Problem Chosen	2022	Team Control Number
B	**MCM/ICM Summary Sheet**	**2206933**

Water and Hydroelectric Power Sharing

The most severe drought in U.S. in the past 800 years occurred in the Colorado Basin, which brings about acute shortage in water and hydropower supplied by reservoirs and dams in this region. The aim of this report is to provide an optimal allocation plan of water and hydropower supplied by the Glen Canyon dam (Lake Powell) and the Hoover dam (Lake Mead) to state AZ, CA, WY, NM, and CO. We are also expected to properly handle the conflicting interests among sectors and among parties involved. Two models are established: Model I: State Allocation Model, Model II: Sector Allocation Model.

For Problem 1, we set up the State Allocation Model by several steps. For preparation, we simulate the structure of the dam, ascertain the demands of resources, and then formulate the transport loss and the hydropower generation function. Then by establishing the Linear programming model, we figure out the optimal allocation scheme (shown in Figure 6), making sure the least water consumption. Next, difference equations are employed to determine the relationship between time, depth, and consumed volume. When the supply nearly runs out, as the initial water level M and P take various possible values, there are 3 cases with different allocation plans, supplying volume, supplying durations, and demanded replenishment. Results are displayed in Figures 6, 7, and 8.

For problem 2, in order to establish the Sector Allocation Model, we first process the data by Lagrange interpolation and validate the interpolation results using the entropy method. Secondly, we set the Spearman correlation coefficient as the criteria to allocate water for each usage and then calculate the correlation between agricultural, industrial, residential water, electricity usage, and GDP for each state. Finally, Table 8 shows how much water supply should be deducted in agricultural, industrial, residential water, electricity generation in water shortage.

For problem 3 we employ the goodness of fit and optimal approximation to determine the function between economy and total water and electricity usage in each state. Based on the average change of function, we apportion the water shortage to each state in water and electricity usage, as is displayed in Table 10.

For Problem 4, we do a robustness test for Model I, in which a disturbance in the independent variables in various scenarios changes the allocation weights. But the allocation plan has no apparent change, showing strong stability. Next, we do sensitivity analysis to Model II. Different variables have different degrees of impact on the plan. Due to its diversity, results are listed in Section 7.

In summary, the paper provides thorough and optimal solutions under diverse initial circumstances. With the consideration of loss and efficiency, models can be accurately applied to general dam supply problems.

keywords: Linear programming; Difference equation; Correlation analysis; Classified discussion

Contents

1 Introduction

1.1 Problem Background

Rivers and dams have always played an essential role in supplying water for multiple uses, controlling floods, and generating hydropower. However, “the past 22 years have been the driest 22 years since 1200 in the Midwest, particularly in the Colorado Basin”, said Patti Aaron, spokesman of the Bureau of Reclamation. The U.S. federal government also declared for the first time that the Colorado River, which supports 40 million people in the western United States, has entered the “Level 1 water shortage” phase. According to the U.S. Drought Monitor, nearly 95 percent of the West is experiencing drought, as is shown in Figure 1 below:

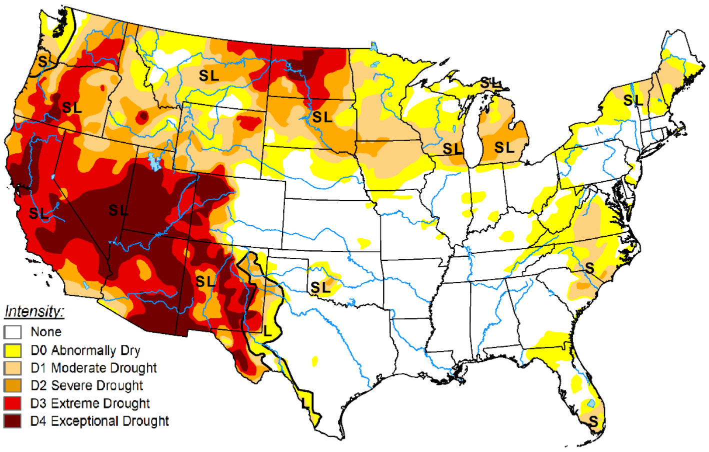


Figure 1: Drought Monitor of the United States Mainland (2021)

Badly still, major rivers and reservoirs, such as the Glen Canyon dam (Lake Powell) and Hoover dam (Lake Mead), are suffering low water levels, which lead to water insufficiency, hydroelectric generation pause, and severe environmental disasters.

Therefore, it is necessary and urgent for us to design an up-to-date, rational, and sustainable plan to allocate water and electricity provided by the Glen Canyon dam and Hoover dam to agriculture, industry, and residences in Arizona (AZ), California (CA), Wyoming (WY), New Mexico (NM), and Colorado (CO), while considering the claim of Mexico downstream.

1.2 Restatement of the Problem

Water and hydropower sharing is an impactful problem with various complexities, related to the conflicting interests and dialectical relationships among different types of water use, among regions, and among countries. Through in-depth investigation and analysis on the topic, combining the basic setting and constraints of the scenario, we can restate the problem as follows:

- Build a mathematical model to allocate fixed water supply and electricity to fixed demands from the 5 states. Figure out how long the adequate supply can sustain and how much additional water needs to be provided after that.

• Based on the model, set standards for allocating water to general (agricultural, industrial, residential) usage and hydroelectricity production.
• What should be done if the supply of water and electricity does not meet the demands?
• What does your model demonstrate under the circumstances of:

⋆ different demands due to population, agricultural, and industrial expansion or contraction.

⋆ increased share of renewable energy technologies.

⋆ other effective water and electricity saving measures.

• Considering the results and conclusions obtained above, prepare one- to two-page article for a magazine on Southwest water-saving.

Noticeably, your solution should consider the coordinating levels of the two lakes, the claim of Mexico, and the outflowing water to the sea. Determine how often your model should be re-run due to supply and demand changes.

1.3 Our Work

The problem requests us to provide optimal water and hydropower allocation plan with the least water resource consumption. Our work mainly includes the following:
• Based on data about supplies, demands, and so on, a model is established to allocate water and electricity from the 2 dams to the 5 states rationally and conservatively.
• By difference equations, the paper determines the relationship among supply period, water depth, and residual volume.
• Through correlation analysis, optimal approximation, and the entropy method, the relationship between resource (water and electricity) usage and economy.

In order to avoid complex and redundant descriptions, the flow chart is displayed:

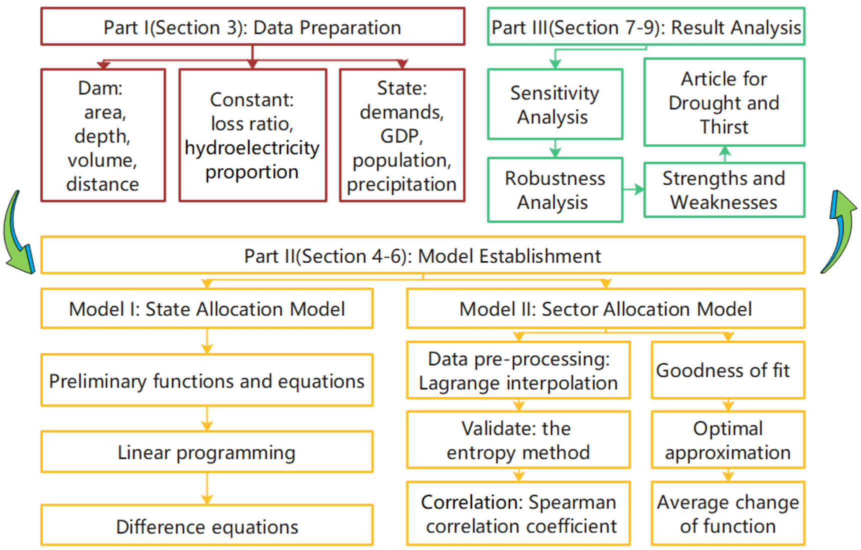


Figure 2: Flow Chart of our work

2 Assumptions and Justifications

Because there are numerous complex situations and factors in reality, we need to make plausible assumptions to simplify the model. Each hypothesis below is supported by proper justification:

• **Assumption 1:** Each community of interest has fixed demands in water and electricity while being supplied by the 2 dams.

Justification: Due to climate and economic circumstances and cycles, demands for water and electricity can be seen as unchanged for a certain length of time. When the demands enter the next phase, we need to adjust and re-run our model, which is presented in

• **Assumption 2:** All water supplied to the 5 states is from the 2 dams. In other words, no additional water (e.g.: rainfall, water from other rivers and etc.) is provided.

Justification: Actually, in the American Southwest, more than 90% water supply comes from the Colorado River system. Besides, additional water supply can be offset by water loss including evaporation, seepage, waste, etc.

• **Assumption 3:** Lake Powell only releases water but has no replenishment. Lake Mead releases water and in the meantime, acquires water that takes part in hydropower generation in Lake Powell.

Justification: Water inflow from upstream and underground can be offset by evaporation, underground outflow, etc. Furthermore, for the reason that Lake Powell is located upstream of Lake Mead, water for generation released from Lake Powell will flow along the river and end up in Lake Mead.

• **Assumption 4:** Water can only be utilized for one of the four usages: agricultural, industrial, residential, and hydropower generation, and cannot be recycled to be put into its second allocation.

Justification: According to the problem requirements, there are conflicting interests in water usage in the four fields. In addition, water that pushes turbines to generate electricity may not have the height and routine to be transported for general use.

We make additional local hypotheses in the following individual sections. They will be mentioned and illustrated at proper locations.

3 Model Preparation

3.1 Notations

Some important notations used in the paper are listed in Table 1:

Table 1: Notations

symbol	description	unit
i	$i=1,2$ Lake Powell, Lake Mead respectively	——
j	$j=1,2,3,4,5$ AZ, CA, WY, NM, CO respectively	——
Ow_{ij}	General Water output from Lake i to State j	m^3
Oe_{ij}	Electricity output from Lake i to State j	kwh
Ve_{ij}	Power Water (consumed for Oe_{ij})	m^3

Tw	Water Mead receives from Powell	m^3
Iw_{ij}	Water input to State j from Lake i	m^3
Ie_{ij}	Electricity input to State j from Lake i	kwh
d_{ij}	Distance between Dam i and State j	km
d_0	Distance between 2 dams	km
t	Time since the allocation process starts	d
V_{0i}	Initial volume in Lake i	m^3
V_{1i}	Available volume in Lake i	m^3
V_i	Volume of Lake i in the allocation process	m^3
V_{res_i}	Residual volume in Lake i	m^3
S_{0i}	Surface area of Lake i	km^2
H_{0i}	Initial depth in Lake i	m
H_{1i}	Available depth in Lake i	m
h_i	depth of Lake i in the allocation process	m
P	Given water level of Powell	m
M	Given water level of Mead	m

Note: Some notations are not listed here and will be explained in detail in the corresponding section.

3.2 Data Collection

The main data we use in modeling include:

• the total and active capacity, the depth, and the surface area of the two lakes.

• the linear distance of the five states from the two dams and the course length between the dams.

• time-series data for the five states on water and electricity demand by sector, GDP by sector, population, and precipitation.

• the energy conversion efficiency of hydroelectric power generation.

The data sources which we referred to are displayed in Table 2.

Table 2: Data Sources

Database Names	Database Websites	Data Type
USGS	https://www.usgs.gov/	Geography
Wikipedia	https://www.wikipedia.org/	Geography
National Weather Service	https://www.weather.gov/	Meteorology
Bureau of Reclamation	https://www.usbr.gov/	Infrastructure
BEA	https://www.bea.gov/	Economy
EIA	https://www.eia.gov/state/	Economy
Wind	https://www.wind.com.cn/	Industry Report
Google Scholar	https://scholar.google.com/	Academic Paper
CNKI	https://www.cnki.net/	Academic Paper

4 Model I: State Allocation Model

4.1 Reservoir Schematic Diagram

By reviewing the information, we obtained the following useful data on Lake Powell and Lake Mead:

Table 3: Lake Information

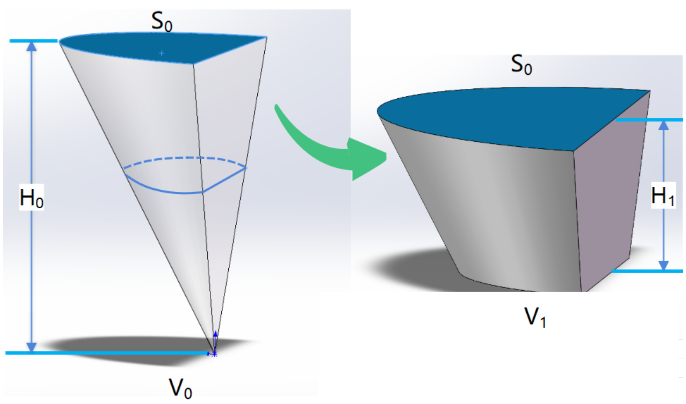
	Lake Powell	Lake Mead
Average depth H_0	$H_{01} = 136.78m$	$H_{02} = 149.63m$
Surface area S_0	$S_{01} = 658km^2$	$S_{02} = 640km^2$
Total Capacity V_0	$V_{01} = 30km^3$	$V_{02} = 31.92km^3$
Active Capacity V_1	$V_{11} = 14.4km^3$	$V_{12} = 16.43km^3$
Available depthH_1	$H_{11} = 26.79m$	$H_{12} = 32.04m$

Plug H_0 and S_0 into the volume formula of semi-cone:

$$V = \frac{1}{3}SH \tag{1}$$

where S is the bottom area of the semi-cone and H is its height. We find the calculated volume is close to the authentic volume of reservoir V_0. Besides, it is documented that the deeper cross section of a lake usually has a smaller horizontal area. In addition, dam stands upright in the water.

Therefore, we can deem the water in the reservoir as a semi-cone. Because water level falls below the available depth H_1, water cannot be utilized. Thus, available water should be considered as a semi-frustum of a cone. In other words, the active capacity V_1 is the volume of a semi-frustum of a cone.



(a) Reservoir: semi-cone (b) Elec-available part:semi-frustum of a cone

Figure 3: Reservoir Schematic Diagram

Due to the triangle similarity principle in mathematics, the volume of cone is proportional to the third power of its height, i.e.:

$$\left(\frac{H_0 - H_1}{H_0}\right)^3 = \frac{V_0 - V_1}{V_0} \tag{2}$$

Consequently, we can figure out the available depth H_1, as displayed in Table 3.

4.2 Regional Information Overview

The locations of Arizona (AZ), California (CA), Wyoming (WY), New Mexico (NM), Colorado (CO), the Glen Canyon dam (Lake Powell), and the Hoover dam (Lake Mead) are shown in Figure 4:

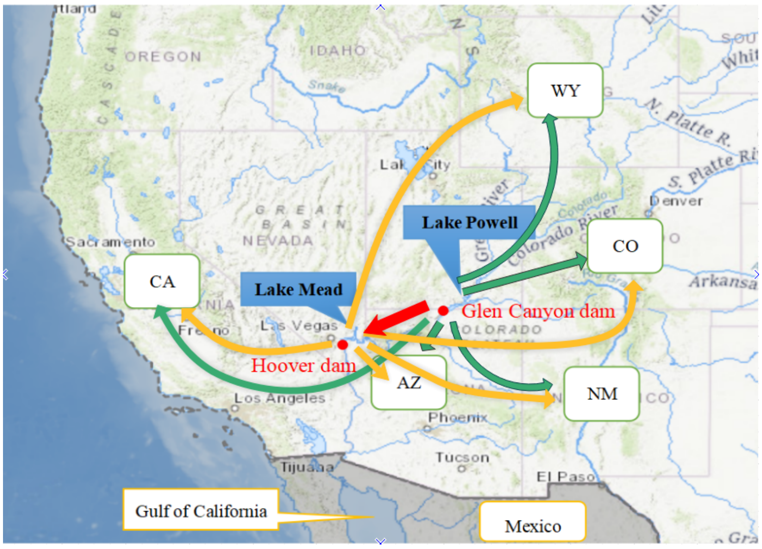


Figure 4: Allocation Schematic Diagram

We assume water and electricity are transported in straight lines, considering the configurations of pipelines and wires. However, water released by Lake Powell flows along the river to Lake Mead, so it runs a course length.

By referring to databases, we acquired the following distance data:

Table 4: Relevant Distances

	Glen Canyon dam (km)	Hoover dam (km)
AZ	$d_{11} = 389.7$	$d_{12} = 373.8$
CA	$d_{21} = 901.4$	$d_{22} = 664.2$
WY	$d_{31} = 742.6$	$d_{32} = 1032$
NM	$d_{41} = 515.3$	$d_{42} = 792.6$
CO	$d_{51} = 646.4$	$d_{52} = 950.2$
	Course distance between the 2 lakes	$d_0 = 446km$

We also find the data of water usage (2015) and electricity usage (2020) by state and sector. Hydroelectricity accounts for $prop = 6.98\%$ of the entire electricity for usage. After disposal and summarization, average daily water and electricity usage in the 5 states are listed below:

Table 5: Water and Electricity Daily Consumption

	Water (10^7 m3)	Electricity ($10^5 Mwh$)	
		All	Hydro-
AZ	2.2341	2.2455	0.1567
CA	9.5577	6.8541	0.4784
WY	0.1181	0.4200	0.0293
NM	1.0835	0.6788	0.0474
CO	3.8948	1.5356	0.1072

Note: Mwh means million watt hour.

4.3 Linear Programming Model

4.3.1 Loss Function

We hypothesize that the transport loss of water and electricity is related to the transport distance and amount through the following equation:

$$Iw_{ij} = Ow_{ij}(1 - k_w d_{ij}) \tag{3}$$

$$Ie_{ij} = Oe_{ij}(1 - k_e d_{ij}) \tag{4}$$

in which $i = 1,2$ representing Lake Powell and Lake Mead respectively, and $j = 1,2,3,4,5$ representing the state AZ, CA, WY, NM and CO in order. Iw_{ij} means the water input to state j from of lake i. Similarly, Ie_{ij} means the electricity input. Ow_{ij} means the water output from lake i to state j, and Oe_{ij} the electricity. k_w and k_e are ratios of water and electricity loss respectively. According to sources and information, they are constant:

$$k_w = k_e = 0.001\% \tag{5}$$

In addition, as we assumed, water for hydropower production released by Lake Powell runs in the Colorado River to Lake Mead. Denote k as the loss ratio in transport. It is also a constant:

$$k = 0.01\% \tag{6}$$

Thus,

$$Tw = Ve_1(1 - kd_0) \tag{7}$$

where Tw means the volume of water that Lake Mead receives from Lake Powell, and Ve_1 means the volume of water that Lake Powell uses to generate electricity and then transports to Lake Mead.

Based on the relationship above, we obtain the water volume of each lake during the allocation process:
For Lake Powell,

$$V_{res_1} = V_{01} - Ow_1 - Ve_1 \tag{8}$$

For Lake Mead,

$$V_{res_2} = V_{02} - Ow_2 - Ve_2 + Tw \tag{9}$$

4.3.2 Hydropower Generation Function

According to the energy conservation law, the gravitational potential energy of the falling water converts into hydroelectric power at a certain conversion efficiency η, which can be estimated through the following data:

Table 6: Separate efficiency in steps

Process	Separate efficiency
Entry	$\eta_1 = 95\%$
Turbine	$\eta_2 = 85\%$
Generator	$\eta_3 = 93\%$

Therefore, we obtain the conversion efficiency η by:

$$\eta = \eta_1 \cdot \eta_2 \cdot \eta_3 \tag{10}$$

We denote the gravitational potential energy as G and follow the previous notation Oe_{ij} to denote the hydroelectric power generated and then output. Then we can list the following equations:

$$G = \rho gh \cdot Ve_{ij} \tag{11}$$

$$Oe_{ij} = G\eta \tag{12}$$

in which ρ represents the water density $1 \times 10^3 kg/m^3$, g denotes the acceleration of gravity $9.8m/s^2$, Ve_{ij} means the volume of water generating Oe_{ij} amount of electricity.

4.3.3 Linear Programming

• Decision Variables

Following the notations above, we use Ow_{ij} and Ve_{ij} to denote the volume of water from Lake i consumed for general usage and electricity generation in state j respectively. They are decision variables, which will be included in our allocation plan and have a lot to do with the future economy and climate.

• Constraints Conditions

$$Iw_{1j} + Iw_{2j} = Iw_j \tag{13}$$

$$Ie_{1j} + Ie_{2j} = Ie_j \tag{14}$$

which manifests for each 5 states ($j = 1,2,3,4,5$), lake 1 and lake 2 provide all the water and hydropower needed. Iw_j and Ie_j can be found in Table 5.

$$Ow_{ij} \geqslant 0 \qquad Ve_{ij} \geqslant 0 \tag{15}$$

$$V_{res_1} \geqslant 0 \qquad V_{res_2} \geqslant 0 \tag{16}$$

which means there are residual water in the 2 lakes.

Besides, equation (3)(4)(5)(6)(7) above are constraints, too.

• **Objective Function**

$$\min \sum_{i=1}^{2} \sum_{j=1}^{5} (Ow_{ij} + Ve_{ij}) \tag{17}$$

which looks for the best allocation approach to guarantee the least water consumption from the 2 lakes.

4.4 Difference Equations

Now we try to find the relationship among time t, depth h, and water volume V.

As water flows out from the dam, the water level drops, which leads to the decrease of its gravitational potential energy. In addition, when the same amount of water is released out, water depth decreases more quickly if the water level is lower. Therefore, we need to split the water output period into small time fragments and investigate the equations through differential methods.

We denote a normal time fragment as Δt. In Δt, dam outputs water for general usage whose volume is Vw and water for hydropower generation whose volume is Ve. V_i means the total output. Thus, we have the following equations:

$$Oe = \eta \rho g h \cdot Ve \tag{18}$$

where Oe means the output volume of water for hydropower production and other notations are the same with those in equation (6) and (7).

For Lake Powell,

$$V_1 = Vw_1 + Ve_1 \tag{19}$$

For Lake Mead,

$$V_2 = Vw_2 + Ve_2 - Tw \tag{20}$$

Define the constant $rate$ by:

$$rate = \frac{\sqrt{S_0}}{H_0} \tag{21}$$

We assume at time t water depth is h. After Δt, depth decreases by Δh, as is shown in Figure 5:

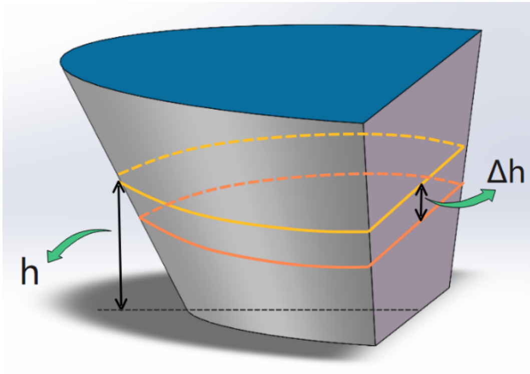


Figure 5: Schematic Diagram of ΔV and Δh

Therefore, based on formula 1 and 2, we calculate the ΔV by the difference between big semi-cone with height h and small semi-cone with height $h - \Delta h$:

$$\Delta V = \frac{rate^2}{3}\left[h^3 - (h - \Delta h)^3\right] \tag{22}$$

by which we solve Δh as follows:

$$\Delta h = h - \left[h^3 - \frac{3}{rate^2}\Delta V\right]^{\frac{1}{3}} \tag{23}$$

The residual amount of water in Lake Powell and Lake Mead during the allocation process is expressed in Formula (18) and (19).

Finally, based on all the discussions and formulas above, we list out the difference equations:

For Lake Powell,

$$V_{res_1} = V_{01} - \int \Delta V_1 \tag{24}$$

$$h_1 = H_{01} - \int \Delta h_1 \tag{25}$$

For Lake Mead,

$$V_{res_2} = V_{02} - \int (\Delta V_2 - \Delta T w) \tag{26}$$

$$h_2 = H_{02} - \int \Delta h_2 \tag{27}$$

So far, we find the relationship among time t, depth h, and residual water volume V_{res}. Results are shown in the next session.

4.5 Results and Analysis

4.5.1 In Sufficient Supply Stage

By solving the Linear Programing Model, we figure out Ow_{ij} and Oe_{ij}, which mean the output of water(for general use) and electricity from Lake i to State j. When there is available water both in Lake Powell and in Lake Mead, the supply of water (for general usage) and electricity from each lake to each state are shown in Figure 6:

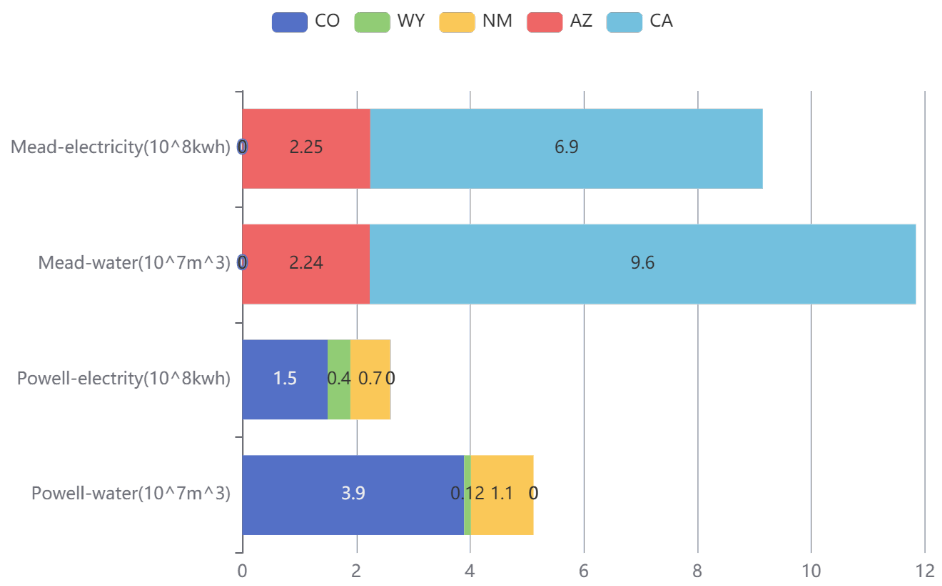


Figure 6: Output of general water and electricity from each lake to each state

The supply of water (for general usage) and electricity from each lake to each state are shown in Figure 6. In sufficient supply stage, Lake Powell provides $5.1283 \times 10^7 m^3$ general water and 2.6511×10^8kwh electricity to WY, NM and CO in total. For Lake Mead, the numbers are $11.8641 \times 10^7 m^3$ and 9.1538×10^8kwh to AZ and CA in total.

4.5.2 In Run-out Supply Stage

Depending on the initial water levels M and P, the end states need to be classified and discussed in three cases.

• **Case 1: Available depths of Powell and Mead reduce to 0 at the same time.**

Lake Powell supplies the water and electricity for three states WY,NM,CO, as is shown in equations (19)and (22).

Lake Mead supplies the water and electricity for two states AZ,CA. In the meanwhile, it receives generation-used water Tw from Lake Powell. Therefore, we have equations (20) and (22).

By marking the position of M and P on the y-axis, we can read the corresponding x-axis number: t_M and $t_P$to find out how released volume and depth change over time and how long it will take before the demands are not met. Besides, we can get the additional supply amount from time t to $t+n$ by reading the number on the y-axis which corresponds with x-axis number t_M+n and t_P+n.

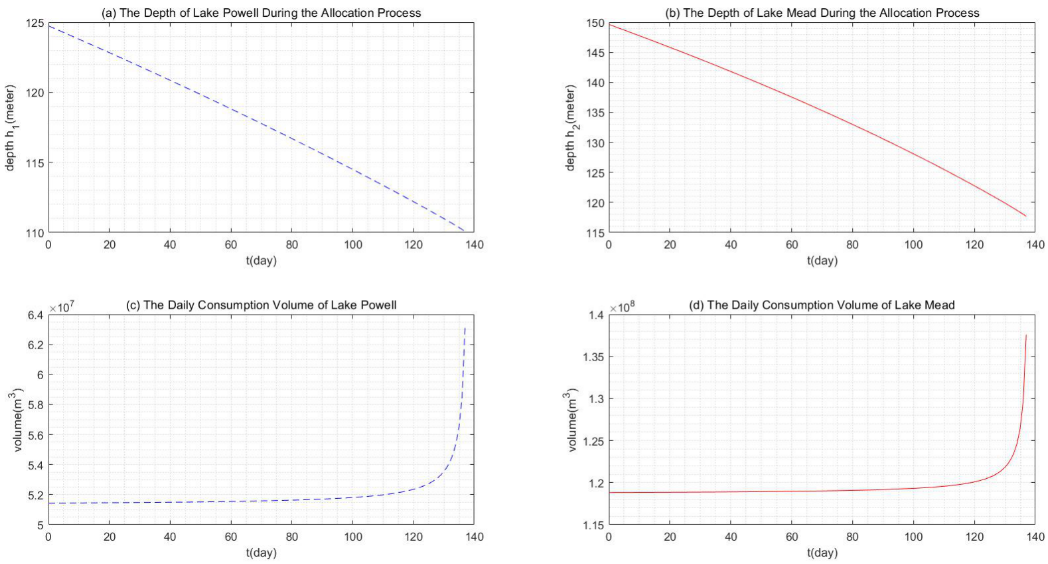


Figure 7: case1

- **Case 2: Available depth of Powell reduces to 0 first.**

  Denote t_1 as the moment when available water in Powell runs out.

  Before t_1, things go the same as Case 1.

  After t_1, the Hoover dam provides water and electricity to all the 5 states AZ,CA,WY,NM,CO as is shown in equations (19) and (22).

  Using the same reading method illustrated in Case 1, we can find the requested results.

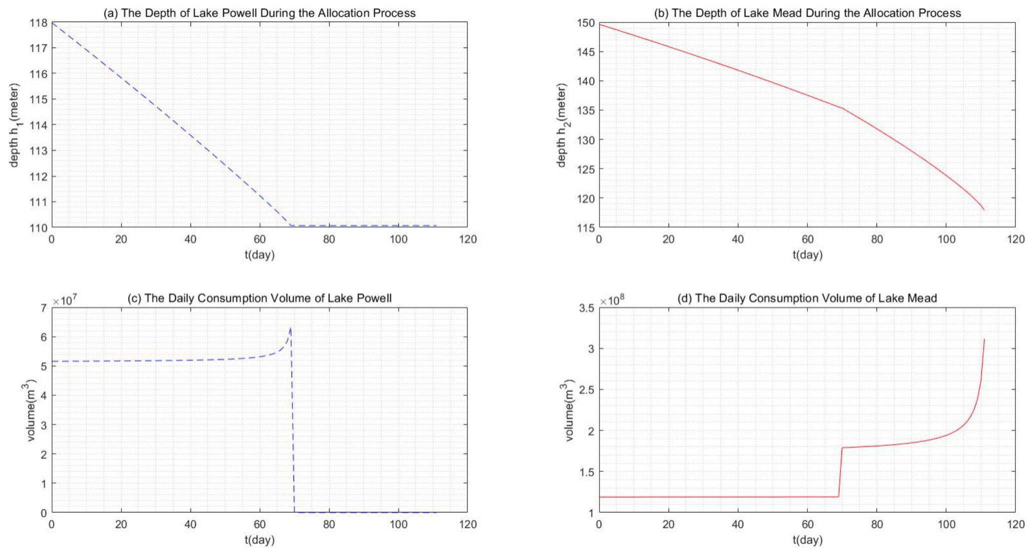


Figure 8: case2

- **Case 3: Available depth of Mead reduces to 0 first.**

  Denote t_2 as the moment when available water in Mead runs out.

  Before t_2, things go the same as Case 1.

After t_2, Lake Powell provides supplies for AZ,WY,NM,CO and a part of CA, as is demonstrated by equations (22) and (28).

$$V_1 = Vw_1 + Ve_1 - Tw\frac{1-k_e d_{22}}{1-k_e d_{12}} \tag{28}$$

Lake Mead accepts generation-used water Tw from Lake Powell and supplies the other part of CA's demands. CA is chosen by Lake Mead for the reason that based on the nature of the hyperbola, the greater the difference of distance, the greater the degree of optimization. Thus, we get the equation:

$$V_2 = Tw \tag{29}$$

Using the same reading method illustrated in Case 1, we can find the requested results.

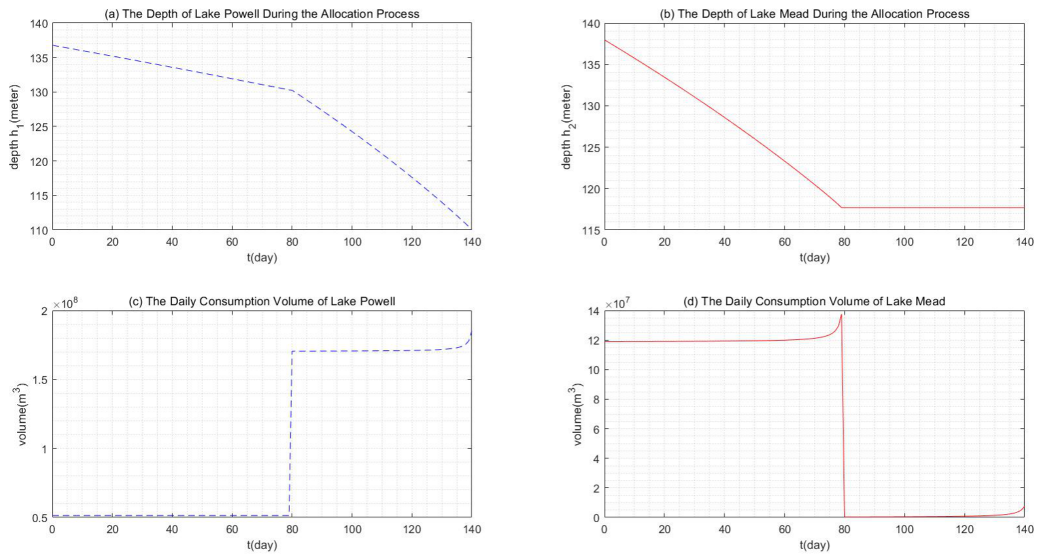


Figure 9: case3

4.5.3 Supplement

• Take the interest of Mexico into consideration. Based on the relationship between depth and volume of consumed water (in section IV), if the amount of water from rainfall or river replenishment exceeds the daily consumption of both dams, then the excess can be delivered to Mexico.
• For the Gulf of California, the generation-used water released from the Hoover dam is discharged into the Gulf of California, whose volume is shown below:

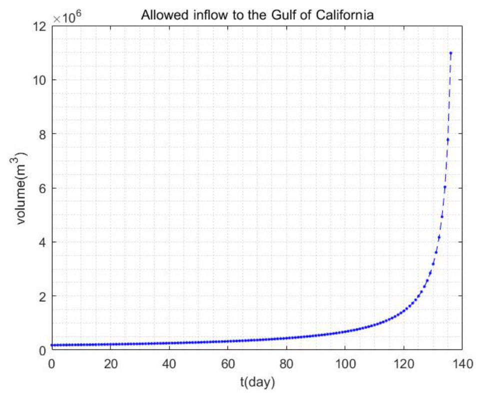


Figure 10: Volume of water discharged to the Gulf of California over time

• As the time-series data we collect about precipitation, water usage, and electricity usage displays, there is an annual cycle and these indicators in an entire year can be split into 2 parts according to their sizes. This means constants used in this model have nonnegligible changes every 6 months. Therefore, the model needs to be re-run every 3 months.

5 Model II: Sector Allocation Model

5.1 Theoretical Basis

• Lagrange interpolation

On $n+1$ different points $x_1, x_2, ..., x_n$ and $f(x_1), f(x_2), ..., f(x_n)$, a polynomial (degree no more than n) $L(x)$ is interpolated at these $n+1$ points, that is,$L(x_i) = f(x_i), i = 0, 1, ..., n$.

Let

$$l_j(x) = \prod_{i=0, i\neq j}^{n} \frac{x - x_i}{x_j - x_i} \tag{30}$$

where $deg(l_i) = n, j = 0, 1, ..., n$. Then

$$l_j(x_k) = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{31}$$

So

$$L(x) = \sum_{j=0}^{n} f(x_j) l_j(x). \tag{32}$$

• Entropy evaluation method

In information theory, entropy is a measure of uncertainty; the greater the amount of information, the smaller the uncertainty and the smaller the entropy. According to the properties of entropy, we can calculate the entropy value to determine the randomness and the degree of disorder of an event. The calculation process is as follows:

(1) Data matrix:

$$A=\begin{pmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & \vdots & \vdots \\ X_{n1} & \cdots & X_{nm} \end{pmatrix}_{n\times m}$$

where x_{ij} is the value of the item j index of the scheme i.

(2) The proportion of the i-th scheme under the item j index is

$$P_{ij}=\frac{X_{ij}}{\sum_{i=1}^{n} X_{ij}} \quad (j=1,2,\cdots m) \tag{33}$$

(3) Calculate the entropy of the item j index:

$$e_j=-k^* \sum_{t=1}^{n} P_j \log\left(P_{ij}\right) \tag{34}$$

where $k>0$, ln is the natural logarithm , $e_j \geq 0$.The constant k is related to the number of samples m, generally let $k=\frac{1}{\ln m}$, then $0 \leq \mathrm{e} \leq 1$.

(4) Calculate the coefficient of difference of the item j index:

$$g_j=1-e_j \tag{35}$$

(5) Number of pleas:

$$W_j=\frac{g_j}{\sum_{j=1}^{m} g_j}, j=1,2\cdots m \tag{36}$$

• Spearman's correlation coefficient

The Spearman correlation coefficient is a non-parametric measure of the dependence of the two variables, and for samples with a sample size of n, n raw data are transformed into hierarchical data with a correlation coefficient of

$$\rho=\frac{\sum_i\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum_i\left(x_i-\bar{x}\right)^2 \sum_i\left(y_i-\bar{y}\right)^2}} \tag{37}$$

When the two variables are completely monotonically correlated, the Spearman correlation coefficient is + 1 or-1. The filtered independent variables were tested for the correlation with the dependent variables for further analysis.

• **Normalization processing**

In order to facilitate the data processing, the data mapping to the range of $[0,1]$ processing, more convenient and fast,making the comparability between the different data more obvious and stronger, we can let

$$x_i^{'} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{38}$$

5.2 Model Building

We have a water shortage in the five states (AZ, CA, WY, NM, CO) of $x_i(1 \leq i \leq 5)$.

First, we interpolates the agricultural, industrial, and residential water data from the existing states by lagrange interpolation to obtain the annual data. Due to agricultural, industrial, and residential water use changes regularly with time, so good data and information should be characterized by low confusion and low information entropy. The above data were detected using the entropy method to verify the credibility of the interpolated data.

Second, the correlations between agricultural, industrial, residential water use, electricity use and GDP were calculated separately for each state using the Spearman correlation coefficient, denoted as a_1,a_2,a_3,a_4. The correlations for each state's agricultural, industrial, residential, and electricity Water shortage is assigned according to the inverse of a_1,a_2,a_3,a_4above, and the correlations are processed using normalization to obtain the assigned weights as $a_1^{'},a_2^{'},a_3^{'},a_4^{'}$,where

$$a_i^{'} = \frac{1/a_i^{'}}{\sum_i 1/a_i^{'}} \tag{39}$$

Finally, we use the weights to calculate the shortage of water in each state as x_i, which corresponds to the shortage of $a_i x_i$ in the four areas of agricultural, industrial, residential water and electricity.

5.3 Model Solutions and Results

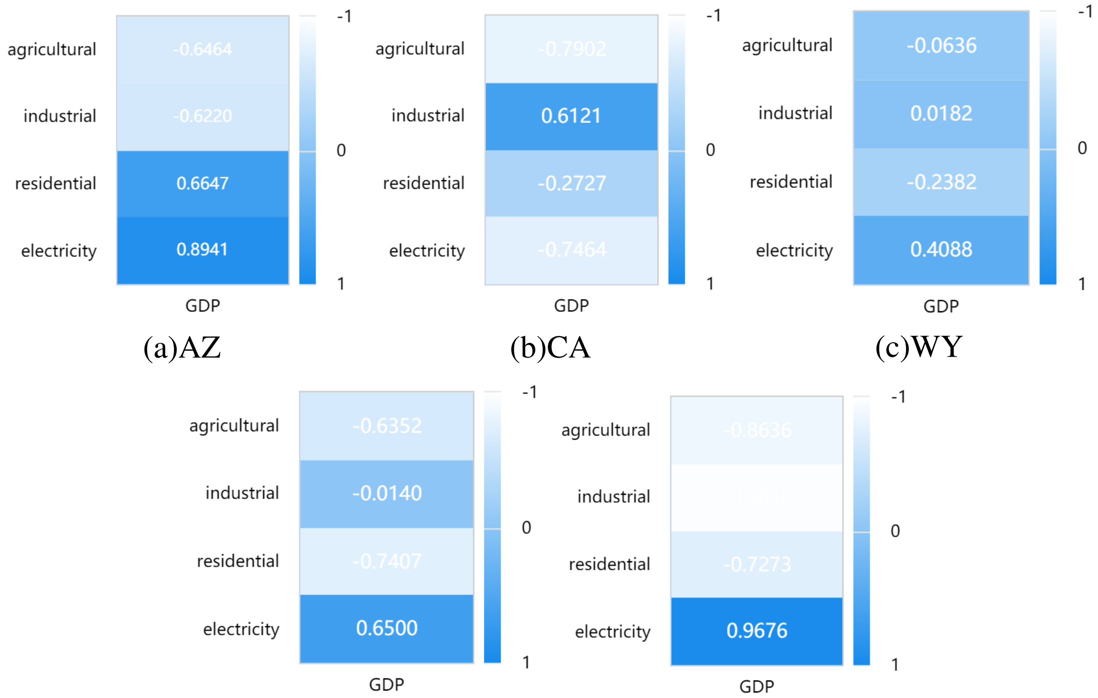
First, information entropy test for data after lagrange interpolation and the results are as follows:

Table 7: Information utility value

	AZ	CA	WY	NM	CO
agriculutre	0.0034	0.0024	0.0304	0.0009	0.0025
industrial	0.0478	0.03244	0.0622	0.0226	0.0047
residential	0.0294	0.0037	0.0003	0.0005	0.0026

As can be seen from the above table, the information utility value is all below 0.05, indicating that the degree of information confusion is low, which indicates that the interpolation effect is good and the data credibility is high, so the following model solution part can use this data.

Secondly, the Spearman correlation coefficient was used to calculate the correlation between agricultural, industrial, residential water and GDP respectively. The results are as follows:



(a)AZ (b)CA (c)WY

(d)NM (e)CO

Figure 11: Correlation calculation results

Finally, the calculation of the water shortage apportion amount of different uses in various states is:

Table 8: distribution outcomes

	AZ	CA	WY	NM	CO
agriculutre	$0.2657x_1$	$0.1594x_2$	$0.2275x_3$	$0.0150x_4$	$0.2525x_5$
industrial	$0.2786x_1$	$0.2064x_2$	$0.6824x_3$	$0.9574x_4$	$0.2262x_5$
residential	$0.2617x_1$	$0.4663x_2$	$0.0567x_3$	$0.0129x_4$	$0.2974x_5$
electricity	$0.1941x_1$	$0.1679x_2$	$0.0333x_3$	$0.0147x_4$	$0.2239x_5$

6 Shortage Apportion Problem

6.1 Theoretical Basis

- **Goodness-of-fit tests**

The goodness of fit is the fit of the regression line to the observations, measured by the dependent coefficient R^2, and the closer the R^2 value is to 1, the better the fit of the regression lines to the observed value.

• **Best cubic function square approximation**

We defined:

(1) variable :$\mathrm{X} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$; $\mathrm{Y} = (\mathrm{y}_1, y_2, \ldots, y_n)^{\mathrm{T}}$

(2) Internal product:$(X,Y) = \sum_{i=1}^{n} w_i x_i y_i$, where w_i are weight coefficient , and $\mathrm{w}_i > 0$.

(3) norm: $\|\mathrm{X}\| = \sqrt{(X,X)}$.

(4) The best approximation function is a polynomial with a number of 3, satisfied

$$\|Y - P_3^*\| = \inf_{P_3^* \in \{\text{all cubic polynomials}\}} \|Y - X\|$$

A polynomial fit, that is, the solution system of equations(take all $\mathrm{w}_i = 1$)

$$\begin{bmatrix} 1 & \ldots & 1 \\ \ldots & \ldots & \ldots \\ x_1^3 & \ldots & x_n^3 \end{bmatrix} \begin{bmatrix} 1 & \ldots & x_1^3 \\ \ldots & \ldots & \ldots \\ 1 & \ldots & x_n^3 \end{bmatrix} \begin{bmatrix} a_0^* \\ \ldots \\ a_3^* \end{bmatrix} = \begin{bmatrix} 1 & \ldots & 1 \\ \ldots & \ldots & \ldots \\ x_1^3 & \ldots & x_n^3 \end{bmatrix} \begin{bmatrix} y_1 \\ \ldots \\ y_n \end{bmatrix}$$

6.2 Problem Solutions and Results

The total water and power consumption of the states were fitted to GDP, and the maximum goodness of fit using the cubic equation was found, so the best cubic approximation was chosen for fit.The corresponding function equation is obtained as:

Table 9: The function fits the results

	GDP (y)
AZ-water(x)	$y = 0.016x^2 - 48.43x + 380290.66$
AZ-electricity(x)	$y = -12.96x^3 + 1253.88x^2 - 32841.23x + 526600.12$
CA-water(x)	$y = -0.002x^2 + 25.45x + 2663528.82$
CA-electricity(x)	$y = 108.32x^3 - 4948.3x^2 + 3046.14x + 3508430.8$
WY-water(x)	$y = -0.43x + 41471.73$
WY-electricity(x)	$y = 13.07x^3 - 956.08x^2 + 22984.64x - 141905.8$
NM-water(x)	$y = -0.03x^2 + 73.58x + 34649.68$
NM-electricity(x)	$y = 19.44x^3 - 1342.22x^2 + 30635.88x - 142355.76$
CO-water(x)	$y = 0.005x^2 - 83.99x + 700434.11$
CO-electricity(x)	$y = 24.62x^3 - 1399.84x^2 + 31158.54x + 12491.29$

By taking the above formula, $\triangle x = 100$, the corresponding $\triangle y$ can be calculated, using $\triangle y$ as the allocation weight, the same as the problem 2 method, using the weights.With the calculation, the total water shortage is L, and the water and electricity corresponding to each state shall be allocated as follows:

Table 10: distribution outcomes

	AZ	CA	WY	NM	CO
water	0.12 L	0.22 L	0.13 L	0.07 L	0.067 L
electricity	0.018 L	0.31 L	0.025 L	0.02 L	0.02 L

7 Sensitivity and Robustness Analysis

- **The Stability of Model I**

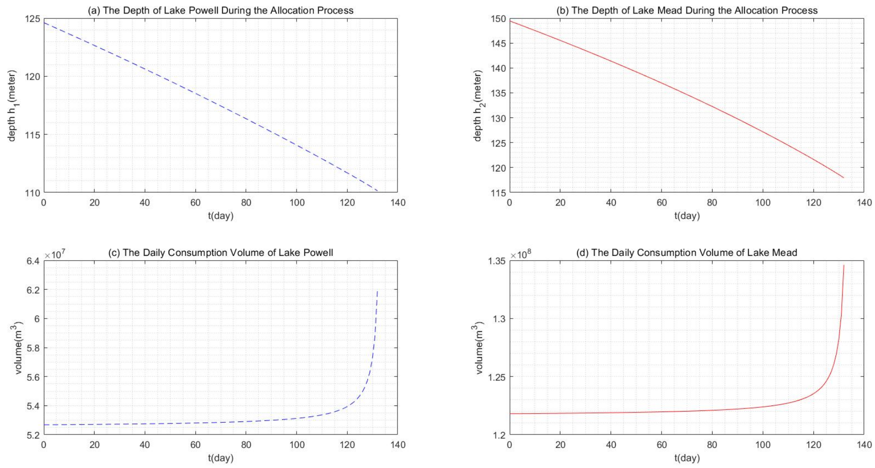


Figure 12: Stability image

When the loss all become five times, consumed four days before, the difference is not obvious.It is not obvious in the image, which also proves the robustness (robustness) of the model and does not vary greatly in the results due to small changes.

The transmission water consumption rate increases, the same arrival quantity dam emission increases (slope increases), both reservoirs are exhausted earlier, the channel consumption rate increases, glen same output, the water quantity to hoover decreases, hoover water consumption accelerates, and Mead consumes earlier.

The e increase, reduced transmission water consumption, both reservoirs are exhausted later; prop increases (less), power demand increases (less), power transmission increases (less), and both reservoirs are depleted earlier (later).

- **Different demands due to population, agricultural, and industrial expansion or contraction.**

When the population is affected, remember the population change a% (including positive and negative), using model two, then the state a_3 becomes $a_3(1+a\%)$, a_4 becomes $a_4(1+a\%)$ to get a new proportion for calculation.According to the formula, when the population increases, the proportion of residential water and electricity will increase relatively, and more water will be proportional than others.The opposite was reversed when the population decreases.

In the case of AZ state, when the population increased by 5%, the distributed water volume of agricultural, industrial, and residential water and electricity are changed by-

7%, -11%, 3.8%, and 25%.

When agriculture is affected, remember the agricultural change b% (including positive and negative), using model two, then the state a_1 becomes $a_1(1+b\%)$, a_4 becomes $a_4(1+b\%)$ to get a new proportion for calculation.The formula can see that when the agricultural economy increases, the proportion of agriculture water and electricity will increase relatively, and the water allocated in proportion is more than others.The opposite conclusion is true when the agricultural economy decreases.

In AZ state, for example, when the agricultural economy increased by 5%, by calculation, the distributed water volume of agricultural, industrial, and residential water and electricity are changed by 4%, -5.4%, -4.8%, and 17%.

When the industry is affected, remember the industrial change c% (including positive and negative), using model two, the state a_2 becomes $a_2(1+c\%)$, a_4 becomes $a_4(1+c\%)$ to get a new proportion for calculation.According to the formula, when the industrial economy increases, the proportion of industrial water and electricity consumption will increase relatively, and the water allocated in proportion is more than the others.

In AZ state, for example, when the industrial economy increased by 5%, the water distribution volumes of agricultural, industrial, and residential water and electricity are changed by-6.8%, -11%, 7.9%, and 19%.

• **Increased share of renewable energy technologies.**

Compared with the initial value of model one, the proportion of renewable energy technology increases by d%, and the hydropower distribution ratio is s: k to s: k (1-d%), and the water used to generate electricity is reduced by d%.According to the formula, when the renewable energy technology increases, the proportion of the water demand provided by the dam will be relatively reduced, and the water distributed in proportion will be less than the electricity consumption.

When the percentage of renewable energy technology increased by 5%, the dam water and electricity use changed by 2%, -2.9%.

• **Other effective water and electricity saving measures.**

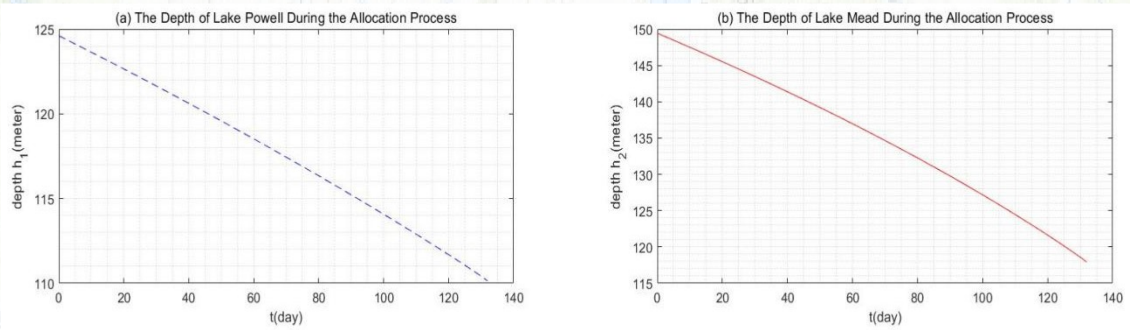
Additional water saving and power saving measures were implemented. With water consumption reduced by e% and f%, the hydropower distribution ratio changed from Model 1 from s: k to s (1-e%): k (1-f%), and the water used for water is reduced by f%.It can be seen from the formula that when the additional water saving and power saving increases, the proportion of water and electricity demand provided by the dam will be relatively reduced, and the reduction proportion is determined by the effect of water saving and power saving measures. The better the effect, the smaller the demand for water provided by the dam.

When the water consumption decreased by 5%, the dam water and electricity consumption changed by-3.1% and 5.6%.

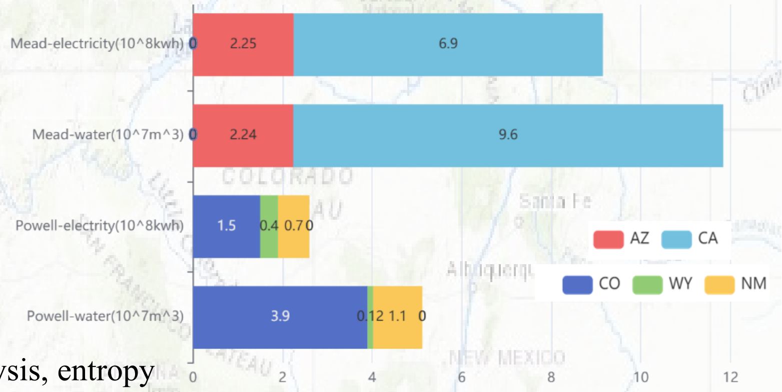
New Insight into Hydro-allocation

Hot weather, dry land, withered plants, inanimate basin…. “The past 22 years have been the driest 22 years since 1200 in the Midwest, particularly in the Colorado Basin”, said Patti Aaron, spokesman of the Bureau of Reclamation. As is known to us, water is an imperative resource not only for it supplies an extensive area for general (agricultural, industrial, residential) usage, but also for it produces hydroelectricity by pushing the turbines and generators when falling and fast flowing. Therefore, a more rational and conservative allocation scheme needs to be formulated urgently.

Consider the case: the Glen Canyon dam (Lake Powell) and the Hoover dam (Lake Mead) provide general- and hydropower-used water to Arizona (AZ), California (CA), Wyoming (WY), New Mexico (NM), and Colorado (CO). Noticeably, Lake Powell sits upstream of Lake Mead, which demonstrates that part of the water released by Glen Canyon dam ends up in the Hoover dam. Assume no additional water is offered to the river system and the five states. There are three possible consequences: Lake Powell cannot supply first; Lake Mead first; both at the same time. Here we briefly introduce the last case. By doing linear programming and solving difference equations, the results are shown in the following figure:



For example, if the initial depth of Lake Powell h_1 is 124.74m and that of Lake Mead h_2 is 149.63m, it will take them 137 days to release all their available volume. As for other initial depths, you could read the result from the figure directly. Besides, the allocation scheme the amount of water and electricity from each lake to each state is displayed on the right:



In addition, if we run into a water shortage, how much water and electricity supply should be deducted in each usage in each state? Through correlation analysis, entropy method, and optimal approximation, we investigate the data of resource demands and GDP by sector and then draw the following conclusion:

How to share the total deduction amount L:

	AZ	CA	WY	NM	CO
water	$0.12L$	$0.22L$	$0.13L$	$0.07L$	$0.067L$
electricity	$0.018L$	$0.31L$	$0.025L$	$0.02L$	$0.02L$

How to share the deduction amount x_i in each state:

	AZ	CA	WY	NM	CO
agriculutre	$0.2657x_1$	$0.1594x_2$	$0.2275x_3$	$0.0150x_4$	$0.2525x_5$
industrial	$0.2786x_1$	$0.2064x_2$	$0.6824x_3$	$0.9574x_4$	$0.2262x_5$
residential	$0.2617x_1$	$0.4663x_2$	$0.0567x_3$	$0.0129x_4$	$0.2974x_5$
electricity	$0.1941x_1$	$0.1679x_2$	$0.0333x_3$	$0.0147x_4$	$0.2239x_5$

9 Model evaluation and promotion

9.1 Strengths

(1)The model offers schemes with different initial values, which makes it applicable to extensive dam supply and allocation problems.
(2)The model takes various factors such as population, climate, economy into consideration. The use of transport loss ratios and energy conversion efficiency makes the model closer to reality.
(3)We use the entropy method to evaluate the interpolation data, which guarantees and enhances the reliability of the data.
(4)With the aim to make the most of water resources, the optimal allocation plan accords with the theme of environmental protection under drought conditions.
(5)The model is made up of rigorous equations, functions, and minor physical and mathematical models, thereby enjoying high interpretability.

9.2 Weaknesses and Further Improvements

(1)If we had more detailed data in each state, such as demands at the county level, rather than the average data of a single state, we could obtain more accurate and detailed results.
(2)It may be improper to simply regard that $Loss = volume(1 + k \cdot distance)$. More complex relations need considering and verifying.

10 References

[1] Zhang Gaofeng, Zheng Hao, Liu Shuangquan, Wang Jinwen. Short-term scheduling algorithm for hydropower cascade systems adapted to different optimization models [J]. Power Grid Technology, 2020.
[2]Zhou Li. Study on Regional Water Resource Optimization Based on Genetic algorithm [D]. Zhengzhou University, 2002.
[3]Xiao Jianhong, Wang Min, Yu Qingdong, Liu Juan. —— of ecological compensation for large-scale hydropower projects based on ecological footprint takes the Three Gorges Project as an example [J]. Journal of Ecology, 2015.
[4]He Chenglong, the energy value conversion rate calculation method of hydropower power generation. Journal of Natural Resources, 2016.
[5]Cheng Haifang, Zhang Zigang. Integrated supplier order model study under multi-supplier condition study [D]. Huazhong University of Science and Technology, 2004.

Appendix

Matlab code

```
dsolve("Dv=(640*10^6)/149.63^2*h*h","h")
dsolve("Dv=(658*10^6)/150.45^2*h*h","h")
gen=[22340818.98,95577105.05,1181275.601,10834984.15,38948101.85];
elec=[224548.1479,685410.0603,42002.78904,67882.61644,153562.3671];
dis=[373.8,664.2,742.6,515.3,646.4];
a=0.1;b=0.1;
function H1 = func1(h, t)
H1 = zeros(3,1);
g = 9.81;
effi=0.95*0.85*0.93;
loss_w=0.001;
loss_e=0.001;
gen=[22340818.98,95577105.05,1181275.601,10834984.15,38948101.85];
elec=[224548.1479,685410.0603,42002.78904,67882.61644,153562.3671];
dis=[373.8,664.2,742.6,515.3,646.4];
rate1=sqrt(640*10^6)/149.63;
H1(1) = h-(1-(elec(1)*(1+loss_e)^dis(1)+elec(2)*(1+loss_e)^dis(2))
/(g*effi*1/3*rate1*h^4))^(1/4);
H1(2)=(gen(1)*(1+loss_w)^dis(1)+gen(2)*(1+loss_w)^dis(2))/rate1/h^2;
H1(3)=H1(1)+H1(2);
end

tspan = [1 20];
h_1 = [0,0,0];
[h,t] = ode45(@func1, tspan, h_1);
plot([0:60],h(35:95)-1,'g--')
xlabel("time"),ylabel("lose"),title("water")


function H2 = func2(h, t)
H2 = zeros(3,1);
g = 9.81;
effi=0.95*0.85*0.93;
loss_w=0.001;
loss_e=0.001;
gen=[22340818.98,95577105.05,1181275.601,10834984.15,38948101.85];
elec=[224548.1479,685410.0603,42002.78904,67882.61644,153562.3671];
dis=[373.8,664.2,742.6,515.3,646.4];
rate2=sqrt(658*10^6)/150.45;
H2(1) = h-(1-(elec(3)*(1+loss_e)^dis(3)+elec(4)*(1+loss_e)^dis(4)+
elec(5)*(1+loss_e)^dis(5))/(g*effi*1/3*rate2*h^4))^(1/4);
H2(2)=(gen(3)*(1+loss_w)^dis(3)+gen(4)*(1+loss_w)^dis(4)+gen(5)*
(1+loss_w)^dis(5))/rate2/h^2;
H2(3)=H2(1)+H2(2);
end
tspan = [1 20];
h_2 = [0,0,0];
[h,t] = ode45(@func2, tspan, h_2);
plot([0:60],h(35:95)-1,'r--')
xlabel("time"),ylabel("lose"),title("water")
```