Research on Water Level Changes in Lake Mead

Siqi Wu*, #, Puyu Fan[a, #], Wenzan Fan[b, #], Yuhan Wu[c, #]

[1]The Affiliated High School of Peking University Jia82 Daniwan Road, Haidian District, Beijing

[2]The Bolles School 7400 San Jose Blvd Jacksonville Fl

[3]The Frederick Gunn School 99 Green Hill Rd, Washington, CT 06793

[4]Quarry Lane School 6363 Tassajara Rd, Dublin, CA 94568.

*Corresponding author, wusiqi@i.pkuschool.edu.cn, [a]mrfanpuyu@outlook.com, [b]frankfan2019@outlook.com, [c]wuyuhan2019@yahoo.com

[#]These authors contributed to the work equally and should be regarded as co-first authors.

Abstract. Lake Mead, the largest reservoir in the U.S., assumes a pivotal role in the water supply. However, in recent years, the reservoir's water volume has gradually decreased, and many states have cut back on water supply. Therefore, it is crucial to explore the pattern of water level changes in Lake Mead and thereby find ways to alleviate water shortage. We first discussed the pattern of development of historical water levels in Lake Mead. We classified the annual minimum and maximum water level data of Lake Mead using hierarchical clustering, and determined the group of clusters using the Elbow Method to perform the analysis, obtaining the best result. We then classified the historical water levels into early drought periods, non-drought periods, and recent drought periods based on the clustering results. Two models were also built to predict future Lake Mead water level changes. To facilitate the overall analysis, we drew scatterplots of the maximum and minimum water levels concerning time and found that the data of the first model is more consistent and regular, so we predicted the data using Linear Fitting-based time series analysis. However, the data needed since 2005 in model 2 showed a very obvious segmentation in roughly 2011, as seen from the scatterplot. We built a breakpoint regression model to make the prediction.

Keyword: Lake Mead, water level, hierarchical clustering, time series analysis.

1. Introduction

As the largest reservoir in the United States, Lake Mead has suffered a severe water loss. Drought, climate change, and tens of millions of residents supplies in the West and northern Mexico have all contributed to the declining water level. The lower Colorado River basin even underwent the first-ever water shortage. Water users, especially Central Arizona farmers, felt the serious impact from the cuts since water delivery in the states of Arizona, Nevada, and New Mexico would reduce [1-3]. To solve this problem, researchers have developed multiple means to recycle water and continuously improve their methods more effectively and efficiently. One way that has already been implemented is to recycle wastewater. After being processed by several treatment facilities, recycled wastewater can serve many people [4].

We need get the data of the elevation of Lake Mead's water level that exceeds the mean sea level, we discover valuable information within these data. We need to define the water level of Lake Mead during drought, which means we must give the critical value of water level during drought, simultaneously discuss the general mode of water level of Lake Mead in the past data [5-7]. We can sort out the different patterns according to water level data to analysis. We also require us to develop two models to anticipate the water level of Lake Mead in the future with different reference data for two different models. Model 1 only consider data from the latest drought and assume the drought pattern would last. Hence, we must find the latest drought data and establish anticipation model. Model two require us to use the water level data from 2005~2020, which requires us to consider pattern during this period and consider this pattern would continue. With different data, the established model should be different as well [8].

According to the provided background and information collection, we need analyze the factors affecting the lake water and establish drought standards and water level prediction models. From the perspectives of the government, enterprises, and users, a waste water recycling plan is formulated. Based on the above background and the searched data, we investigated and analyzed Lake Mead, found out the relationship between drought and reservoirs, and considered recycling wastewater as a solution to water shortage.

2. DATA

It assumed that the data given in the title is true and valid, and the lowest and highest water level of each year can reflect the comprehensive situation of the water level of Lake Mead that year. The water level of Lake Mead has time continuity. The wastewater recycling plan can be carried out smoothly and on schedule.

It requires us to discuss matter of Lake mead pattern based on past data. We can adopt clustering method to classify water level data. First, to understand past water level data macroscopically, we drew the graph of water level data over time as shown in Figure 1.

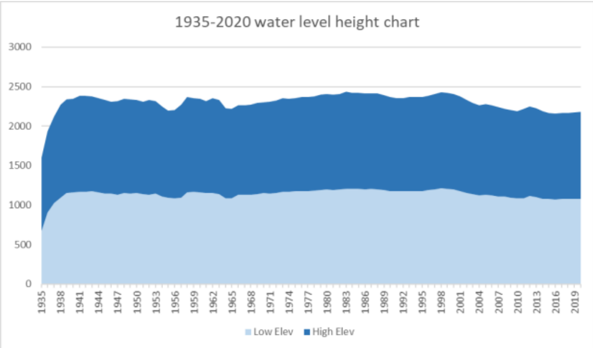

Fig. 1 Curve chart of water level data over time

3. Models

3.1. Data preprocessing

Because the lake water volume has the characteristics of regional time, which is he unceasing property during drought [9-10]. Therefore, during the process of clustering, time, as a variable should be considered. To eliminate the influence that the maximum water level and the lowest water level have little difference in the adjacent time, we need to standardize the data. Let the data be:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

Standardize it, the matrix is written as $\tilde{Z}$, and each element of z_{ij} is:

$$\tilde{z}_{ij} = x_{ij} \Big/ \sqrt{\sum_{i=1}^{n} x_{ij}^2}$$

Eventually gives this standard matrix:

$$\tilde{Z}=\begin{bmatrix}\tilde{z}_{11} & \tilde{z}_{12} & \cdots & \tilde{z}_{1m}\\ \tilde{z}_{21} & \tilde{z}_{22} & \cdots & \tilde{z}_{2m}\\ \vdots & \vdots & \ddots & \vdots\\ \tilde{z}_{n1} & \tilde{z}_{n2} & \cdots & \tilde{z}_{nm}\end{bmatrix}$$

3.2. Mode division based on Hierarchical Clustering Method

After processing the data, we use the Hierarchical Clustering Method to cluster indicators after standardization to distinguish different modes of Lake Mead. With the help of Hierarchical Clustering Method analyzation, we can avoid the effect of original and isolated points, which gives a good cluster effect.

Hierarchical clustering, also known as systematic clustering or hierarchical clustering, is mainly divided into three steps:

i. Each point in the system is treated as a class, and the Euclidean distance between the classes is calculated separately.

ii. Use the distance calculated above to calculate the distance between the two shortest points and then group them together.

iii. Finally, calculate the distance between each class and repeat the above steps until they are grouped together.

4. RESULTS

4.1. Development of water level of Lake Mead

To get a better classification result, we use **Elbow Method** to analysis, this method first needs to define different degree of distortion sum:

$$\sum_{i\epsilon C_k}|x_i-u_k|^2\ k=(1,2,\cdots,K)$$

C_k represents the k-th class. x_i represents element in k-th class, u_k represent center of mass of k-th class. Calculation of center of mass is:

$$u_k=\frac{\sum_{i=1}^{n}x_i}{n}$$

After the calculation of degree of distortion of k-th class, we add degree of distortion of each class to get total degree of distortion J. Calculation is presented as:

$$J=\sum_{i=1}^{K}\sum_{i\epsilon C_k}|x_i-u_k|^2$$

Eventually we decide where to get better cluster effect through relationship between J and k. Cluster effect visually is shown in Figure 2.

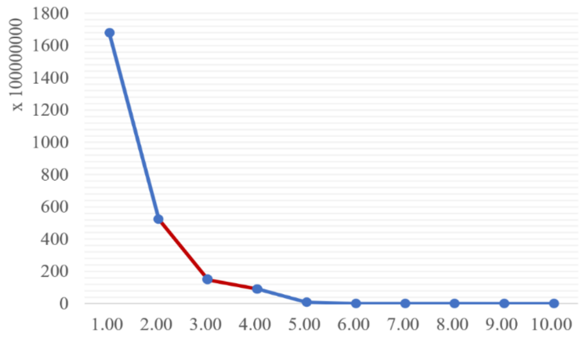

Fig. 2 Results of Elbow Method

We discovered that when k=3, the cluster effect is better, so we divide data into three categories. First model is defined as early drought period, second is defined as non-drought period, and third is defined as recent drought period. Synthesize the critical values for each category. The classification results are presented in Table 1.

Table 1. Development of water level of Lake Mead

Categories	Range
Early Drought Period	[1935, 1937]
Non-drought Period	[1938, 2008]
Recent Drought Period	[2008, 2020]

After clustering, the maximum value of the lowest water level and the highest water level in the drought period is found, which is the boundary. The statistics can be obtained. The drought standard is that the lowest water level is less than 1094.2 and the highest water level is less than 1126.

As showed in the result, Lake Mead water level experiences a "Low-High-Low" general mode. However, the characteristic of two low water level was different. According to clustering analysis result, early drought and recent drought were in different categories. During early drought period, water lever was extremely low, which creates a great bluff with data gained later. Possibly because Lake Mead wasn't as fresh, and lack of influx estuary or other geological reasons. Later, lake mead went to fresh era. Nevertheless, since 2008, Lake Mead enters recent drought period, this is also presented in data analysis. Data variation is not a bluff variation. This is caused by a yearly reduction, which leads that the reason of this recent drought period is not due to geological factors. According to news reports, recent drought and poor rainfall are the main causes of the recent drought in Lake Mead.

4.2. Water level prediction model of Lake Mead

First, in order to understand how the water level of Lake Mead changes over time, we draw a scatter plot of the highest and lowest water levels with respect to time as shown in Figure 3.

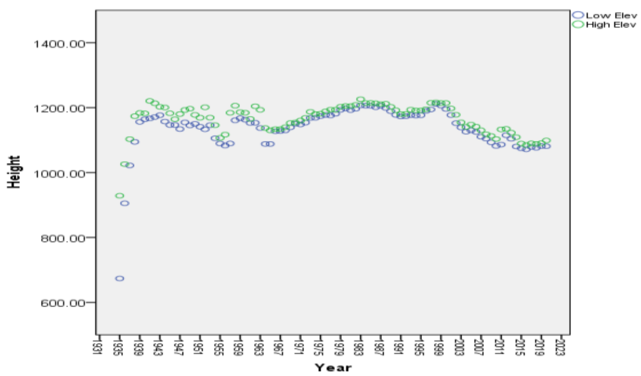

Fig. 3 Scatter plot of the highest and lowest water levels with respect to time

It can be seen from the figure that the water level change of Lake Mead has a certain historical law in Table 2.

Table 2. Water level change law

Period	Law
1931-1940	A period of rapid rise.
1941-1963	Changing greatly every year, and there was a big jump, but the overall pattern remained balanced.
1964-2011	Wave-like characteristics, and the regularity was obvious.
2012-2020	Tending to balance after falling.

The title requires us to consider only the most recent drought period data, and assumes that the most recent drought period continues. From the previous question, we know that the recent drought pattern started in 2009. If it continues, we need to consider the data after 2012. From the scatter diagram analysis, we can see that the recent drought pattern is in the regularity, so we focus on the data model that is in the regularity. Due to the small amount of data, we used the Curve Fitting toolkit of MATLAB to perform a fitting test. We set 2012 as the starting point and returned to zero. We found that Linear Fitting has the best fitting effect in Figure 4.

At the 95% confidence level, a = -3.628, b = 0.3232, c = 1070.

$$y = a \cdot (\sin(x-\pi)) + b \cdot ((x-10)^2) + c$$

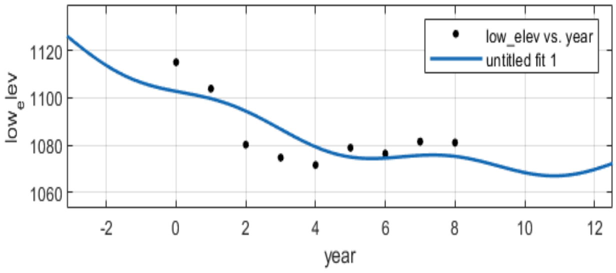

Fig. 4 Fitting results

Substituting into the formula, the lowest water level in 2025 is 1074.43, 2030 is 1070.18, and the lowest water level in 2050 is 1073.98.

From the above scatter plot, it can be seen that the annual maximum and minimum water levels of Lake Mead in 2005-2020 have "breakpoints" in 2011, so we can consider using breakpoint regression to predict, because the breakpoint regression appears The time of the breakpoint is not easy to determine. We can analyze the reasons for the breakpoint in 2011 to derive it. We use Stata software to solve the problem with 2011 as the coordinate 0 point, and the result is shown in Figure 5.

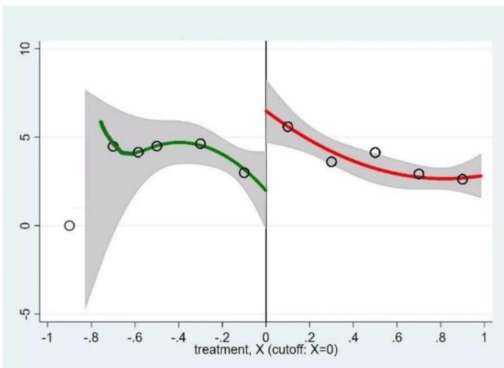

Fig. 5 Breakpoint regression prediction results

From the above figure, we can see that in 2011, the water level of Lake Mead experienced a slight increase due to rainfall and reduced water use. However, the drought criteria we previously determined showed that despite a slight rebound, Lake Mead is still at a low level. The water level, and with the continued drought, the water level has shown a pattern of decreasing year by year. From this, we can predict according to the 9-year cycle, 2030 and 2050, and are at the beginning of the cycle. The comparison table of the results of the two models is shown in Table 3.

Table 3. Comparison of the results of the two models

$Year$	$Model\ A$	$Model\ B$
2025	[1074.43,1091.02]	[1075.01,1090.62]
2030	[1070.18,1083.55]	[1087.45,1114.27]
2050	[1073.98,1085.74]	[1086.1102.81]

5. Conclusions

In this paper, when using the analytic hierarchy process to delineate the water level model of Lake Mead, we creatively used the Elbow Method to analyze and determined the value of k, making the number of model types more scientific. When predicting the water level of Lake Mead, we use the breakpoint regression method to effectively avoid endogenous problems and better reflect the relationship between variables. However, For the sake of convenience, we only use the lowest and highest water level data of Lake Mead each year, instead of using the average water level information for each month, and under certain circumstances, we have lost part of the information.

References

[1] Liang Ziliang, Xie Kun, Fu Chao, Wang Chun. Research on lake water level monitoring and prediction methods based on satellite altimetry[J]. Surveying and Spatial Geographic Information, 2021, 44(08): 172-175.

[2] Zhang Fei, Wang Jianqiang, Luo Han. Analysis and prediction of Poyang Lake water level based on time series[J]. Surveying and Spatial Geographic Information, 2016, 39(08): 35-37+41.

[3] Zhou Lusheng, Wang Qingchun. Analysis and prediction of the interannual variation of Qinghai Lake water level[J]. Plateau Meteorology, 1996 (04): 93-99.

[4] Mei Hong, Sun Yihan. The Orlando Wetland Park in the United States, famous for its wastewater recycling and treatment [N]. China Ocean News, 2011-12-16(004).

[5] Dai Kezhi. Industrial wastewater pretreatment plan in the United States [J]. Building Technology Communications (Water Supply and Drainage), 1990(02): 38-41.

[6] Zhang Aiguo, Que Lili, Li Xin, Wang Yuefeng, Cui Weiwei. A cost-benefit evaluation model for reclaimed water projects based on the perspective of stakeholders and its application [J]. Hydropower Energy Science, 2021, 39(06): 136-139.

[7] Xu Nan,Ma Yue,Zhang Wenhao,Wang Xiao Hua,Yang Fanlin,Su Dianpeng. Monitoring Annual Changes of Lake Water Levels and Volumes over 1984–2018 Using Landsat Imagery and ICESat-2 Data[J]. Remote Sensing,2020,12(23):

[8] Wang Zhuangling. Greece plans to study waste water reuse [N]. China Water Resources News, 2008-09-25(004).

[9] Zhao Wenjuan. Analysis of the development status of my country's water enterprises under the new situation [J]. Clean the World, 2021, 37(10): 93-94.

[10] Du Qingkun, Ge Mei. Application of Renewable Water Resources in Agricultural Irrigation in Water-Deficient Areas: Taking Tianjin as an Example [J]. Renewable Resources and Circular Economy, 2021, 14(10): 41-44.