Tackling the Drought

The drought in the Lake Mead region has had a great impact on the Southwestern United States, which has been experiencing long-term water shortages. Given the circumstances, this report is mainly an investigation of the recent drought in the Lake Mead region.

In order to evaluate the volume of Lake Mead, this survey first identified three factors that affect the volume of Lake Mead: inflow, outflow, and loss. These factors were then used to establish an IOL model. By analyzing the factors that affect the inflow, outflow, and loss, we established the Inflow, Outflow, and Loss Models, respectively. In addition, we used the micro-element method to initiate a model of the water volume of Lake Mead. One other factoid we need to solve the problem is the water level. We also used statistical methods such as unary and multiple linear regression models to estimate the volume of Lake Mead based on its real-time water elevation and area. Ordinary Least Squares (OLS) were applied to measure the coefficients in linear regression models. Then we found that the R^2 of linear regression was relatively close to 1, indicating a strong regression impact.

In order to assess the elevation of Lake Mead, we first visualized the data given in the attachment and summarized the historical water level pattern of Lake Mead. By using Bollinger Bands strategy to define the drought criteria, we found that the recent drought will be more serious and last longer. Time series analysis was applied to create two models for predicting water levels, namely the short-term time series model (STTS) and the long-term time series model (LTTS). STTS has only selected data for the past 3 years, and the predicted result is that the water level of Lake Mead will remain above 1080 feet in the future; LTTS has collected data for nearly 15 years, and its predicted result shows that the water level of Lake Mead will gradually decline.

In order to assess the impact of wastewater recycling, this report first analyzes the five major ways that wastewater is recycled, and then proposes some policies and plans to solve the drought problem. The plan includes the development of wastewater discharge standards, cleaning strategies, efficiency monitoring, and other major parts. we defined a recovery rate (RR), a proportion of recycled water to total water consumption, to quantify the effect of waste water recycling.

Finally, I conducted a sensitivity analysis on the Bollinger Bands strategy, and the results reflect what I expected. Some possible advantages and disadvantages of our model are described.

Keywords: Mead Lake, Linear Regression, Time Series, Bollinger Bands, Recovery Rate

Contents

1 Abstract

Lake Mead
Drought Impact and Possible Response

Our report generally describes the possible causes of a drought around the Southwestern region of America, specifically the Mead Lake area. The report also analyzes the basic factors needed to evaluate the volume, which was mostly used to describe the drought level. Previous data of water levels was used to build two time series prediction models for future water levels. In the response to the drought, a waste water recycling plan has been raised up as well.

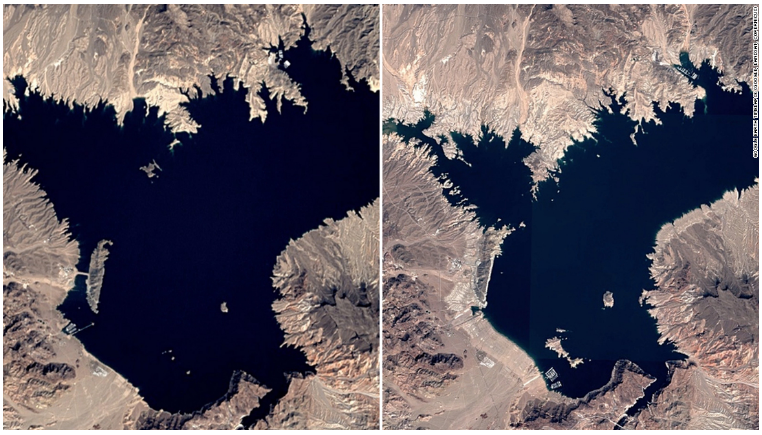


Figure 1.1: Comparison of the water volume of Lake Mead in 1984 and 2021. The picture on the left was taken in 1984, and the picture on the right was taken in 2021[1].

The volume of a lake is mostly affected by inflow, outflow, and loss. The increasing rate of water loss is directly caused by the rise of temperature and inversely proportional to inflow, which results in a smaller volume over a long period of time. The report analyzes the factors that affect the volume of Lake Mead, and carefully analyzes the impact of these factors on the volume and the relationship between them. Taking account of the irregular shape and different depths of Lake Mead, we used the micro-element method to obtain a way to calculate the volume, and used linear regression to verify the previous data. The results they obtained were accurate and robust.

The report also analyzes the past water level of Lake Mead, summarizes several patterns, and uses Bollinger Bands strategy to define drought. It then illustrates two time series prediction models to predict future changes in the water level of Lake Mead. The first model is based on short-term data, and the predicted results are relatively optimistic; The second model is based on long-term data. The predicted results show that the water level of Lake Mead will continue to decline until 2050 meters. The water level of Lake Mead will be lower than 950 feet above sea level. In this case, we need to pay closer attention because of the drought. We also analyzed the impact of wastewater recycling. The report highlights the five major ways that wastewater recycling happens in today's society, and gives practical suggestions to local leaders.

So, everyone, no matter what occupation or age they are, shouldn't argue that there is no need to worry about water. We should strive to make it through the drought and fight for out own future.

2 Introduction

2.1 Background

In today's world, drought has become a new challenge facing mankind. Drought is a long-term dry period in the natural climate cycle. It can occur anywhere in the world. It is a slow-onset disaster characterized by insufficient precipitation, leading to water shortages. Drought can have serious effects on health, agriculture, economy, energy, and the environment. It is estimated that 55 million people worldwide are affected by drought every year, and drought is the most serious hazard to livestock and crops in almost every part of the world. Drought threatens people's livelihoods, increases the risk of disease and death, and encourages mass migration. Water scarcity affects 40% of the world's population, and by 2030, up to 700 million people will be at risk of being displaced by drought[2].

According to the National Oceanic and Atmospheric Association, in the U.S., nearly half the mainland is currently afflicted. The situation is especially dire in the Northwest, which is facing some of its driest conditions in over a century following a heat wave that killed hundreds of people[3]. Also, the Southwest part suffers from the drought. In the summer of 2021, Lake Mead reaches its lowest level on record since its initial filling in the 1930s. On August 16, 2021, the Bureau of Reclamation announced the first-ever water shortage declaration on the Colorado River[4]. Drought—low precipitation leading to water shortage—has persisted for millennia. But scientists say that rising global temperatures and shifting precipitation patterns are leading to more frequent and serious droughts.

In this case, It's important to get well prepared to handle such drought impact and take the reality that it might be a long term situation for us to face drought impact. Among many of the solutions to drought impact, waste water recycling is one of the effective ones. And it's been well researched and conducted by government for years[5].

2.2 Problems

We considers using Lake Mead to investigate and learn more about the impact of drought on the reservoir. We set up wastewater recycling as a solution to water shortages. In this project, We am provided with the volume, water level(elevation), and area of Lake Mead, along with two data sheets containing the time series data of Lake Mead's water levels. The data sheets describe monthly elevation and the lowest and highest water level each year.

Step 1 The volume of water in Lake Mead is a function of inflow, outflow and loss. We first need to identify and describe the factors that affect these variables, and then discuss the relationship between these factors and their impact on the volume of water in Lake Mead and relative influence of water level. In addition, Lake Mead has a particularly irregular shape and varying depths throughout. We need to consider how to verify the relationship between water level, area and water volume.

Step 2 Modeling the water level and elevation of Lake Mead. The data file (online-https://www.comap.com/highschool/contests/himcm/2021problems.html) provides information about the water level. We first need to briefly discuss the overall pattern of historical data on the water level of Lake Mead. Then we need to define the criteria for the drought period and determine the beginning and end of it. Finally, we need to comment on how the recent drought period compares with the earlier drought period.

Step 3 Modeling and planning of wastewater recycling. Based on our model and water level forecast in Problem 2, we need to focus on whether wastewater recycling can make up all or part of the future shortage. First, we need to identify and describe the factors that will be included in the wastewater recycling plan, and consider the decisions that local leaders need to make and the priorities they may set that will affect your plan. Then we need to elaborate our plan in detail and measure the impact of implementing the plan.

2.3 Our work

For Step 1, we first regarded the water volume of Lake Mead as a function of inflow, outflow, and loss, and then analyzed in detail the factors that affect these variables. We then established an IOL model. Secondly, in order to verify the relationship between the water volume of Lake Mead and the water level, altitude, and area, we established a VAE model, using the idea of the infinitesimal method to determine the relationship between them.

For Step 2, we first visualized the attached data and summed up the overall pattern of the historical water level of Lake Mead from these data. Then we used the Bollinger Bands strategy to define the drought criteria, and summarized several drought periods in history, and found that the recent drought was a bit more serious. After that, we established short-term and long-term time series models STTS and LTTS to predict water level elevations in 2025, 2030, and 2050. The prediction effects of the two models were inconsistent, and we discussed them.

For Step 3, we first analyzed the source of waste water to find the factors that affect waste water recycling, and then analyzed the policy measures that government decision makers may take for waste water recycling. We described our plan and evaluated the impact of wastewater recycling on future water resources.

3 Assumptions and Justifications

1. Assume that the water volume of Lake Mead is only caused by inflow, outflow and loss. According to the given information description, these three parts are the main factors affecting the water volume of Lake Mead. Therefore, we can simplify the problem as mentioned above.

2. Assume that the inflow caused by precipitation is a constant value within the time t. Since we often consider the changes in the volume of water in a lake on a scale of one year, although there are rainy and dry seasons in a year, the total annual precipitation is not much different and can be considered as a constant.

3. Assume that the water consumed directly from the lake during the time t is a constant value. Since the water directly consumed in the lake may be taken by nearby residents, the uncertainty is large, so we considered it as a constant here.

4. Assumed that the previous water level data of Lake Mead is accurate, and that the trend of the dry period will continue. From the data chart, we can see that entering the 21st century, the dry period has been continuing, and in the long run, the water volume of Lake Mead has shown a downward trend.

4 Lake Mead Volume Factors and Survey

4.1 Factors Influencing Lake Mead Volume

In this part we need to model the water volume of Lake Mead, which may be affected by many factors, including natural factors and unnatural factors.

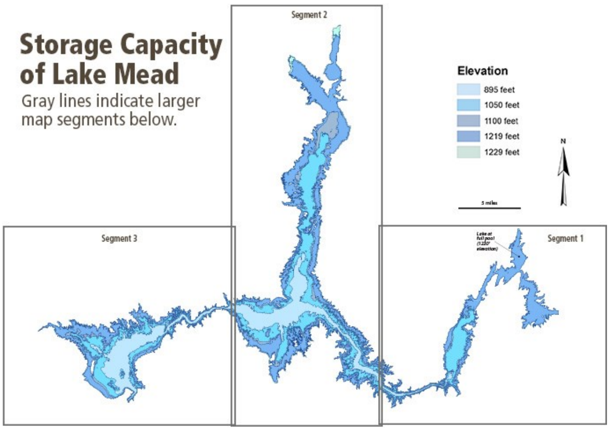


Figure 4.1: Natural Landscape and Water Volume of Lake Mead[6]

For simplicity, we only consider water volume as a function of inflow, outflow, and loss. The Colorado River provides more than 96% of the inflow and additional water from the other three tributaries, as well as direct precipitation from the lake. Outflow occurs through the release of water (for example, through a dam) and directly from the consumption of the lake. Loss occurs through evaporation.

Regarding the water volume of Lake Mead as V. As mentioned in the assumptions, there are three main factors affecting the water volume: inflow, outflow and loss, which satisfy

$$V = V_0 + I - O - L \tag{4.1}$$

where V_0 represents the initial water volume of Lake Mead, I represents inflow, O represents outflow, and L represents loss.

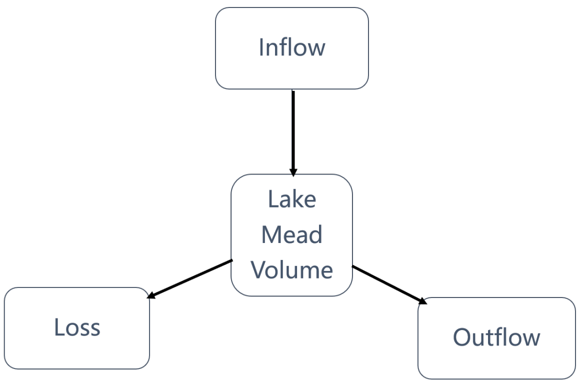


Figure 4.2: Modeling and analysis of Lake Mead's water volume.

Next we model and analyze these aspects separately.

4.1.1 Inflow Model

Now we analyze the factors that affect the inflow. According to the description, the inflow of Lake Mead is mainly composed of two parts: the inflow of other rivers (I_R), and precipitation (I_P), namely

$$I = I_R + I_P \tag{4.2}$$

where I_R represents the inflow of other rivers, and I_P represents precipitation.

Based on the known information, since more than 96% of the inflow flow in through other rivers, we ignore the changes in precipitation as a secondary factor and treat the precipitation as a constant.

Then we analyze the factors that affect the inflow of other rivers, remember

$$I_R = I_C + I_1 + I_2 + I_3 \tag{4.3}$$

where, I_C represents the inflow of the Colorado River, and I_1, I_2, I_3 represent the inflow of the other three tributaries.

The factors that affect the inflow of a river are the cross-sectional area of the river, the average flow rate, and time. The relationship is:

$$I_R = (S_C v_C + S_1 v_1 + S_2 v_2 + S_3 v_3)t \tag{4.4}$$

Among them, S_C, S_1, S_2, S_3 represent the cross-sectional area of the intersection of the Colorado River and the other three tributaries. v_C, v_1, v_2, v_3 represent the average flow velocity of Colorado River and the other three tributaries during the time t at the confluence of Lake Mead.

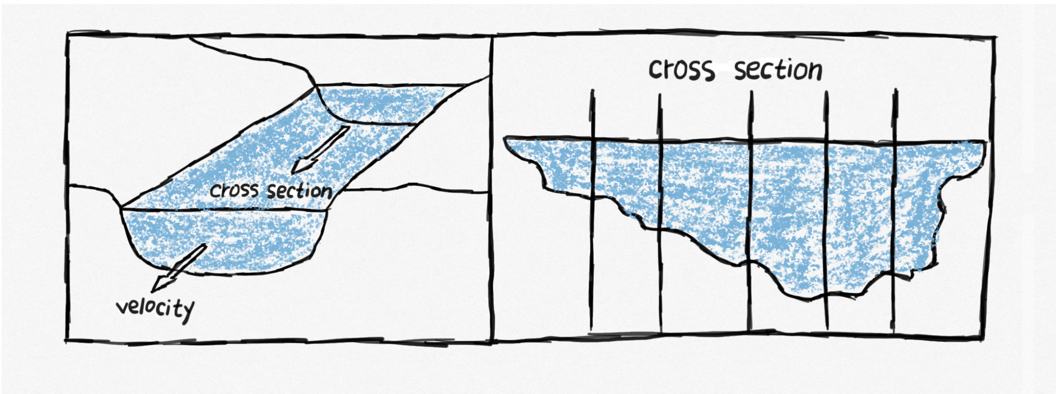


Figure 4.3: The cross-section of the river confluence

Now we consider how to obtain the water flow velocity. The speed of water is an important data in river hydrology. Studies have found that, due to the resistance of the mud at the bottom of the river, the water velocity at different locations on the same cross-section is different.

- The velocity at the center of the river surface is greater than the velocity of the river near the bank.
- The velocity of the river surface is greater than that at the bottom.

- Due to the difference in the river water velocity, the pressure at the center of the river surface is smaller, making the height of the river center higher than the sides.
- As the river water is mixed with sand, there is a continuous accumulation of sand sinking to the bottom of the river as the river flows, which will change the shape of the river bottom and affect its flow in turn.

In the prior art, we can characterize the flow rate of river water by averaging at multiple points.

$$v = \frac{1}{n}\sum_{i=1}^{n} v_i$$

Where $\{v_i\}_{i=1}^{n}$ represents the flow velocity of n points sampled in the river.

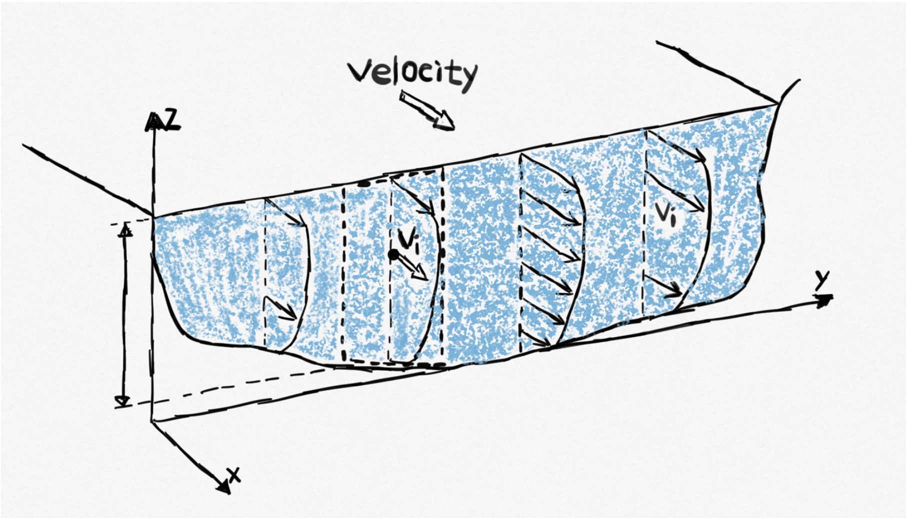


Figure 4.4: Water flow velocity on a cross-section of a river

As for the cross-sectional area of the confluence of the river and Lake Mead, we have the following steps[7].

1. Establish an $x - y$ coordinate system with x representing the width direction and y representing the height direction f the river channel;
2. Measure the contour line of the water surface($f_{up}(x)$) at the measurement site;
3. Measure the river bottom contour line ($f_{down}(x)$) at the measurement site;
4. Depict $f_{up}(x)$ and $f_{down}(x)$ on the $x - y$ coordinate system. The resulting closed area is exactly the cross-section of the measurement site;
5. Calculate the area of the closed area as the cross-sectional area through the formula 4.5.

$$S = \int_{x_1}^{x_2} f_{up}(x) - f_{down}(x)dx \tag{4.5}$$

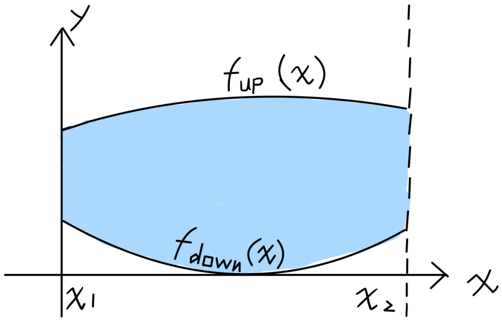


Figure 4.5: Cross-sectional Area Calculation.

The rate of flow v here is given based on the mean of the velocity of each point in the field experiment. In fact, we can also get the flow through model test or simulation, which is more accurate than calculating the mean directly.

For the simulation method, a 21mm wide and 1000mm long river model is first established in CFD commercial software FLUENT for illustrating the principle instead of the real size. The river bottom contour is a spline curve. For a formal survey, the depth at each point across the river can be measured to fit its bottom contour.

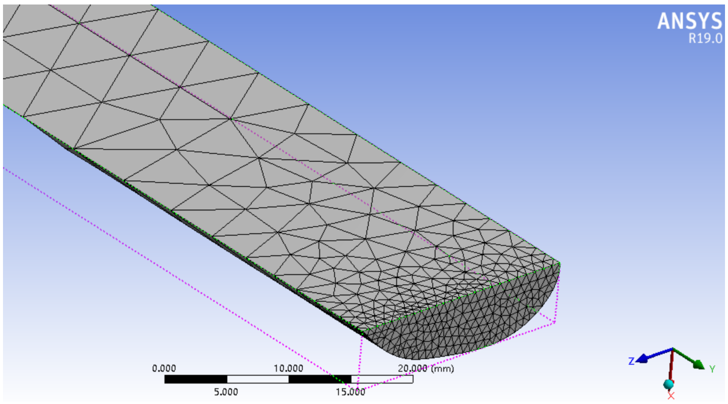


Figure 4.6: Meshing

Due to the viscosity of the bottom wall, the velocity at the bottom of the river is 0. As shown in Figure 4.6, the change of flow velocity along the river direction is not as sharp as that on the cross-section, so when conducting finite element simulation, a more compact mesh should be set on the cross-section, which is partitioned by patch generation method. Certainly, we can use some fluid mechanics knowledge, such as boundary layer theory and potential flow theory, to calculate the theoretical solution of the flow.

In fact, we build a scaled model through experimental mapping, and then set the appropriate pressure differential to simulate the driving force of the river. Then the flow Q_{exp} of this experiment is measured by the flow meter. In order to reduce the pressure difference loss caused by the flow meter, ultrasonic flow meter is selected. Then, according to the similarity principle, the flow Q_{fact}will be obtained.

First, we select four independent variables Δ, A, Q, ρ, whose basic dimensions are L, M, T.

$$\begin{cases} [\Delta P] = MT^{-2}L^{-1} \\ [A] = L^2 \\ [Q] = MT^{-1} \\ [\rho] = ML^{-3} \end{cases}$$

In order to obtain infinite dimensional constants $\Pi = \Delta P^a A^b Q^c \rho$, we set out the equations (4.6).

$$\begin{cases} a + c + 1 = 0 \\ -2a - c = 0 \\ -a + 2b - 3 = 0 \end{cases} \tag{4.6}$$

Solve the system of equations. The final solution is $\Pi = \Delta P A^2 Q^{-2} \rho$. As the density of experimental water flow is the same as that of the real one, the flow rate of the model can be calculated by controlling the pressure difference and area difference.

$$\frac{Q_{fact}}{Q_{exp}} = \sqrt{\frac{\Delta P_{fact}}{\Delta P_{exp}} \frac{A_{fact}}{A_{exp}}}$$

4.1.2 Outflow Model

Now we analyze the factors that affect the outflow. Given the information, we know the outflow is combined by released water (such as a dam) and direct consumption (consumption), so

$$O = O_R + O_C \tag{4.7}$$

Among them, O_R represents the water released from the lake, and O_C represents the water directly consumed in the lake. Since it is difficult to estimate the water consumed directly from the lake, we believe that its annual change is not large, so it is regarded as a constant.

Next, consider the water released from the lake O_R, which is mainly similar to the release of water for irrigation in a dam manner. Similar to the above analysis:

$$O_R = S_R v_R t_R \tag{4.8}$$

where, S_R represents the cross-sectional area of the pipe where the water is released (usually it can be directly measured), and t_R represents the time the water is released.

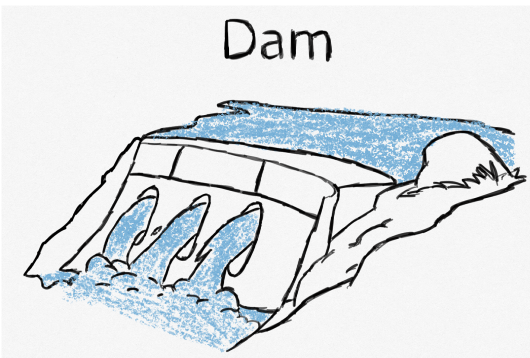


Figure 4.7: Water Releasing through a dam.

4.1.3 Loss Model

Now we consider the amount of water evaporated in Lake Mead L, which depends on the surface area of the lake A and the evaporation rate E. In fact, the evaporation rate E is influenced by many factors, including dimensions, temperature, etc.

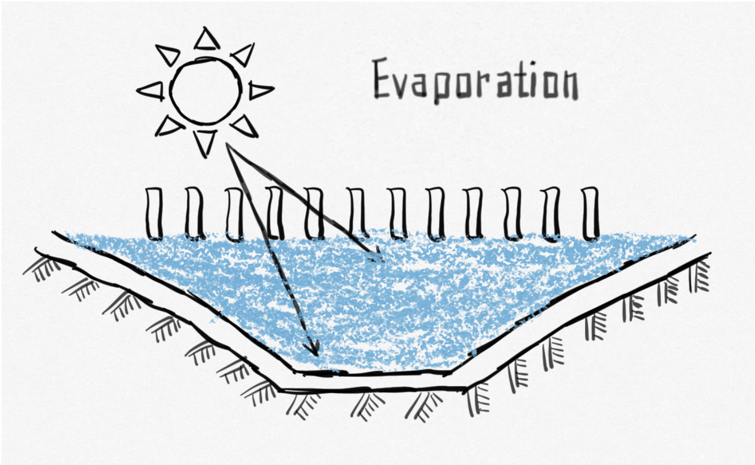


Figure 4.8: Lake Evaporation.

The Penman equation[8] describes evaporation E from an open water surface, which requires daily mean temperature, wind speed, air pressure, and solar radiation to predict E.

Numerous variations of the Penman equation are used to estimate evaporation from water, and land. Specifically the Penman-Monteith equation refines weather based potential evapotranspiration (PET) estimates of vegetated land areas. It is widely regarded as one of the most accurate models, in terms of estimates.

The equation for evaporation given by Penman is:

$$E_{\text{mass}} = \frac{mR_n + \rho_a c_p(\delta e) g_a}{\lambda_v(m+\gamma)} \tag{4.9}$$

where:

$m =$ Slope of the saturation vapor pressure curve $\left(\mathrm{PaK}^{-1}\right)$

$R_{\mathrm{n}} =$ Net irradiance $\left(\mathrm{Wm}^{-2}\right)$

$\rho_{\mathrm{a}} =$ density of air $\left(\mathrm{kgm}^{-3}\right)$

$c_{\mathrm{p}} =$ heat capacity of air $\left(\mathrm{Jkg}^{-1}\,\mathrm{K}^{-1}\right)$

$\delta e =$ vapor pressure deficit (Pa)

$g_{\mathrm{a}} =$ momentum surface aerodynamic conductance $\left(\mathrm{ms}^{-1}\right)$

$\lambda_{\mathrm{v}} =$ latent heat of vaporization $\left(\mathrm{Jkg}^{-1}\right)$

$\gamma =$ psychrometric constant $\left(\mathrm{PaK}^{-1}\right)$

which (if the SI units in parentheses are used) will give the evaporation E_{mass} in units of $\mathrm{kg}/\left(\mathrm{m}^2 \cdot \mathrm{s}\right)$, kilograms of water evaporated every second for each square meter of area.

This model considers more comprehensive factors, but it is also relatively complicated. In fact, as an estimate of the evaporation rate, there is a simpler model.

The Penman formula for the evaporation rate from a lake is simplified as[9]:

$$E_0 = \frac{700T_{\mathrm{m}}/(100 - L_a) + 15\,(T - T_{\mathrm{d}})}{(80 - T)}\,\left(\mathrm{mm\ day}^{-1}\right) \tag{4.10}$$

where $T_{\mathrm{m}} = T + 0.006\,\mathrm{h}$, h is the elevation (metres), T is the mean temperature, L_a is the latitude (degrees) and T_{d} is the mean dew-point. Values given by this formula typically differ from measured values by about 0.3 mm day $^{-1}$ for annual means, 0.5 mm day $^{-1}$ for monthly means, 0.9 mm day $^{-1}$ for a week and 1.7 mm day $^{-1}$ for a day. The formula applies to a wide range of climates. Monthly mean values of the term $(T - T_{\mathrm{d}})$ can be obtained either from an empirical table or from the following empirical relationship, provided precipitation is at least 5 mm month $^{-1}$ and $(T - T_{\mathrm{d}})$ is at least $4°\mathrm{C}$:

$$(T - T_{\mathrm{d}}) = 0.0023h + 0.37T + 0.53R + 0.35R_{\mathrm{ann}} - 10.9°\mathrm{C} \tag{4.11}$$

where R is the mean daily range of temperature and R_{ann} is the difference between the mean temperatures of the hottest and coldest months. Thus, the evaporation rate can be estimated simply from values for the elevation, latitude, and daily maximum and minimum temperatures.

Through formulas (4.10) and (4.11), we can estimate the evaporation rate E_{Mead} from the altitude, latitude and daily maximum and minimum temperature of Lake Mead to measure the area of Lake Mead A_{Mead}, and then through equation 4.12 to get the amount of evaporated water in the time period t.

$$L = A_{Mead}E_{Mead}t \tag{4.12}$$

4.2 Lake Mead Volume Survey

Now we consider the relationship between the water volume of Lake Mead V, the area A, and the water level E. If Lake Mead is a regular cuboid, the height from the bottom of the lake to the sea level is recorded as H_0. Thus, the water volume V can be calculated as

$$V = A \cdot H = A \cdot (E - H_0) = A \cdot E - A \cdot H_0 \tag{4.13}$$

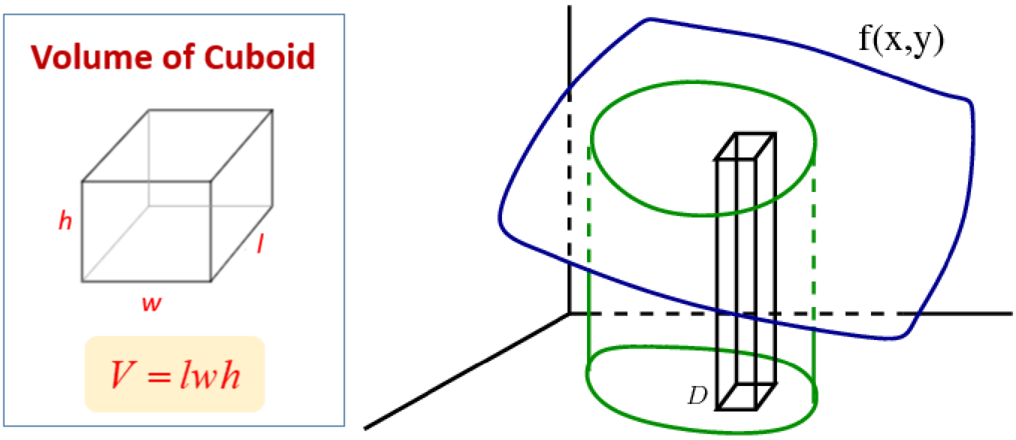


Figure 4.9: Left: The volume of the cuboid[10]. Right: how to calculate the volume of the lake by the infinitesimal method[12].

However, Lake Mead's shape is very irregular, and its depth varies at different sites. Therefore, equation (4.13) can not accurately estimate the water volume V. we need to modify (4.13).

We can use the idea of calculus to get a new expression of volume V. Find points on the lake surface with a boat, measure the depth of the lake, and obtain a topographic map of the lake bottom.

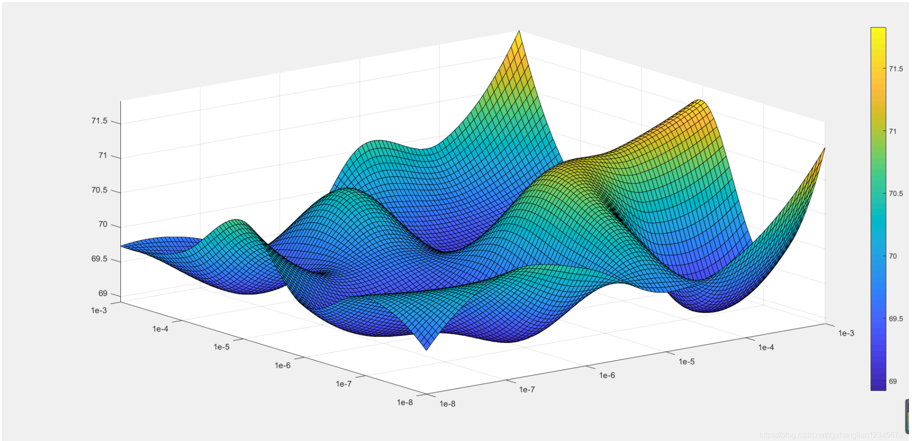


Figure 4.10: Topography of the lake bottom[11].

Since the bottom of the lake generally changes continuously, we can divide the entire lake into N slender cuboids, use (4.13) for each cuboid to calculate the amount of water, and then add them up:

$$V = \sum_{i=1}^{N} A_i \cdot H_i = \sum_{i=1}^{N} A_i \cdot (E - \hat{H}_i) = A \cdot E - \sum_{i=1}^{N} A_i \cdot \hat{H}_i \tag{4.14}$$

Among them, H_i is the height of the cuboid i, which can be obtained by sampling and measuring on the lake surface. $\hat{H}_i = E - H_i$ is the i The distance from the bottom of a small lake cuboid to the height of sea level. This method is the same as the idea of binary integral.

In fact, the information and data we need can be measured by hydrological methods. In 1935, BOR and the Soil Protection Bureau conducted a mapping project[13] to calculate the water storage capacity of the almost formed Lake Mead. As the water behind the dam began to flow backwards, surveyors recorded measurements on the ground based on stereo aerial photos and water level detection (Brown 1941).

Then we consider a data-driven approach to analyze the relationship between volume, elevation, and area from another perspective. We visualize the data given and find that the points are roughly distributed in a straight line. Next, we use linear regression model to fit them.

We assume that the water volume V and water level E of Lake Mead and the lake area A satisfy:

$$V = \beta_1 \cdot A \cdot E + \beta_0 \tag{4.15}$$

Our goal is to find a combination of (β_1, β_0) that minimize the following expression:

$$\min_{\beta_1, \beta_0} \sum_{i=1}^{4} (V_i - \beta_1 \cdot A_i \cdot E_i - \beta_0)^2$$

In fact, there is an analytical solution to this problem:

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{i=1}^{4}\left(A_i \cdot E_i \cdot V_i - \overline{A \cdot E} \cdot \bar{V}\right)}{\sum_{i=1}^{4}\left((A_i \cdot E_i)^2 - \overline{A \cdot E}^2\right)} \\ \hat{\beta}_0 &= \overline{V_i} - a\overline{A \cdot E}\end{aligned} \tag{4.16}$$

where

$$\overline{V} = \frac{1}{4}\sum_{i=1}^{4} V_i$$

$$\overline{A \cdot E} = \frac{1}{4}\sum_{i=1}^{4} A_i \cdot E_i$$

So we get $a = 0.16, b = -1960044.94$. Figure 4.10 shows the results. There is a significant linear relationship between volume V and product $A \cdot E$

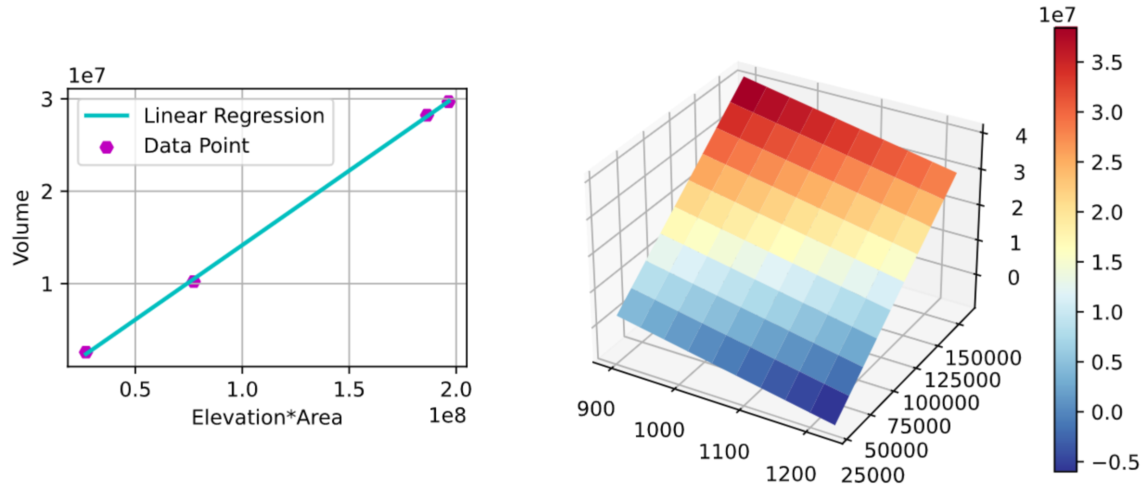


Figure 4.11: Left: Linear Regression $Volume \sim Elevation \cdot Area$. Right: The Multiple Linear Regression.

On the other hand, we can build a binary linear regression model with water volume V as the dependent variable, water level E and lake area A as independent variables:

$$V = \beta_0 + \beta_1 \cdot A + \beta_2 \cdot E \tag{4.17}$$

And the simulation results are in the 3D map Figure 4.11.

5 Water Level Patterns and Prediction

Actually, the prediction of water level has been conducted for many times. They could use only water level time series, or the can use snow line or some other variables to help predict the water level. In this case, we will apply the monthly recorded water level data from 1939 to now to do the prediction. Two result would be generated.

5.1 Overall Patterns of Lake Mead Water Levels

Based on the data given, we plot the monthly Mead Lake's water level and highest/lowest water elevation in Figure 5.1.

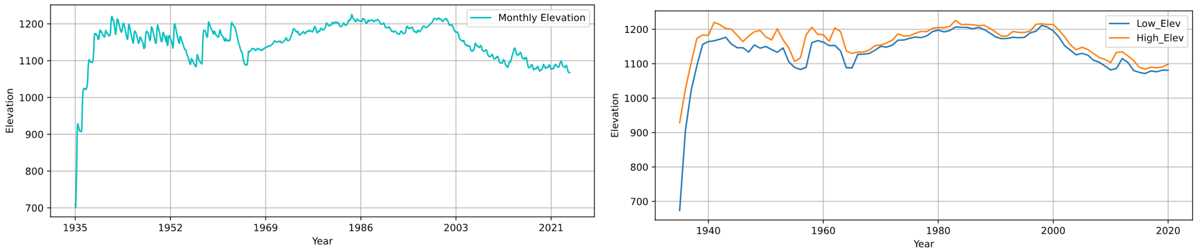


Figure 5.1: Left: Monthly Water Elevation. Right: The Highest and Lowest Water Elevation.

Considering the first few years of the reservoir conducted, the water level is much less than the average. To amplify the variation of water elevation after 1939, we exclude the data before 1939 and get left Figure 5.2, which has already been subtracted by the monthly average water level. Also we draw right Figure 5.2 for the first order difference of water level to see the change of water level among years.

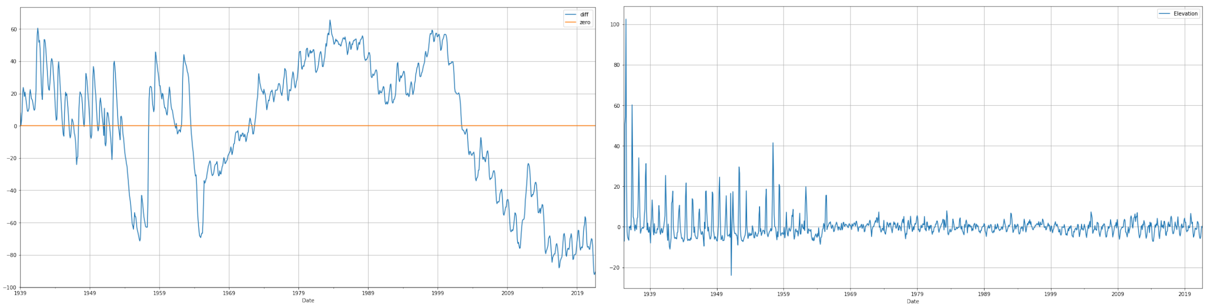


Figure 5.2: Left: Water Level After 1939. Right: the Water Level After Difference.

5.1.1 Overall Patterns

1. Since 1935, the water level of Lake Mead began to rise sharply to 1150 feet.
2. After the water level became stable, it fluctuated between 1100 feet and 1200 feet.
3. The water level of Lake Mead fluctuates seasonally.
4. After year 2000, the water level of Lake Mead declined in a fluctuation.
5. The annual maximum water level and minimum water level of Lake Mead fluctuate synchronously.

5.1.2 Criteria for Drought Periods

We come up with *threshold value strategy* and *Bollinger Bands strategy* for the drought periods prediction, which focus on the absolute value and the gradient of the water level respectively.

- **Criterion 1: Threshold Value 0**

  Based on left Figure 5.2, we have calculated the monthly mean of water level, and plot the water level minus average as in right Figure 5.2. In this case, the yellow line on 0 could be our threshold value to justify if the reservoir is running through a drought period. when water level is below average water level this period, Lake Mead is in the middle of a relatively dry period.

  - A drought period started/ended when (water level - average) broke through line 0 downwards/upwards and failed to recover back to average level soon enough.
  - A flood period started/ended when (water level - average) broke through line 0 upwards/downwards and failed to recover back to average level soon enough.

  Then we identify three periods as drought period in Table 5.1.

  Table 5.1: Drought Period

Period	1	2	3
Start Date	1953-02-01	1963-08-01	2002-07-01
End Date	1957-06-01	1972-08-01	Present

  This criteria is relatively simple and focuses on macro data

- **Criterion 2: Bollinger Band**

  Besides, we considered that drought does not completely depend on the absolute level of water level. According to formula 4.3, the decrease of water level is caused by the decrease of inflow and the increase of outflow, which can be identified as drought. When we extract dry periods according this idea, we should filter out micro changes in water levels, which may cause the dry period to be too fragmented and lose continuity. In order to solve this problem, we propose Bollinger bands strategy.

  Bollinger band is a term in stock trading, which is composed of an upper (pressure) line and a lower (support) line for price. We can calculate the upper and lower line by

  $$MA \pm \beta \cdot std$$

  where MA is moving average and std is standard error, . In this section, we set the rolling window of the Bollinger Band as 4 years, and the standard deviation factor $\beta = 1$. The results are shown in Figure 5.3.

  According to the Pauta Criterion, if random variable $X \sim N(\mu, \sigma^2)$, then

  $$P(X \in (\mu - 3\sigma, \mu + 3\sigma)) \approx 0.9973.$$

  For $X \in (\mu - \sigma, \mu + \sigma)$, the probability becomes 0.6826. So we claim that

  - A drought period started/ended when $MovingAverage(MA)$ broke through the lower Bollinger Bands downwards/upwards with a high slope.

- A flood period started/ended when $MovingAverage(MA)$ broke through the upper Bollinger Bands upwards/downwards with a high slope.

As shown in Figure 5.3, the red shaded area are the drought periods. This method is more sensitive than Criterion 1, and some short term fluctuations below lower Bollinger Band are also taken into account.

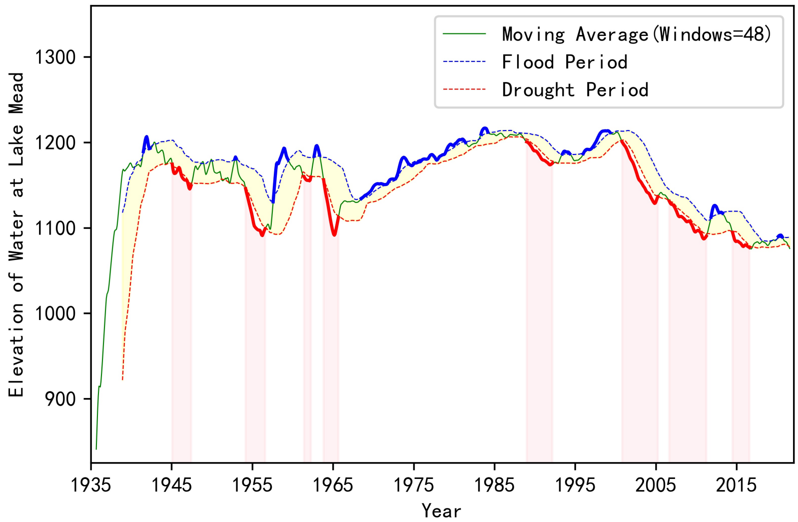


Figure 5.3: Bollinger Bands to Identify Drought/Flood Periods

Based on the results above, we can find that the most recent drought periods are longer more frequent than earlier ones.

5.2 Water Level Prediction

Since the data we collected is chronological and auto-correlated, we can study these time series, discover its development rules, and predict its future trend through the time series model.

5.2.1 Short Term Time Series (STTS) Model

Consider data from only the most recent drought period and assume the most recent drought period's pattern continues, we can build Short Term Time Series (STTS) Model. First, we only take the data in recent three years into account, and draw the water level as follows.

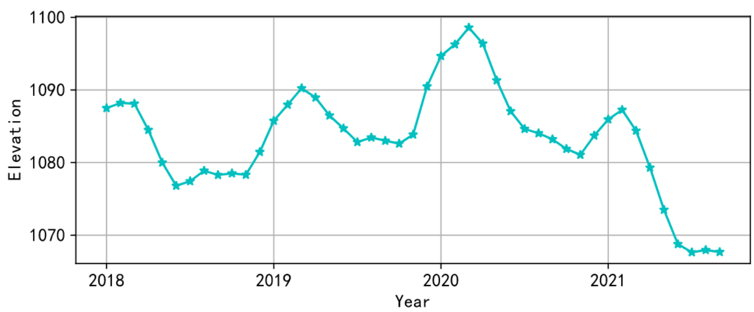


Figure 5.4: Water levels in recent 3 years

From Figure 5.4, we can see that the water level in recent three years is lower than 1100 feet, which is in the drought period. It fluctuated between 1070 feet and 1100 feet, with an average of 1080 feet. If we assume that this trend will continue, it can be predicted that the water level in the next few decades will also remain near 1080 feet, which seems unreliable. We will use the long-term data from 2005 to 2020 later in the following LTTS model for verification.

Now we adopt stationary time series model to predict the water level. *m - step prediction* of time series is to estimate $X_{k+m}(m > 0)$ based on $\{X_k, X_{k-1}, \cdots\}$, and the estimation is a linear combination of $\{X_k, X_{k-1}, \cdots\}$.

Suppose stationary time series $\{X_t, t = 0, \pm 1, \pm 2, \cdots\}$ with expectation $E(X_t) = \mu$, satisfies the following ARMA model:

$$(X_t - \mu) - \phi_1 (X_{t-1} - \mu) - \cdots - \phi_p (X_{t-p} - \mu) = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-i} \quad (5.1)$$

where ε_t is a white noise with zero mean. Now we use (5.1) to predict water level in the following steps.

1. Draw a line graph of the original data (see Figure 5.4), and identify whether it is a stationary time series through ADF-test.

2. Calculate auto-correlation Function (ACF) and partial auto-correlation function (PACF) as Figure 5.5.

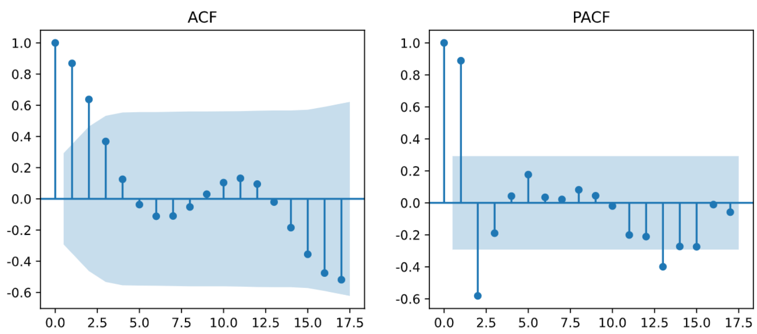


Figure 5.5: ACF and PACF

3. Based on AIC/BIC criteria and ACF/PACF, we choose ARMA(2,4) as our prediction model. The calculation results are shown in Figure 5.6.

ARMA Model Results

Dep. Variable:	y	No. Observations:	45
Model:	ARMA(2, 4)	Log Likelihood	-89.786
Method:	css-mle	S.D. of innovations	1.700
Date:	Sun, 14 Nov 2021	AIC	195.573
Time:	16:34:01	BIC	210.026
Sample:	0	HQIC	200.961

	coef	std err	z	P>|z|	[0.025	0.975]
const	1082.6772	2.773	390.388	0.000	1077.242	1088.113
ar.L1.y	1.4516	0.547	2.655	0.008	0.380	2.523
ar.L2.y	-0.5735	0.430	-1.333	0.182	-1.416	0.269
ma.L1.y	0.4209	0.568	0.741	0.459	-0.693	1.535
ma.L2.y	0.0538	0.617	0.087	0.931	-1.155	1.262
ma.L3.y	0.0287	0.451	0.064	0.949	-0.855	0.913
ma.L4.y	-0.1530	0.279	-0.548	0.583	-0.700	0.394

Figure 5.6: ARMA Model Results

4. Check whether the residual is close to the normal distribution, for which we can adopt Ljung-Box test.

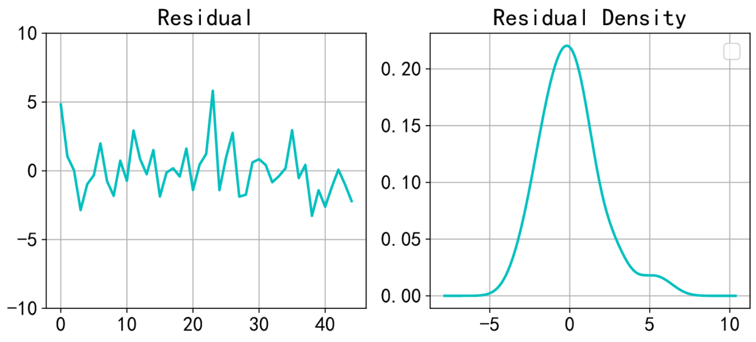


Figure 5.7: Residuals' Values and Distribution

5. Use the model above to predict. The original data and its predicted values are shown in the Figure 5.8.

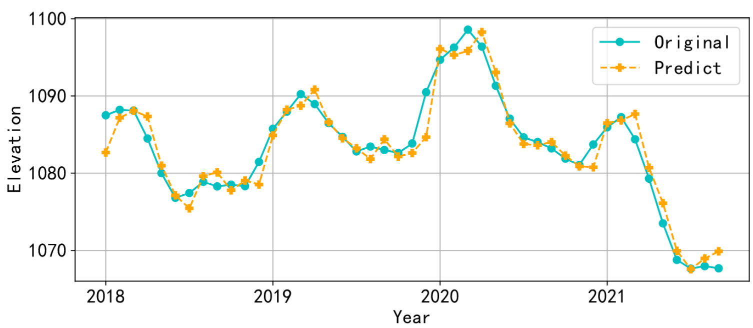


Figure 5.8: Short Term Prediction

Our STTS Model predicts that the annual average water level in 2025, 2030 and 2050 will be 1082.40, 1082.37, 1082.31 respectively, basically maintained near 1080, which meets our earlier expectation.

5.2.2 Long Term Time Series (LTTS) Model

In contrast to the short term model above, we build Long Term Time Series (LTTS) Model in this section with water level data from 2005 to 2021 in the following steps.

1. Draw a line graph of the original data in Figure 5.9. We find this time seires isn't stationary, with a significant descending trend.

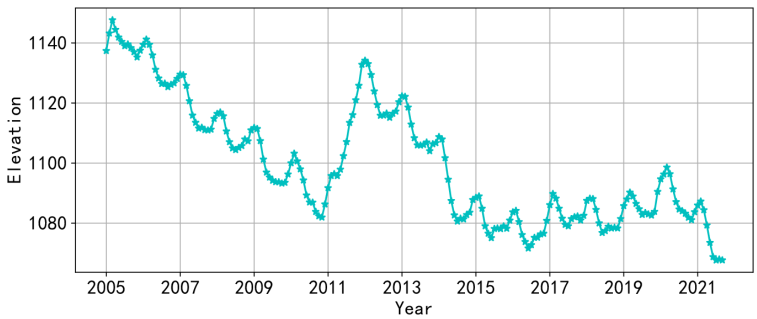


Figure 5.9: Water Levels from 2005 to 2021

2. Since the data shows a linear trend, we decide to do first order difference on the data and build ARIMA(p,q) model.

3. Based on AIC/BIC criteria and ACF/PACF, we choose ARIMA(2,1,0) as our prediction model. The calculation results are shown in Figure 5.10.

ARIMA Model Results

Dep. Variable:	D.y	No. Observations:	200
Model:	ARIMA(3, 1, 2)	Log Likelihood	-404.184
Method:	css-mle	S.D. of innovations	1.787
Date:	Sun, 14 Nov 2021	AIC	822.369
Time:	17:13:43	BIC	845.457
Sample:	1	HQIC	831.712

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.3122	0.410	-0.761	0.447	-1.116	0.492
ar.L1.D.y	1.6626	0.056	29.823	0.000	1.553	1.772
ar.L2.D.y	-1.6544	0.057	-28.879	0.000	-1.767	-1.542
ar.L3.D.y	0.6608	0.055	11.987	0.000	0.553	0.769
ma.L1.D.y	-0.9150	0.024	-38.539	0.000	-0.962	-0.868
ma.L2.D.y	0.9998	0.031	32.632	0.000	0.940	1.060

Figure 5.10: ARIMA Model Results

4. Use the model above to predict. The original data and its predicted values are shown in the Figure 5.11.

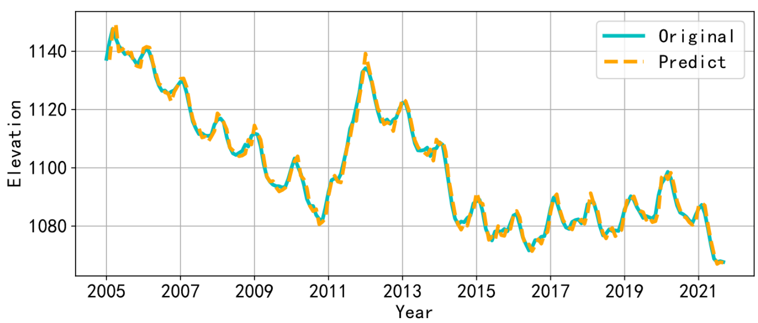


Figure 5.11: Long Term Prediction

Our LTTS Model predicts that the annual average water level in 2025, 2030 and 2050 will be 1050.84, 1029.09, 942.10 respectively, which shows a significant descending trend and meets our earlier expectation. Compared to STTS Model, LTTS Model has a more accurate prediction because more history data are used.

6 Future Water Usage and Wastewater Recycling

Based on our models and water level predictions in Section 6, the water level of Lake Mead will be less than 1000 feet in the long run. Thus, it is quite important to recycle wastewater.

6.1 Factors in Recycling Wastewater

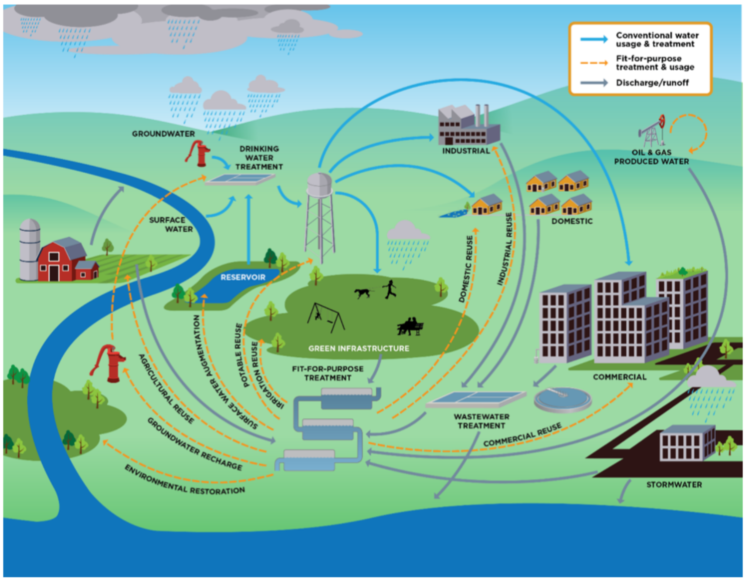


Figure 6.1: Wastewater Recycle

From Figure 6.1, we can see that the grey lines are ways of wastewater generation. There are five in total:

$$WR = WR_I + WR_D + WR_C + WR_S + WR_O \tag{6.1}$$

where

WR = Wastewater Recycle
WR_I = Industrial Wastewater Recycle
WR_D = Domestic Wastewater Recycle
WR_C = Commercial Wastewater Recycle
WR_S = Surface Wastewater Recycle such as Rainstorm
WR_O = Oil and Other Wastewater Recycle

For wastewater recycle, we need the help of local leaders to make the following decisions:

1. Enterprises are required to comply with regulatory requirements such as adopt wastewater treatment facilities to meet strict pretreatment or direct discharge standards.

2. Reduce or eliminate wastewater treatment surcharges. Industrial discharged wastewater shall be treated on site as far as possible, so as to reduce (or even eliminate) the surcharges related to the discharge of untreated wastewater to the local municipal treatment plant.

3. Reuse water. For the treated water that can be reused, the factory will recover high-quality effluent to reduce the operation cost of the factory.

4. Convert wastewater into green energy. Valuable biogas can be produced from wastewater through anaerobic digestion, which has become a cost-saving renewable energy.

6.2 Measurement of Recycling plan's Impact

The Sustainable Development Goals[14] (SDGs) are the latest attempt by the international community to mobilize government, private and non-governmental actors at the national, regional and local levels to improve the quality of life of billions of people in developed and developing countries. These goals are the ambitious, challenging and urgently needed "Humanity, Planet and Prosperity" action plan by 2030.

Among the 17 Sustainable Development Goals, the sixth goal is to "ensure that all people have access to water and sanitation facilities and sustainably manage them". In view of the importance of clean water to overall socio-economic development and quality of life (including health and environmental protection), even a partial realization of this goal will greatly benefit mankind.

We are able to quantify the impact of our plan using wastewater recycling as a percentage of water demand.

$$RR = \frac{WR}{WD}$$

where RR is the evaluation index of wastewater recycling, representing the percentage of waste water recycling to the water supply demand, WR is the reclaimed water obtained from waste water recycling, and WD is the total water supply demand.

If our plan can be implemented efficiently, it can be foreseen that wastewater recycling can effectively alleviate the water demand caused by drought. For example, in California, most of the state has a rapidly increasing population and a desert climate, so efficient water use is essential. Faced with such demand and limited supply, water recycling is becoming more and more common throughout California. Reclaimed water in parts of the state has already met approximately 7% of the water supply demand.

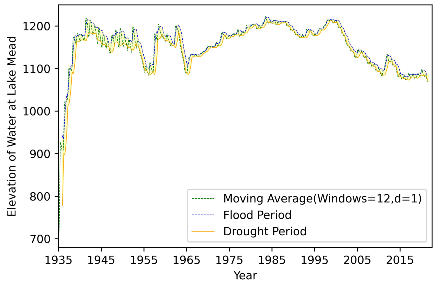
7 Sensitivity Analysis

Figure 5.3 uses two hyperparameters: Windows(moving average length) and d(magnification factor). The result of Figure 5.3 has been relatively ideal after adjustment. In order to conduct sensitivity analysis on this part, we first keep the amplification factor d = 1 to unchanged and set the moving average length to 12, 24, 48 and 96 respectively, so as to make Figure 7.1.

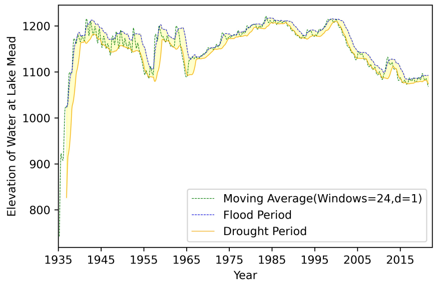
Figure 7.1 shows that when the window length is too small, such as figure 7.1a, the Bollinger band will follow the moving average line of water level closely, and the moving average line will frequently cross the upper and lower limits. When the window length is too large, such as Figure 7.1d, in which the Windows length is 96, the period of the fluctuation is too long and too macro, so that some relatively micro rise and drop are easy to be ignored and cannot be captured.

Otherwise,when we set moving average length to 48 and magnification factor D to 0.5, 1, 2, 3 respectively, figure 7.2 was obtained. We consider that d = 1 is relatively appropriate. When d is too large, it is difficult for the moving average line to break through the Bollinger Bands. when d is too small, the moving average line will have frequent contact with the upper and lower boundaries.

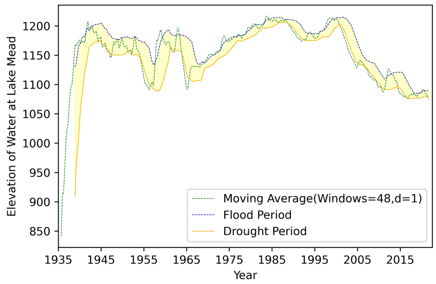
To sum up, Windows=48 and D =1 are relatively reasonable values of hyperparameters.



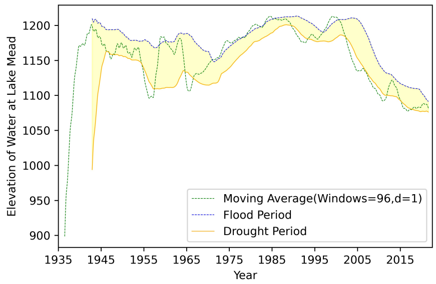
(a)



(b)

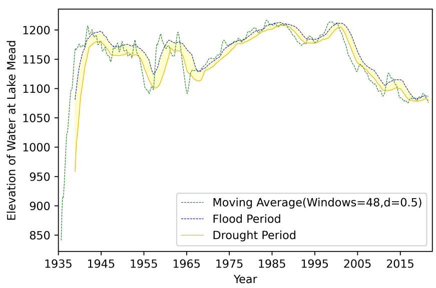


(c)

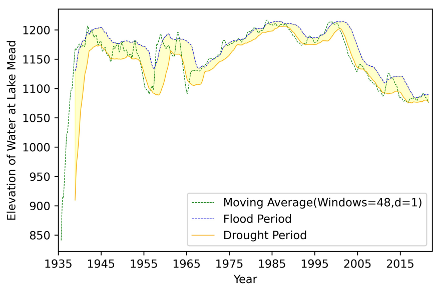


(d)

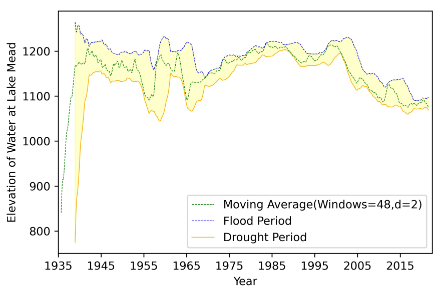
Figure 7.1: Bollinger Band at Different Moving Average Lengths



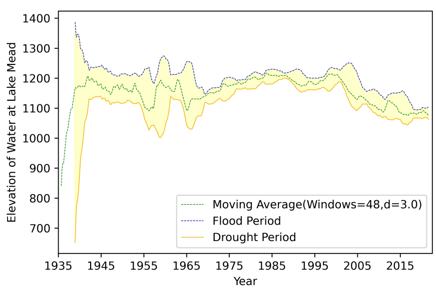
(a)



(b)



(c)



(d)

Figure 7.2: Bollinger Bands at Different magnifications

8 Conclusion and Discussion

In this paper, We propose multiple solutions to a single problem,such as *threshold value strategy* and *Bollinger Bands strategy* to determine the drought period, experimental and simulation methods to measure the flow rate. A comprehensive application of multiple methods reflects the comprehensiveness of our consideration. And according to section 7, when we set different hyperparameters for Bollinger band strategy, the extracted dry periods are roughly the same, and it is impossible to misjudge the wet period as the dry period.What's more, STTS Model and LTTS Model also have strong robustness. And the model we used does not have lots of subjective parameters such as those in Analytic Hierarchy Process(AHP), which are directly extracted from the data. Those are all the advantages we have in this research or in our modeling. However, there are still some weakness. In the process of modeling and analysis, we assumed that the annual precipitation is constant, and the amount of water directly consumed from the lake is also constant each year. Although these quantities are uncertain, they rely on more information in actual situations. It should be possible to portray them better. In the future, we are looking forward to containing more factors that could contribute to the change of the volume of the lake and further generalize a common model that could be used to all the Lakes.

References

[1] Google search. https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcTX09 ypWuGl67xJI0pKS72ci5Hi_lcosDWynA&usqp=CAU

[2] World Health Organization. https://www.who.int/health-topics/drought#tab=tab_1

[3] Celina Tebor. "the current drought is worldwide.". https://phys.org/ news/2021-07-current-drought-worldwide.html. November 14, 2021.

[4] Bureau of Reclamation. “Reclamation Announces 2022 Operating Conditions for Lake Powell and Lake Mead.” August 16, 2021. Found at: https://www.usbr.gov/newsroom/#/news-release/3950.

[5] United States Environmental Protection Agency. “Basic Information about Water Reuse.” Updated June 4, 2021. Found at https://www.epa.gov/waterreuse/basic-information-about-water-reuse.

[6] National Park Service. “Storage Capacity of Lake Mead.” Updated August 5, 2019. Found at https://www.nps.gov/lake/learn/nature/storage-capacity-of-lake-mead.htm.

[7] Google search. http://www.xjishu.com/zhuanli/52/202110413234.html

[8] Google search. https://en.wikipedia.org/wiki/Penman equation

[9] Linacre, Edward T.. “A simple formula for estimating evaporation rates in various climates, using temperature data alone.” *Agricultural Meteorology* 18 (1977): 409-424.

[10] Google search. https://www.onlinemathlearning.com/image-files/volume-cuboid.png

[11] Google search. https://mathinsight.org/media/image/image/double_integral_volume_under _surface_box.png

[12] Google search. https://mathinsight.org/media/image/image/double-integralvolumeunder-surfacebox.png

[13] Google search. https://www.nps.gov/lake/learn/nature/storagecapacityoflake-mead.htm

[14] Tortajada, C. Contributions of recycled wastewater to clean water and sanitation Sustainable Development Goals. npj Clean Water 3, 22 (2020). https://doi.org/10.1038/s41545-020-0069-3