex nonlinear construction mode to correctly capture the characteristics of the system's state changes.

9 MEMO

To: IJC

From: ICM TEAM

Subject: The establishment of the Great Lakes water level control algorithm

Data: Monday, February 5, 2024

At the invitation of the IJC, we are the team responsible for leading the management plan for developing the model and implementing it. We have established a new algorithm for determining and controlling optimal water levels in the Great Lakes. This memo will introduce you to this algorithm and related instructions for using it.

1. **Algorithm overview**

   The algorithms is MPC optimal water level control model.The MPC control model updates the natural and decision state-transition matrix in real time by establishing discrete state expressions of the Great Lakes water level. By minimizing the loss function through quadratic programming, the decision variable is obtained, and then the decision variable of the next state is obtained iteratively through the state expression, forming a dynamic optimization process. This is how the MPC optimal water level control algorithm works.

2. **Historical data use and parameter construction**

   For the broad applicability of the model, we collected data on monthly precipitation, evaporation, river flows, annual snow and ice thickness, water use, and economic benefits in and around the Great Lakes. The data comes from NOAA's Great Lakes Laboratory, the University of Michigan, and the Great Lakes Water Use Database. Using this data, we build a state-transition matrix through regression analysis and customize the weights for use in dynamic MPC optimization processes.

3. **Model features and comparison**

   Our models accurately capture water level changes, set optimal levels and make decisions to keep the Great Lakes close to optimal levels. We used the model to simulate the data of the 2017 rainfall change, and the results show that the model can regulate water levels quickly and stably, which is better than the dam strategy at that time. In addition, the model has strong robustness, and can cope with emergencies or parameter changes.

4. **Instructions for use**

   Our model makes decisions based on monthly data, but practical applications require daily or even hourly decisions, which requires us to obtain water level data from each observatory in real time and adjust the model parameters in real time. In the real world, weather conditions such as heavy rainfall and ice jams can have an impact, and we need to monitor these data in real time and build a predictive and corrected MPC system to achieve better water level control in the real world.

It is hoped that our MPC control algorithm will provide a more efficient and flexible solution for water level management in the Great Lakes region. If you have any questions or need further information, please feel free to contact me. Thank you!

References

[1] A. R. Neshastehriz, M. Cantoni and I. Shames, ”Water-level reference planning for automated irrigation channels via robust MPC,” 2014 European Control Conference (ECC), Strasbourg, France, 2014, pp. 1331-1336, doi: 10.1109/ECC.2014.6862420.

[2] X. Liu, D. Jiang and K. Y. Lee, ”Quasi-min-max Fuzzy MPC of UTSG Water Level Based on Off-Line Invariant Set,” in IEEE Transactions on Nuclear Science, vol. 62, no. 5, pp. 2266-2272, Oct. 2015, doi: 10.1109/TNS.2015.2466658.

[3] Ke Hu, Jingqi Yuan,Multi-model predictive control method for nuclear steam generator water level,Energy Conversion and Management,Volume 49, Issue 5,2008,Pages 1167-1174， ISSN 0196-8904.

[4] ”Updated Files,” NOAA Great Lakes Environmental Research Laboratory, 2024. [Online]. Available: https://www.glerl.noaa.gov/pubs/tech_reports/glerl-083/UpdatedFiles/.

[5] Marchand, D., Sanderson, M., Howe, D. et al. Climatic change and great lakes levels the impact on shipping. Climatic Change 12, 107–133 (1988).

[6] Jorge A. Garcia-Hernandez , Roy Brouwer (2021) A multiregional input–output optimization model to assess impacts of water supply disruptions under climate change on the Great Lakes economy, Economic Systems Research, 33:4, 509-535, DOI: 10.1080/09535314.2020.1805414.

[7] K. Horváth, M. Petreczky, L. Rajaoarisoa, E. Duviella and K. Chuquet, ”MPC control of water level in a navigation canal - The Cuinchy-Fontinettes case study,”?2014 European Control Conference (ECC), Strasbourg, France, 2014, pp. 1337-1342, doi: 10.1109/ECC.2014.6862483.

[8] Jorge A. Garcia-Hernandez , Roy Brouwer (2021) A multiregional input–output optimization model to assess impacts of water supply disruptions under climate change on the Great Lakes economy, Economic Systems Research, 33:4, 509-535, DOI: 10.1080/09535314.2020.1805414.

[9] ”Great Lakes Jobs Report,” University of Michigan School for Environment and Sustainability, 2024. [Online]. Available: https://seas.umich.edu/sites/default/files/users/user30/MICHU-20-203%20Great%20Lakes%20Jobs%20Report%20(1).pdf.

[10] CHEN Ming-Zhen.Multi-objective multi-objective fuzzy dynamic model for rivers[J].Advances in Science and Technology of Water resources,1993,13(02):69-72.

[11] MATLAB, version 9.10.0 (R2021a). Natick, Massachusetts: The MathWorks Inc., 2021