Problem Chosen	2024 MCM/ICM Summary Sheet	Team Control Number
D		2418844

A network flow control model for the water level of the Great Lakes

Summary

The water resources of the Great Lakes are of great value to the economy, ecology, shipping and other aspects of the surrounding areas. Due to natural factors and human intervention, the water level of the Great Lakes will change, which will change the distribution of water resources. Different interest groups have different purposes and ways to benefit, and they have different dynamic demands for the water level and flow of the Great Lakes, and even contradict each other. Aiming at this problem, this paper has developed a control mechanism that directly affects the water level of the Great Lakes flow network.

Firstly, we set up the network model of water flow in the Great Lakes. Based on the conservation idea, we construct the equation of water resource quantity change in each lake. There is no dam between some subsystems that can regulate the flow. Through the study of the monthly flow of rivers and the monthly water level of lakes, we find that the river flow is generally positively correlated with the water level of upstream lakes. In order to make the system solvable, we explore the characteristics of water flow self-distribution of these subsystems based on linear fitting. The goodness of fit between the discharge of Detroit River and the water level of Lake Michigan and Lake Huron was **0.8353428**. The goodness of fit ratio between the flow of the Niagara River and the water level in Lake Erie was **0.9278661**.

We considered the needs of the various stakeholders for the water level and flow of the Great Lakes, and the impact of water level changes on them is different. To simulate this mechanism, we used five utility functions to quantify the satisfaction of each stakeholder. Combining the GDP and environmental protection needs of the stakeholders' industries, we weighted and summed their utility. On this basis, a single objective nonlinear programming problem is proposed, whose decision variable is the flow through two DAMS in 12 months (24 variables in total), the optimization objective is the weighted utility function, and the constraint condition is based on the drainage operation mechanism. Using historical data, we calculated upper and lower limits for water levels and flows.

We use genetic algorithms to solve this NLP problem. Finally, the optimal water level of the Great Lakes is obtained, and the control strategy of the two DAMS is formulated. Compared with the actual data, our model provides a control scheme that improves the overall benefit. With the exception of environmentalists, who suffered a small loss in benefits, the benefits of almost all other groups increased. One possible reason is that environmentalists have less weight and are at odds with the needs of multiple other stakeholders.

The sensitivity and robustness analysis of the model are carried out. The results show that the model has the characteristics of high precision, good robustness and flexible parameters, which can adapt well to inaccurate input and different conditions. This could be a practical way for the International Joint Commission to better control the Great Lakes water system.

Keywords: Network flow; The Great Lakes; Water level; Utility; Genetic algorithm

Content

1 Introduction

1.1 Problem Background

The Great Lakes are located in central North America and are made up of five adjacent freshwater lakes. The Great Lakes flow roughly from west to east, through many cities in the United States and Canada, and eventually through the St. Lawrence River into the Gulf of St. Lawrence in the Atlantic Ocean.

The water of the Great Lakes is an important natural resource, providing fresh water for nearby cities, shipping lanes [1], and playing a key role in agriculture and industry, and its waters are used for shipping, fishing, drinking, power generation, ecology and many other purposes. Therefore, the water level of the Great Lakes is of concern to a wide variety of stakeholders.[2]

The water level of each lake is determined by the amount of water entering and leaving the lake, and is the result of the complex interaction of atmospheric conditions, water conditions, water resources management and other factors. People can do this by Compensating Works of the Soo Locks at Sault Ste. Marie and the Moses-Saunders Dam at Cornwall, to control the lake level. The effects of rainfall, evaporation, erosion, ice jams and other water flow phenomena need to be considered.

1.2 Restatement of the Problem

Regulating water levels in the Great Lakes is a complex and challenging dynamic network flow problem with changing dynamic needs and conflicting interests among stakeholders, and requires many uncertainties. Through in-depth analysis and research of the background of the problem, combined with the background and constraints of the problem, we need to solve the following problems:

- Problem 1: Model the network of the Great Lakes and their connected rivers, taking into account the needs of different stakeholders, to determine the optimal water level for the Great Lakes at any time of the year.
- Problem 2: Construct a regulation algorithm to maintain the optimal water level of the Great Lakes, and regulate the water output of the two control DAMS according to the inflow and outflow data of the lakes. The sensitivity of the control algorithm to the water yield is analyzed, and the degree to which the model meets the water level needs of various stakeholders is assessed based on historical data.
- Problem 3: Assess the sensitivity of the control algorithm to changes in environmental conditions, e.g., precipitation, winter snow cover, ice jams.
- Problem 4: Focus on Lake Ontario's stakeholders and factors and propose targeted management recommendations.

1.3 Structure of Our Model

The structure diagram of our model is shown in Figure 1.

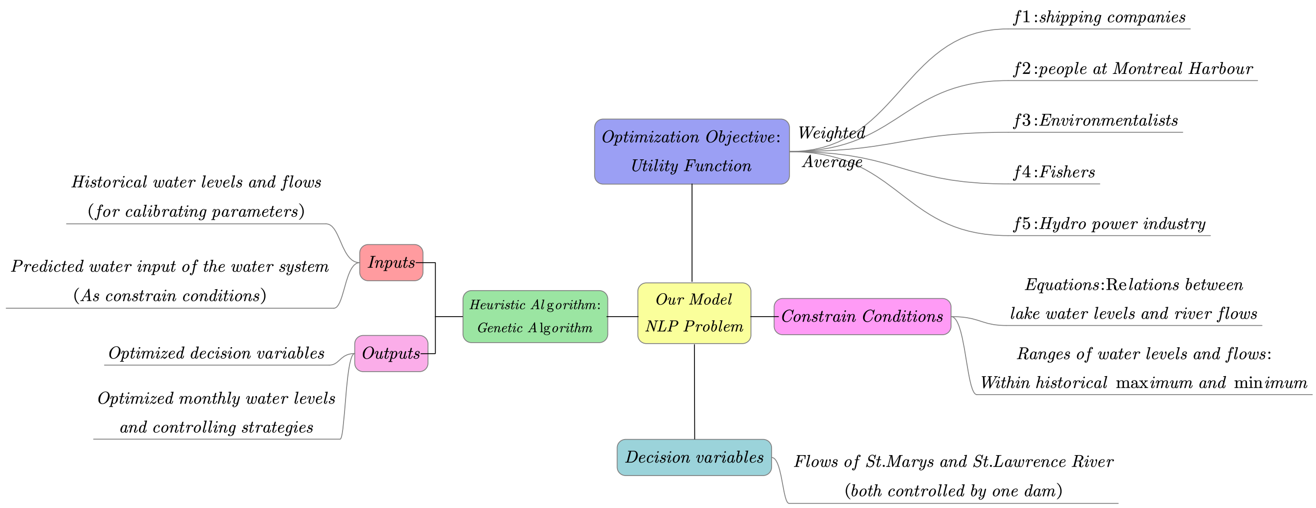


Figure 1: Flow Chart

2 Assumptions

Assumption 1: The rise of fall of water levels is slight, so the changeable water body of the Five Great Lakes can be simplified as cylinders with different bottom area.

Justification1: From the data table *Problem_D_Great_lakes.xlsx*, we can see most water level values of the same lake have a range of no more than 1 meter. When the water rises and falls in such a little scale, the changed volume can be seen as a cylinder, and it's bottom area is equal to the lake's.

Assumption 2: When considering benefits of stakeholders, we only take those whose benefit is related to Lake Ontario into account. As for stakeholders living near other lakes, their problems were relatively minor, so they were ignored in our model.

Justification2: This is a requirement of the question. We need to focus our analyses on stakeholders living near Lake Ontario and its related rivers.

Assumption 3: Each type of stakeholder has their desired water level value in each month, and the desired water level can be deduced using data up to year 2014.

Justification3: From the Addendum we knew that major voices about problems of Lake Ontario started after Plan 2014.So it's very likely that stakeholders might be satisfied with water levels before that year. That is, we could use these former data to evaluate each stakeholder's desired water level value in detail.

Assumption 4: Water input of each lake consists of water from the upstream lake and water from other sources can be calculated accurately.

Justification4: If we do not know how much water will be introduced into the Five Great Lakes system from rainfall, snowpack, groundwater, etc., apparently we are not able to make appropriate control strategies for dams.

Assumption 5: Only 2 major hydro power plants at Niagara Waterfall and Beauharnonis make contributions to the hydro-power generation companies' benefits. Other minor plants near rivers were ignored.

Justification5: We knew if there is a big drop in height, the water will be much more powerful. So we only consider these major power plants when analyzing hydro-power companies.

Assumption 6: Only 2 dams: the Compensating Works and Moses-Saunders Dam are available in strategies to control water levels of lakes and rivers in the whole system. Other minor dams can be ignored.

Justification6: Although there might be minor dams in the middle of the system, they would not be so important as these 2 major ones. The 2 dams are much more giant in size and much more capable in controlling water flow. Also, their geographic locations: the "entrance" and "exit" of the system are vital.

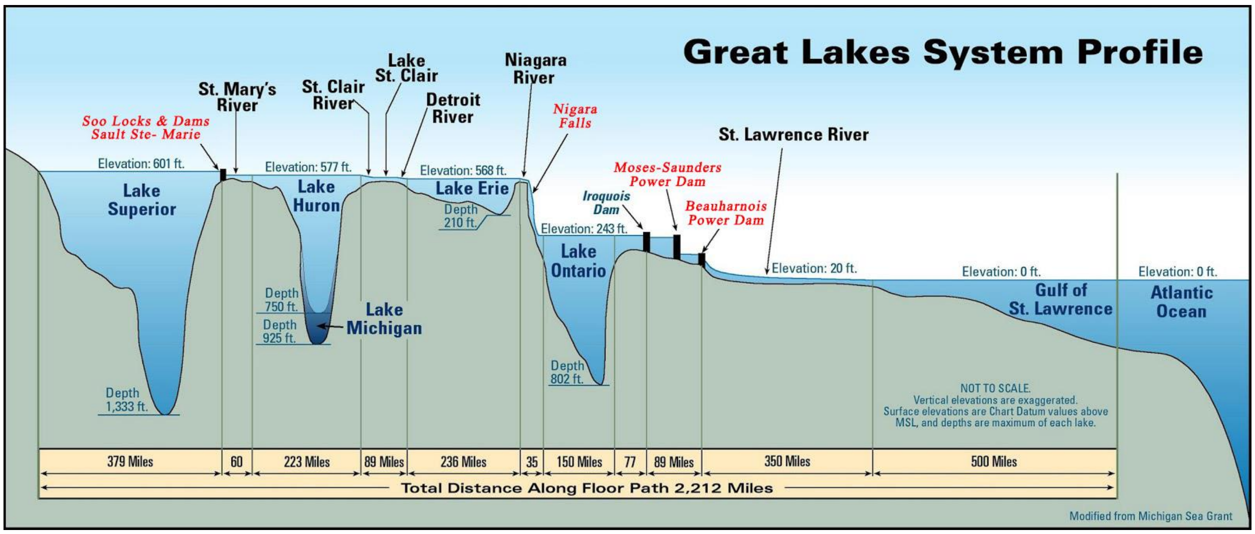


Figure 2: The DAMS and power stations considered in this paper are marked in red

3 Notations

The key mathematical notations used in this paper are listed in Table 1.

Table 1: Notations used in this paper

Symbol	Description
i	Represents the river label, where $i = 1$ represents the St. Marys River, $i = 2$ represents the Detroit River, $i = 3$ represents the Niagara River, $i = 4$ represents the St. Lawrence River, and $i = 5$ represents the Ottawa River
q_i	The discharge of river i

Continuation of table 1

Symbol	Description
j	Represents the lake label, where $j = 1$ indicates Lake Superior, $j = 2$ indicates Lake Michigan and Lake Huron, $j = 3$ indicates Lake Erie, and $j = 4$ indicates Lake Ontario
w_j	Represents the water level of Lake j
S_j	Equivalent simplified cylinder base area of Lake j
k	The month label
I	The water brought into the system by rainfall, snow, underground water, etc.
s	Interest groups are represented, where $s = 1$ means shipping companies, $s = 2$ means residents of the Port of Montreal, $s = 3$ means environmentalists, $s = 4$ means fishing port operators and fishermen, and $s = 5$ means hydroelectric companies
$\Delta h_{A,B}$	The height difference between lake A and Lake B
f_p	The utility of interest group p
f_{i_r}	the monthly utility calculated using actual data
f_{i_m}	the monthly utility calculated using model outputs

4 Modeling the Water System

We started our model from building a network for the water system. As Figure 3 shows, we used a network to model the five lakes and rivers connecting two of them. Because we need to consider stakeholders at Montreal, St. Lawrence River (act as an "exit lake" in fact) and Ottawa River were also taken into account.

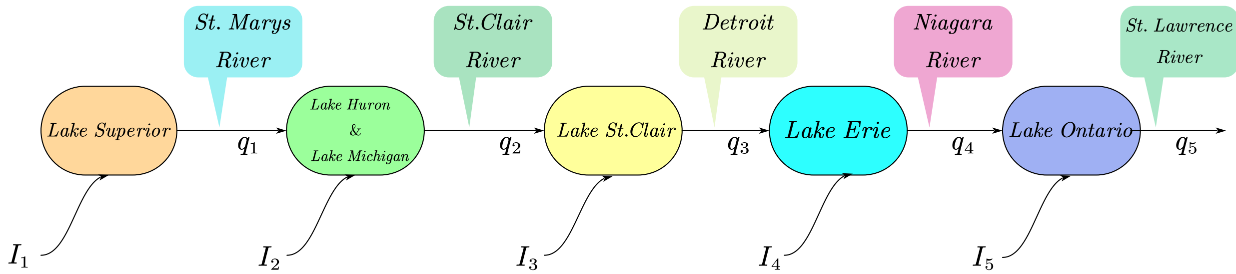


Figure 3: The flow network of the Great Lakes and their connected rivers

In this model, Water is transported from one lake to another through rivers. As shown in Figure 4, each lake receives liquid water from its upstream lake (except for Lake Superior, it does not have a upstream lake) and other sources. Also it gives out its water to downstream lake by river. Although evaporation and other factors will also consume some water in the lake, we merged these lost water into the input from other sources (with negative sign, of course).

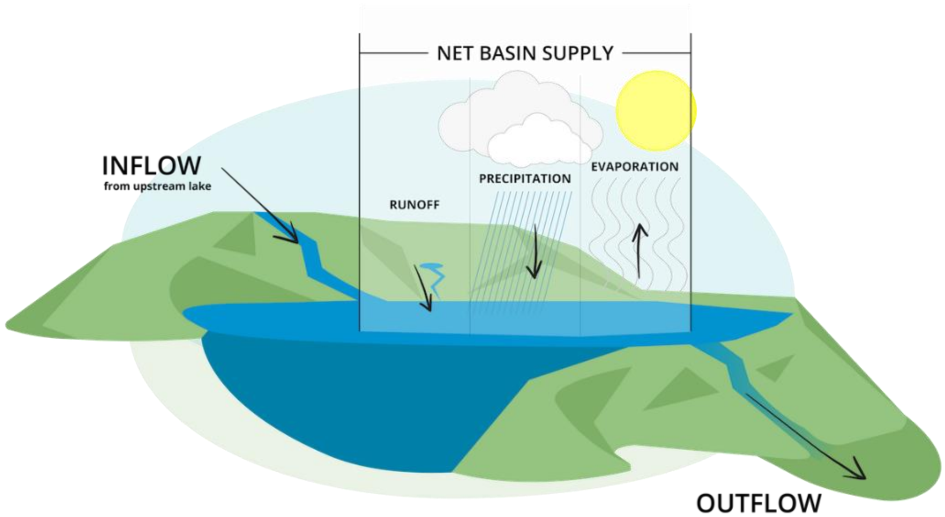


Figure 4: Water level Net basin supply[3]

These theory can also be expressed in formula.

$$\begin{cases} S_j * \Delta w_j = I_j + q_{i-1} - q_i \\ q_i = a_j w_j + b_j \end{cases} \tag{1}$$

Where a and j are parameters; the lake label and the river label in general exactly correspond, in this formula, i often has the same value as j, that is, the river of the same number flows from the lake of the same number

In actual practice, water input from other sources can be estimated using amount of melted snow, rainfall, ground water flows, evaporation rates and other data [4-6]. For sake of convenience, when solving the model, we used historical data to calculate water input directly, and deduced control strategies for that past year (year 2017, to be exact) using the calculated input.

The value we want to optimize is the water level of Lake Ontario. Since water levels of other lakes and flows of rivers will alter this value, it is necessary to reflect on all these decision variables.

Note that although we could take full control of variable w, q using 2 major dams (we assumed these giant structures were adequately capable in controlling water flow), other variables can't be directly adjusted since we do not have major controlling facilities in the middle of the system. However, they could be deduced since when input and output of the subsystem consisted by the latter 4 lakes are clear, the subsystem will reach a determinant state through a self-balancing process.

Now that we know how to control the water, the only problem is: how high are the optimal water levels? Or, how to describe our optimization objectives mathematically?

According to **Assumption 2**, people will have an expectation water level (or a range they hope the level to keep in) for each month, and these expectations or ranges can be deduced using the average, minimum and maximum values in that month up to year 2014.So we first calculated the five stakeholders' expected water level value in each month. Results are shown in Table 2:

Table 2: Expected water level values for each month for the five stakeholders

month	$w_{O,high}$	$w_{O,low}$	$w_{O,mid}$	$q_{s,high}$	$q_{s,low}$
1	75.05	74.33	74.66	33201617267.16	21383106072.51
2	75.12	74.29	74.72	30083553166.43	20119366826.78
3	75.18	74.35	74.76	31257016731.31	21879379869.02
4	75.35	74.62	74.94	36393332092.70	24032954087.75
5	75.80	74.71	75.11	39953702693.19	24075667869.02
6	75.91	74.71	75.15	34928186614.64	20320440755.52
7	75.80	74.79	75.10	34857452109.58	19635596807.55
8	75.53	74.65	74.97	31959423309.58	19240307234.39
9	75.24	74.50	74.78	29026278909.59	18703545643.68
10	75.04	74.34	74.63	28369887685.41	20568077495.87
11	75.04	74.29	74.56	28508031499.50	20512016679.31
12	75.00	74.28	74.57	28916997423.93	21394403314.93

Then, since the actual water level is likely to be different from people's expectation, they might get unsatisfied sometimes. To simulate this mechanic, we used five functions to calculate each stakeholder's degree of satisfaction. Note that when a stakeholder cares about water level of Lake Ontario, the exact value will appear in the function, while for those who care about water levels of rivers, we used the flow of rivers to replace their water levels in formulas since we did not get data about these rivers' average water level.

The shipping companies, who need to operate from Montreal to Lake Ontario, want the water to be high and stable. Therefore, the utility function of shipping companies is as follows:

$$\begin{cases} f_{1,k} = max\{ (\frac{q_{4,k} + q_{Ottawa}}{(q_{4,k} + q_{Ottawa})_{hismax}})^2, 1\} - \frac{CV(q_{4,k})}{CV_{hisaver}(q_{4,k})}, \quad 3 <= k <= 11 \\ f_{1,k} = 1, \ k = 12, 1, 2 \end{cases} \tag{2}$$

where CV represents the coefficient of variation, which is the ratio of the standard deviation of the data to the mean of the data.

The people of the Port of Montreal hated flooding and wanted the water level at the mouth of the Montreal River to be low and stable. Their utility function is as follows

$$f_{2,k} = 1 - max\{(\frac{q_{4,k}}{q_{4,k,hismin}})^2, 1\} - \frac{CV(q)}{CV_{hisaver}(q)} \tag{3}$$

Conservationists want Lake Ontario's water level to change seasonally to help sustain the species. Their utility function can be expressed by the following formula.

$$\begin{cases} f_{3,k} = max\{\frac{Error_{tol}}{100\left|w_{4,k} - w_{4,k,hismax}\right|}, 1\},\ 3 <= k <= 8 \\ f_{3,k} = max\{\frac{Error_{tol}}{100\left|w_{4,k} - w_{4,k,hismin}\right|}, 1\},\ others \end{cases} \tag{4}$$

Fishing port operators and fishermen want the water level to be medium and stable, so their utility function is shown below.

$$f_{4,k} = (\frac{\left|w_{4,k} - w_{4,k,hisavere}\right|}{w_{4,k,hisaver}})^2 - \frac{CV(w)}{CV_{hisaver}(w)} \tag{5}$$

The water and electricity company wants the water level to maintain a certain height difference, and their utility function is shown in the following formula.

$$f_{5,k} = \frac{q\,\Delta h_{Niagara,Onatrio} + q\,\Delta h_{Ontario,St.Lawrence}}{(q\,\Delta h_{Niagara,Onatrio} + q\,\Delta h_{Ontario,St.Lawrence})_{hismax}} \tag{6}$$

The utility function is obtained by weighted average of GDP:

$$f = \sum_{k=1}^{12}\sum_{s=1}^{5} C_s f_{s,k} \tag{7}$$

Where C_s represents each satisfaction coefficient, for the entity industry, its value is equal to the industry's proportion in GDP; For environmentalists, its value can be set according to environmental needs. According to the relevant data [7-10], the value of parameter C_s adopted in this paper is shown in Table 3.

Table 3: Weight values of different stakeholders

Parameter	Stakeholder	Weighted value
C_1	The shipping companies	0.6
C_2	The people of the Port of Montreal	0.02
C_3	Conservationists	0.2
C_4	Fishing port operators and fishermen	0.08
C_5	The water and electricity company	0.1

The CV operator represents the coefficient of variation of the calculation sequence. The variables marked with the subscript $hismax$ are obtained by taking the maximum value of historical data before 2014. Similarly, $hismin$ represents the minimum value of historical data, and $hisaver$ represents the average value of historical data. $Errtol$ means the maximum change in water level that environmentalists can tolerate, and we'll take 0.2 meters (20 centimeters) here.

So these five functions are our optimization objectives. If we consider the problem as a multi-object optimization, the final result would be a Pareto solution set. Each solution is Pareto optimized, meaning they were not better for all stakeholders than any others in this set. However, since solutions in Pareto set might be varied, it was hard to pick one “best” solution and make control strategies for it.

Therefore, we choose to add weights to the five degrees of satisfaction, and transform the problem to a single-object nonlinear programming problem (NLP in short). Based on theories above, here is the **complete mathematical expression of this problem:**

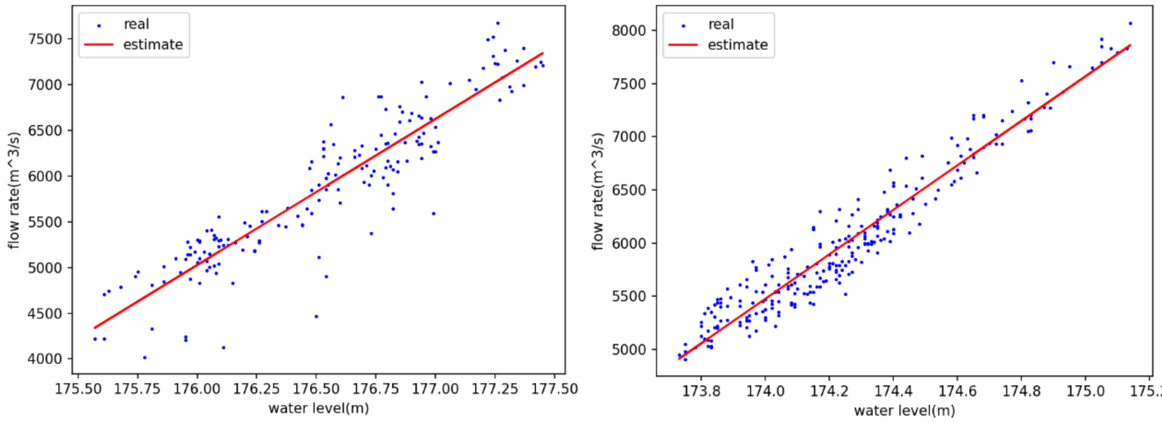
$$max\ f = \sum_{k=1}^{12} \sum_{s=1}^{5} C_s f_{s,k} \tag{8}$$

Note that we regulate the maximum and minimum water levels of lakes to avoid theoretically better but ridiculous solutions. The water level should be no higher than historically highest level and no lower than historically lowest level. After all, no one wants his house washed by overflowing waves, neither will they be happy if the crops were drought to death, right?

Through the above method, we can get the best water level of Lake Ontario in each month. In reality, in order to achieve the optimal level of the lake, it is necessary to regulate the discharge of the two DAMS in the system to affect the flow of the downstream river.

In the process of studying the regulation of the system, we find that there is no dam between Lake Michigan and Lake Huron and Lake Erie that can regulate the flow, and the water flow is completely self-distributed in this subsystem. Therefore, we need to explore the characteristics of the traffic self-distribution of the above subsystems in order to make the system solvable.

By studying the data of monthly discharge of rivers and monthly water level of lakes, we found that River discharge was generally positively correlated with the water level of upstream lakes, especially Lake Michigan and Lake Huron and Detroit River, Lake Erie and Niagara River. We fit it linearly, and the result is shown in the figure below:



(a) Detroit River - Lake Michigan & Lake Huron (b) Niagara River - Lake Erie

Figure 5: Fitting the relationship between river discharge and upstream lake level

If the water level of Lake Michigan and Lake Huron is x and the discharge of Detroit River is y, the functional relationship between the two is approximately as follows:

$$y = 1597.7297x - 276174.0481 \tag{9}$$

The goodness of fit (R^2) is 0.8353428072297593.

Similarly, if the water level of Lake Erie is x and the discharge of Niagara River is y, the functional relationship between the two is approximately as follows:

$$y = 2088.9166x - 357993.8242 \tag{10}$$

The goodness of fit (R^2) is 0.9278661425210013.

From this, we get the characteristics of regulating the flow self-distribution of the neutron system, and then we can construct the system equation of the whole control network [11]. Let's say that in the j month of the year, Compensating Works of the Soo Locks at Sault Ste. Marie controls the flow of St. Mary's River as $q_{1,k}$, the Moses-Saunders Dam at Cornwall controls the flow of St. Lawrence River as $q_{2,k}$; The water level of Lake Superior is $w_{1,k}$, and the runoff flow from precipitation, evaporation, underground and surface runoff is $I_{1,k}$. The water level of Lake Michigan and Lake Huron is $w_{2,k}$, and the runoff from precipitation, evaporation, underground and surface runoff is $I_{2,k}$,. The water level of Lake Erie is $w_{3,k}$,, and the runoff from precipitation, evaporation, underground and surface runoff is $I_{3,k}$. The water level of Lake Ontario is $w_{4,k}$, and the runoff from precipitation, evaporation, subsurface and surface runoff is $I_{4,k}$.

It should be emphasized that in the process of research, we found that the Lake area of Lake St. Clair between Lake Michigan and Lake Huron and Lake Erie is too small, and the linear relationship with the downstream Detroit River is not good. Therefore, we directly established the linear relationship between Lake Michigan and Lake Huron and Detroit River, and approximately regarded Lake Michigan and Lake Huron and Detroit River as direct connection. It skipped the St. Clair River and Lake St. Clair

According to the principle of water flow balance, the following equations can be listed:

$$\begin{cases} S_1(w_{1,k}-w_{1,k-1})=I_{1,k}-q_{1,k} \\ S_2(w_{2,k}-w_{2,k-1})=I_{2,j}+q_{1,k}-1597.7297w_{2,k}+276174.0481 \\ S_3(w_{3,k}-w_{3,k-1})=I_{3,k}+1597.7297w_{2,k}-276174.0481-2088.9166w_{3,k}+357993.8242 \\ S_4(w_{4,k}-w_{4,k-1})=I_{4,k}+2088.9166w_{3,k}-357993.8242-q_{2,k} \end{cases} \tag{11}$$

Where S_1, S_2, S_3, S_4 are the water area of the above four lakes respectively. Taking the above equations as constraints, and then genetic algorithm can be used to solve $w_{4,j}, q_{1,j}$, $q_{2,j}$.

5 Solution of the Model

NLP problems were hard to solve using traditional methods. So we turned to computer and used a popular algorithm called **Genetic Algorithm** (GA in short).

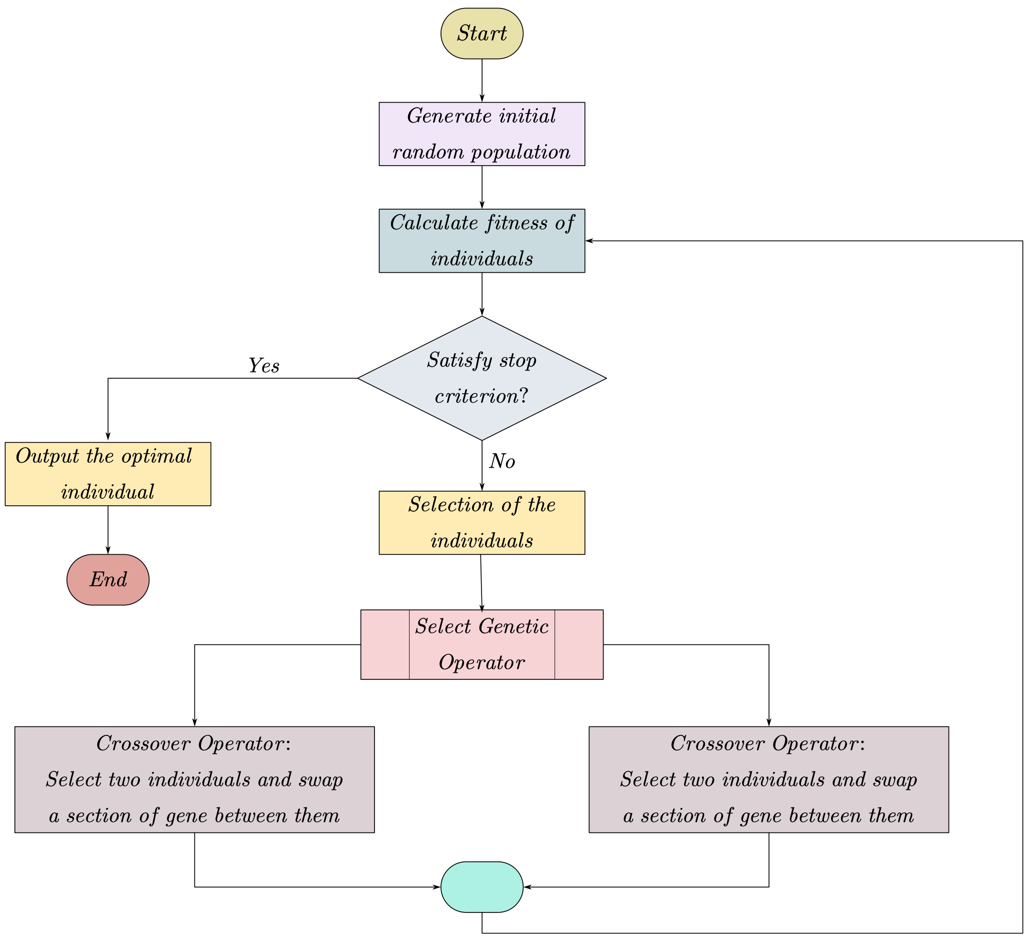


Figure 6: Genetic algorithm flow chart

The procedure of GA is:

- **First,** Generate a random initial population wit 40 individuals. Each individual in the population has a "chromosome" wit 24 genes. Each gene stands for the flow of St. Marys River or St. Lawrence River in 12 months of year 2017(the former 12 genes for flows of St.

Marys River and the others for St. Lawrence River, to be exact). So each individual corresponds with a solution of the problem.

- **Second,** judge each individual's feasibility using the constraint conditions. If an individual is not feasible, it is eliminated. As for feasible solutions, we calculate their utility in order to estimate how satisfying they are. Then all individuals will "breed" and copy themselves. The number of copies of individuals is weighted averaged by their utility values.
- **Third,** pick 40 individuals from the copied population. Since individuals with higher utility are copied more, these "elites" have a much higher chance to be selected in this step.
- **Fourth,** like creatures in nature, these "elites" will hybridize with each other, generating even better offspring. When two individuals hybridize, some of their genes will be exchanged. Also, there is a small chance that some genes mutate and become different from genes in same location in parental generation. This mutation mechanical is introduced to avoid being mislead to partial but not global optimized solutions.
- **Last,** after generating the new generation of population, go to the second step. The loop continues until no obvious better solutions were found in more than the top 100 best "elites", or the generation (times of loops) is beyond 1000.

We used data of year 2017 as input. Note that variables I_i were calculated using historical data. Input data samples are shown in Table 4:

Table 4: Inputs of the model

The unit of $I_{i,j}$ is:cubic meters per month				
Month j	$I_{1,j}$	$I_{2,j}$	$I_{3,j}$	$I_{4,j}$
1	2.7459E+09	1.0007E+10	3.4481E+09	3.6690E+09
2	2.5845E+09	1.3666E+10	2.7209E+09	3.9336E+09
3	8.1746E+09	2.4539E+10	4.9377E+09	9.3544E+09
4	1.8239E+10	2.5878E+10	6.1721E+09	1.1151E+10
5	1.4417E+10	2.0068E+10	3.3517E+09	3.5703E+09
6	1.2813E+10	2.2366E+10	2.1698E+09	5.4875E+09
7	1.0929E+10	9.3825E+09	-4.0711E+08	3.5465E+09
8	1.0196E+10	4.9027E+07	-1.0307E+09	2.4027E+09
9	7.7749E+09	2.2865E+09	-1.9556E+09	3.5244E+09
10	5.4608E+09	6.1346E+09	6.5453E+08	5.6237E+09
11	2.6518E+09	1.5012E+09	-6.9444E+08	2.0194E+09
12	2.5454E+08	3.1349E+09	-3.4065E+08	-1.4574E+12

Then we run the program, coded with Python package Geatpy2. Table 5 shows the results:

Table 5: Outputs of the model

The unit of $q_{i,j}$ is:cubic meters per month
The unit of $w_{i,j}$ is: meters

Month j	$q_{1,j}$	$q_{4,j}$	$w_{1,j}$	$w_{2,j}$	$w_{3,j}$	$w_{4,j}$
1	6.2745E+09	2.1382E+10	183.470	176.470	174.290	74.620
2	5.9877E+09	2.0240E+10	183.452	176.459	174.392	74.547
3	4.1977E+09	2.3426E+10	183.409	176.509	174.448	74.493
4	5.8193E+09	1.6851E+10	183.458	176.620	174.580	74.651
5	6.4672E+09	1.6794E+10	183.608	176.757	174.752	75.247
6	7.0843E+09	1.8926E+10	183.705	176.841	174.798	75.536
7	8.4566E+09	2.1715E+10	183.774	176.951	174.804	75.796
8	8.2743E+09	2.4507E+10	183.804	176.954	174.725	75.847
9	7.5232E+09	2.5909E+10	183.827	176.875	174.639	75.674
10	7.8766E+09	2.4346E+10	183.831	176.817	174.523	75.432
11	3.9248E+09	2.2973E+10	183.801	176.792	174.524	75.375
12	4.4634E+09	2.3739E+10	183.786	176.699	174.469	75.175

The optimized utility f is 2.4162. **Keep the water levels around results above, and the optimized utility should be achieved.** We recommend the dams check if the water level is appropriate every day and adjust their flow accordingly, since it's obvious that we cannot let all redundant water go away at the end of every month.

We compared our results with the actual water level, using the same satisfaction formula set (See Section 4). Here are comparisons of some indexes in Table 6:

Table 6: Comparison of indexes

Month	f_{1_r}	f_{1_m}	f_{2_r}	f_{2_m}	f_{3_r}	f_{3_m}	f_{4_r}	f_{4_m}	f_{5_r}	f_{5_m}
1	1.0000	1.0000	0.8262	0.8667	0.0069	0.0069	-0.0057	-0.0050	0.0678	0.0733
2	1.0000	1.0000	0.8220	0.8551	0.0038	0.0058	-0.0056	-0.0050	0.0648	0.0694
3	1.0000	1.0000	0.7318	0.7090	0.0111	0.0120	-0.0056	-0.0050	0.0786	0.0830
4	-0.1713	-0.0545	0.6251	0.8310	1.0000	0.0088	-0.0056	-0.0050	0.0885	0.0825
5	-0.1731	-0.0486	0.3919	0.7654	1.0000	0.0031	-0.0056	-0.0050	0.0997	0.0890
6	-0.1731	-0.0256	0.3101	0.8075	0.0200	0.0024	-0.0056	-0.0050	0.0904	0.0791

Continuation of table 6

Month	f_{1_r}	f_{1_m}	f_{2_r}	f_{2_m}	f_{3_r}	f_{3_m}	f_{4_r}	f_{4_m}	f_{5_r}	f_{5_m}
7	-0.1731	-0.0418	0.2259	0.7125	0.0182	0.0028	-0.0056	-0.0050	0.0914	0.0826
8	-0.1731	-0.0617	0.3899	0.6892	0.0200	0.0023	-0.0056	-0.0050	0.0859	0.0816
9	-0.1731	-0.0713	0.5223	0.5591	0.0034	0.0024	-0.0056	-0.0050	0.0784	0.0814
10	-0.1697	-0.0723	0.7378	0.7996	0.0038	0.0026	-0.0056	-0.0050	0.0764	0.0782
11	-0.1731	-0.0704	0.6749	0.7211	0.0034	0.0026	-0.0056	-0.0050	0.0789	0.0797
12	1.0000	1.0000	0.7032	0.7720	0.0041	0.0033	-0.0056	-0.0050	0.0807	0.0802

Index	f_{sum_r}	f_{sum_m}
12 Months in total	2.2230	2.4162

Figure 7: Stakeholders' potential reaction towards our controlling scheme

Green values mean our model performed better in this index than real water level, while red ones mean opposite. From the indexes we can see our model did harm to some stakeholders(the environmentalists for example) to get better weighted average performance(illustrated in Figure 7).Although our models may satisfy fishers and shipping companies, it should be reconsidered that if these adjustments will endanger some species. The weight of stakeholders can also be revised to get more balanced schemes.

6 Model Analysis

6.1 Sensitivity Analysis

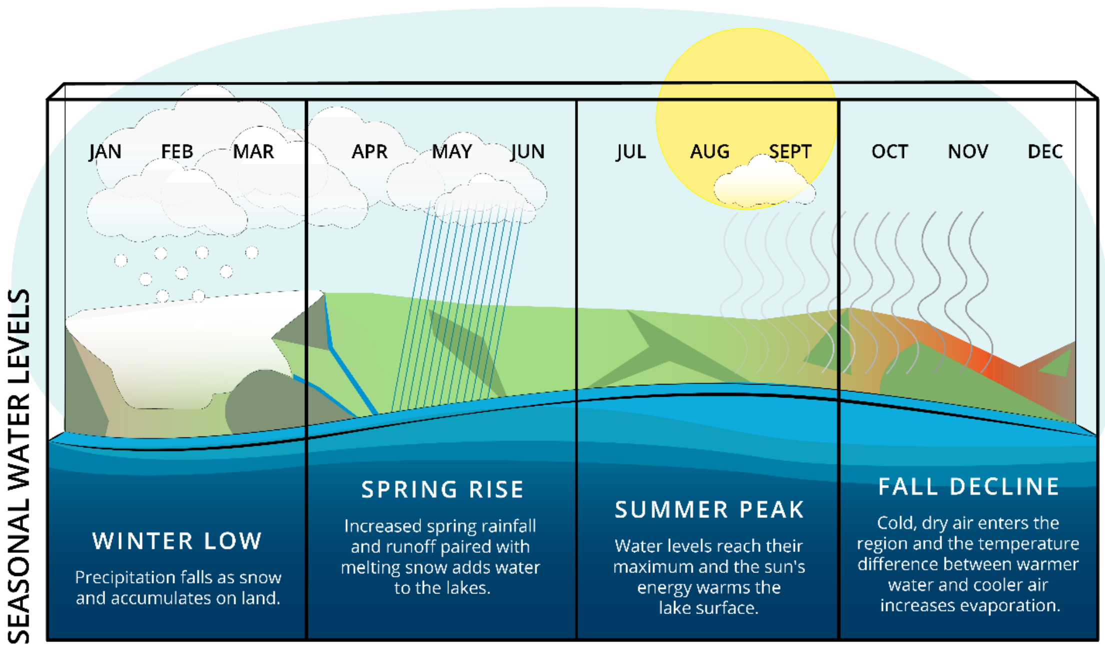


Figure 8: The change of seasons causes the water level to change

We know that the rainfall and melted snow may bring more or less water into the system. So will these changes make results different, and how? To figure out this problem, we conducted sensitivity analysis regarding input variables $I_{i,j}$.In four trials, all input were multiplied by a factor $ratio$,and output optimized utility of the model in these different conditions were compared with the original condition. Table 6 shows the results:

Table 6 Result of Sensitivity Analysis

ratio	New utility f	Changed by
1.0(base)	2.4162	0%
0.8	2.19657	-9.09%
0.9	2.20977	-8.5%
1.02	2.27672	-5.7%
1.05	No feasible Solutions	

We can see that more or less water will both result in the deterioration of optimized solution. When the water input is more than 105%, the model may not work since too much water will definitely exceed the dams' controlling capacity.

6.2 Robustness Analysis

Our model relies on accurate prediction of rainfall and amount of melted snow, etc. to calculate exact input $I_{1,j}$. However, in practice the prediction is always inaccurate to some extent. Will our model deduce a wrong scheme given a set of inputs with some errors? To figure this out, we conduct robustness analysis. In 4 experiments, each factor is multiplied by a random bias factor. The bias factor is subject to Gaussian Distribution $N(1, \sigma)$, and the value of σ varies with experiments to simulate different degree of relative error. Each experiment was repeated 5 times. Table 7 shows the results:

Table 7 Result of Robustness Analysis

σ	Result 1	Result 2	Result 3	Result 4	Result 5	Average	Changed by
0(base)	N/A	N/A	N/A	N/A	N/A	2.4162	0
0.020	2.267	2.345	2.307	2.310	2.302	2.306	-4.5%
0.050	2.214	2.344	2.348	2.132	2.268	2.261	-6.4%
0.070	2.351	2.144	2.376	2.346	2.183	2.280	-5.6%
0.100	2.306	2.133	2.220	2.149	2.236	2.209	-8.6%

We can see that our model could still output a satisfying control strategy even if σ reached 0.1, which means some inputs may be 20% more or less than its exact value. That means that, given an inaccurate prediction of water input, our model is unlikely to be misled to a completely wrong solution (as long as the relative error is not too large). Therefore, the model is robust and practical.

7 Strengths and Weaknesses of our Model

In the following sections, we will describe some of the advantages and disadvantages of the model in detail.

7.1 Strengths

Balanced: This model takes water levels and flows all year long into account, and consider the benefits of all stakeholders, making its solution more generalized.

Flexible: The weights of stakeholders and other parameters can be adjusted to generate schemes fitting with different groups' needs. Also the constrain conditions can be revised so the model can be applied to other water systems.

Robust: Given slightly (relative errors less than 20%) inaccurate prediction values of input, the model can still output a satisfying solution.

Practical: Calculation of the model is fast enough to be deployed in modern computer systems.

7.2 Weaknesses and Prospect

Complex: The model includes a complex nonlinear programming problem without analytical solutions, and it seems that only heuristic algorithms could deal with it. That is, the final solution may be partial and not global optimized, if the parameters of algorithm is not proper.

Arbitrary: Some operators and parameters may be not scientific. For example, the forms of stakeholders' satisfactory functions may not fit well with these people's actual wises. These weakness can be improved if social surveys (RP or SP survey, for example) were conducted to figure out how much water people really want.

Utilitarian: We care only about the GDP they contributed when calculating weights of stakeholders. This may result in disasters of wildlife or dissatisfaction of industries with lower GDP contribution. Of course this problem can be settled by revising the weights according to the government's major goal: to promote economy, to enhance social fairness, or to protect the ecosystem.

8 Conclusion

As the largest group of freshwater lakes, it's really difficult to control the water levels and flows in the Five Great Lakes system properly. Considering different stakeholders' benefits, we created functions to estimate their degree of satisfaction. The total utility is the weighted average of all satisfactory functions. In our solution, the weight depends on GDP contribution of each industry (except for the environmentalists).

Using the utility, we are able to evaluate specific control strategies. Then we analyzed how water run in the system, and deduced lakes' water levels given the flow of St. Marys and St. Lawrence River (both controlled by dams). Then a nonlinear programming problem was proposed. The decision variables were flows passing through the two dams in 12 months (24 variables in total). The optimization objective was utility function f. The constrain conditions were based on the running mechanic of the water system. Upper and lower bounds of water levels and flows were calculated using historical data.

Genetic algorithms are used to solve this nonlinear programming problem. Finally, the optimal water level of the Great Lakes is obtained, and the control strategy of the two DAMS is formulated. Compared with actual data, our model's output scheme was more satisfying for shipping companies and less favorable for environmentalists. Maybe this is because we set the weight of shipping companies relatively large.

Then we analyzed the model. Our model is an accurate and robust model, which is flexible in parameters and can fit in well with inaccurate inputs and different conditions. It might be a practical method for governments of Canada and America to better control the Five Great Lakes water system.

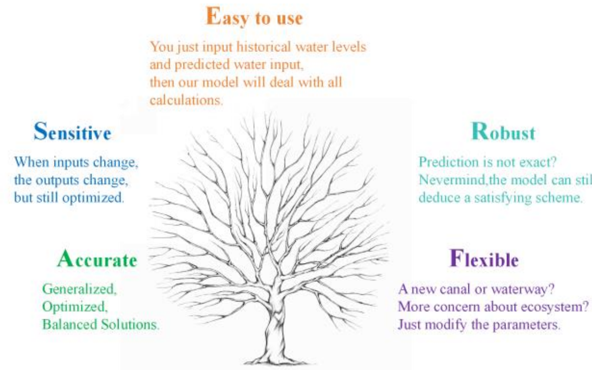
9 Memorandum

To: The IJC leadership

From: Team # 2418844

Subject: Introduction to the characteristics of the Great Lakes water level control model

Date: February 5, 2024



Reasons to choose our model:A,S,E,R,F

Dear IJC Leadership:

I am writing to recommend the water level management model developed by our team to address the water level management challenges facing the Great Lakes Basin. We transform your requirement problem into a dynamic network flow optimization problem, and build the corresponding control method.

First, our model is **accurate**. This model takes water levels and flows all year long into account and consider the benefits of all stakeholders, making its solution more generalized.

Second, our model is **sensitive**. When the model is faced with changes in input data, it can effectively adjust accordingly and maintain stability, and the result is still optimized.

Third, our model is **easy to use**. You just input the historical water level and the predicted water level, and then our model will process all the calculations and provide the appropriate management scheme. It should be emphasized that the model is fast enough to be deployed in modern computer systems. The convenience of the model will greatly help you in your work.

Fourth, our model is **robust**. Our model relies on accurate predictions of rainfall, snowmelt, ice jams, etc., but even if the predicted value of a given input is slightly inaccurate (with a relative error of less than 20%), the model can still output a satisfactory solution. This means that even if the prediction of precipitation and ice loss is not accurate or timely, our model can still provide reasonable, efficient and robust water level management solutions.

Finally, our model is flexible. We consider the utility of different interest groups, and provide variable weighting parameters when adding the utility of different interest groups. You can adjust the weight parameters according to the main goals of the department. At the same time, the utility of each interest group is independent, even if the demand of a certain group changes, you only need to modify the utility function for this group, but do not need to change the utility function of other groups. Also the constrain conditions can be revised so the model can be applied to other water systems.

Thank you for reading our memorandum in your busy schedule, hoping our suggestions are useful to you.

Yours Sincerely,

Team # 2418844

Reference

[1] James A. Polidori, Hannah L. Paulson, Andrew D. Gronewold. (2024). Assessing trends in urban municipal water use across the Great Lakes Basin. Journal of Great Lakes Research, 50(1), 102243.

[2] Chen Jiemin, ZHAO Jiuzhou, Liu Genshui et al. Experience and implications of integrated management of the Great Lakes Basin in North America [J]. Wetland Science,2010,8(02):189-192.

[3] Net-basin supply https://www.flickr.com/photos/noaa_glerl/51912395471/

[4] Great Lakes ice concentration data https://coastwatch.glerl.noaa.gov/statistics/great-lakes-ice-concentration/

[5] Average surface water temperature data https://coastwatch.glerl.noaa.gov/statistics/average-surface-water-temperature-glsea/

[6] Great Lakes evaporation, precipitation, runoff data https://www.glerl.noaa.gov/pubs/tech_reports/glerl-083/UpdatedFiles/

[7] Cargo transportation on the Great Lakes Waterway

https://greatlakes-seaway.com/en/the-seaway/economic-impacts/

[8] Great lakes fisheries http://www.glfc.org/pubs/factsheets/FACT%202_14-0913_HR.pdf

[9] Hydroelectric Power in the US - Market Size https://www.ibisworld.com/united-states/market-research-reports/hydroelectric-power-industry/

[10] The Port of Montreal generates nearly CAD $2.6 billion (US $1.93 billion)

https://www.port-montreal.com/m/rapport-annuel/2018/en/benefits-and-traffic/economic-driver

[11] Hu X J. Research on multi-objective coordinated optimization method of unified network flow in regional integrated energy network [D]. Dalian university of technology, 2021.

Appendices

Appendix 1

Introduce: Calculation variables (including model performance analysis)

```
import geatpy as ea
import numpy as np
import statistics as sta
import pandas as pd
import random as rnd

Data = pd.read_csv("C:\\Users\\Administrator\\Desktop\\Data.csv")
Data = np.array(Data)
Data = Data.tolist()

a = [1597.7297, 1952.9788, 2088.9166]
b = [-276174.0481, -336406.5109, -357993.8242]
W = [[183.47, 176.47, 175.19, 174.29, 74.62]]
for i in range(1, 13):
    W.append([])

S = [82414000000, 117616000000, 1210000000, 25700000000, 19554000000]
I = []
for j in range(0, 12):
    II = []
    for i in range(1, 6):
        II.append(Data[j][i])
    I.append(II)

sigma = 0.1
for j in range(0, 12):
    for i in range(0, 5):
        I[j][i] *= rnd.gauss(1, sigma)

ratio = 1.02
for j in range(0, 12):
    for i in range(0, 5):
        I[j][i] *= ratio

C = [0.6, 0.02, 0.2, 0.08, 0.1]
Errtol = 0.2
Powmax = 46731468830443
wmax = [183.88, 177.45, 176.04, 175.14, 75.91]
wmin = [182.79, 175.57, 174.44, 173.73, 74.28]

Wau = []
```

```
Wal = []
Waa = []
Qsu = []
Qsl = []
q6 = []
q1max = []
q1min = []
q5max = []
q5min = []
for j in range(0, 12):
    Wau.append(Data[j][6])
    Wal.append(Data[j][7])
    Waa.append(Data[j][8])
    Qsu.append(Data[j][9])
    Qsl.append(Data[j][10])
    q6.append(Data[j][11])
    q1max.append(Data[j][12])
    q1min.append(Data[j][13])
    q5max.append(Data[j][14])
    q5min.append(Data[j][15])
print(I)
print(Wau)
print(Wal)
print(Waa)
print(Qsu)
print(Qsl)
print(q6)
```

Appendix 2

Introduce: Genetic algorithm solution

```
def evalVars(Vars):
    f = 0
    dayofmon = [31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31]
    q1 = []
    q2 = []
    q3 = []
    q4 = []
    for j in range(0, 12):
        q1.append(Vars[j])
    q5 = []
    for j in range(12, 24):
        q5.append(Vars[j])
    for j in range(1, 13):
```

```
        q2.append((a[0] * W[j - 1][1] + b[0]) * 3600 * 24 * dayofmon[j - 1])
        q3.append((a[0] * W[j - 1][1] + b[0]) * 3600 * 24 * dayofmon[j - 1])
        q4.append((a[2] * W[j - 1][3] + b[2]) * 3600 * 24 * dayofmon[j - 1])
        w1 = W[j - 1][0] + (I[j - 1][0] - q1[j - 1]) / S[0]
        w2 = W[j - 1][1] + (I[j - 1][1] + q1[j - 1] - q2[j - 1]) / S[1]
        w3 = W[j - 1][2] + (0 + q2[j - 1] - q3[j - 1]) / S[2]
        w4 = W[j - 1][3] + (I[j - 1][3] + q3[j - 1] - q4[j - 1]) / S[3]
        w5 = W[j - 1][4] + (I[j - 1][4] + q4[j - 1] - q5[j - 1]) / S[4]
        w = [w1, w2, w3, w4, w5]
        W[j] = w
    w = []
    for j in range(0, 12):
        w.append(W[j][4])
    for j in range(0, 12):
        q5[j] += q6[j]
        if (j <= 2 or j == 11):
            f1 = 1
        else:
            f1 = ((q5[j] - Qsu[j]) / Qsu[j]) ** 2 - sta.stdev(q5) / sta.mean(q5)
        f2 = 1 - ((q5[j] - Qsl[j]) / Qsl[j]) ** 2 - sta.stdev(q5) / sta.mean(q5)
        if (j <= 1 or j >= 8):
            f3 = Errtol / 100 / abs(w[j] - Wal[j])
        else:
            f3 = Errtol / 100 / abs(w[j] - Wau[j])
        if (f3 > 1):
            f3 = 1

        f4 = ((w[j] - Waa[j]) / Waa[j]) ** 2 - sta.stdev(w) / sta.mean(w)
        f5 = (99 * q4[j] + 68 * q5[j]) / Powmax
        f += (C[0] * f1 + C[1] * f2 + C[2] * f3 + C[3] * f4 + C[4] * f5)

    upperbound = []
    lowerbound = []
    for i in range(0, 5):
        upperbound.append(W[0][i])
        for j in range(1, 12):
            if (W[j][i] > upperbound[i]):
                upperbound[i] = W[j][i]
        lowerbound.append(W[0][i])
        for j in range(1, 12):
            if (W[j][i] < lowerbound[i]):
                lowerbound[i] = W[j][i]
    judge = -1
    for i in range(0, 5):
        upperbound[i] -= wmax[i]
        lowerbound[i] = wmin[i] - lowerbound[i]
```

```
            if (upperbound[i] > 0 or lowerbound[i] > 0):
                judge = 1
        f = np.array(f)
        judge = np.array(judge)
        return f, judge


VarTypes = []
for j in range(0, 24):
        VarTypes.append(0)
LB = []
for i in range(0, 12):
        LB.append(q1min[i])
for i in range(12, 24):
        LB.append(q5min[i - 12])
UB = []
for i in range(0, 12):
        UB.append(q1max[i])
for i in range(12, 24):
        UB.append(q5max[i - 12])
problem = ea.Problem(name='Task 1',
                                        M=1,
                                        maxormins=[-1],
                                        Dim=24,
                                        varTypes=VarTypes,
                                        lb=LB,
                                        ub=UB,
                                        evalVars=evalVars)

algorithm = ea.soea_SEGA_templet(problem, ea.Population(Encoding='RI', NIND=40),
MAXGEN=1000,
                                        logTras=10, trappedValue=1e-6,
maxTrappedCount=100)
res = ea.optimize(algorithm, seed=None)
```