E3S Web of Conferences **352**, 03001 (2022) https://doi.org/10.1051/e3sconf/202235203001
*ESAT 2022*

# Research on lake protection scheme based on clustering algorithm

Rongchang Ji [1], Zhien Yan [2], Haihao Lin [3], Hao Zheng [4]

[1] Shanghai Weiyu International School, Shanghai, 200232, China
[2] Furen International School, Singapore
[3] International Department, The Affiliated High School of SCNU, Guangzhou, China
[4] Shanghai International Studies University, Xiwai International School, 201620, Shanghai, China

**Abstract.** In recent years, global drought, particularly in the lake region, has become a crucial issue of discussion for a few specialists. In this paper, based on this background, establishes a clustering algorithm to determine the effect of major factors, uses a multiple regression model and game model analysis to determine the change of water in the area and waste water recycling, as well as discusses the gains of drought problem solving in the actual process of solving the wastewater problem.

**Keywords:** Lake Mead; clustering algorithm; river protection.

## 1. Introduction

Mead Lake, the largest reservoir in the United States of America, is a Colorado River reservoir located on the Nevada-Arizona border. In the summer of 2021, its water level is the lowest it has been since the 1930s, and the reservoir's capacity is significantly reduced due to the drought induced by global climate change and increased demand for water. Water scarcity depletes water supplies, which has a detrimental effect on agricultural areas, which will be impacted first.

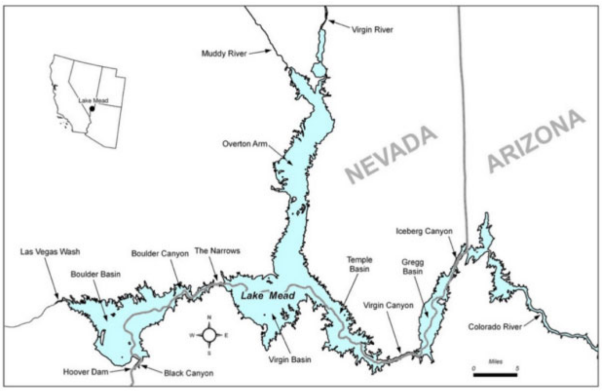


**Figure 1:** Lake Mead Overview (National Park Service [2])

## 2. A Lagrangian Model Based on Clustering Algorithm for Investigating the Intrinsic Relationships between Lake Factors

### 2.1 Calculation of Input & Outflow & Loss

Begin with dry weather and water consumption. It will diminish as a result of the dry weather at the time of entry. Outflows grew as demand increased. It is lost because to the weather, and hence evaporates.

Inflow

It enters via the Mississippi River. It supplies an adequate supply of irrigation and industrial water for daily living. Due to the distinct climate, landform, and other circumstances, the east and west tributaries have radically different hydrological characteristics. Due to the difference in flow direction, the silt content of the river will be varied.

Outflow

The outflow originates from a dam, which has the capacity to store water in response to fluctuations in river flow. Between upstream and downstream water, there is a large height difference. Irrigation, power generating, and water supply can all benefit from increased stability. However, because each location has unique geographical and climatic characteristics, the water production will vary for natural reasons.

Loss

The lake will have its own evaporation loss. Droughts and rising temperatures, for example, result in increased water demand, creating a vicious cycle. Thus, this is the graph of the economy's demand root supply. In the latter part of the figure, the intersection indicates that supply and demand are exactly equal. However, supply is currently less than demand, creating a vicious spiral.

## 2.2 Model Preparing

MacQueen's K-means clustering algorithm is also referred to as the K-means algorithm. It is mostly used to classify data objects. Clustering is the process of grouping data with similar properties into a single set using predefined rules. This algorithm is iterative. It determines the distance between the sample object and the cluster's center point first, and then splits each sample into the cluster domain that is closest to the sample object.

### *2.2.1 The Anatomy of the Clustering Problem*

A data matrix is made up of $n$ samples' p-element observation data:

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

Each row corresponds to a sample, each column to an attribute, and $x_{ij}$ corresponds to the observed value of the j-th attribute of the i-th sample. In this question, each row has Lake Mead data for a separate time period, and each column contains Lake Mead-related parameters such as inflow, outflow, and loss.

### *2.2.2 Set Up the N-sample Observation Data for the P-element*

At this stage, each sample can be thought of as a point in the P-element space, corresponding to a P-dimensional vector. The distance between the two vectors is $d(x_i, x_j)$. The Euclidean distance is the most often used distance measure in cluster analysis, that is:

$$d(x_i, x_j) = \left[ \sum_{k=1}^{p} (x_{ik} - x_{jk})2 \right]^{1/2}$$

Prior to doing K-means clustering, it is necessary to calculate the categorization number k based on the actual instance. The selection samples within each class are referred to as cluster points, and they are normally chosen using the minimum and maximum principles.
The following approaches are used to determine the similarity of two data objects in a two-dimensional sample space: Euclidean distance, Minkowski distance, and Mahalanobis distance.
Among them, $x_i^m$ is the center of gravity of class $G_i^{(m-1)}$, $x_i^m$ is not necessarily a sample. When m gradually increases, the classification also tends to be stable. At the same time, $x_i^m$ can be regarded as the center of gravity of $G_i^{(m)}$ approximately, that is, $x_i^{(m+1)} \approx x_i^{(m)}, G_i^{(m+1)} \approx G_i^{(m)}$ calculation can stop at this time.

## 2.3 Model Solving

To begin, load the data in the attachment into MATLAB and run the clustering method to obtain the result graph displayed below.

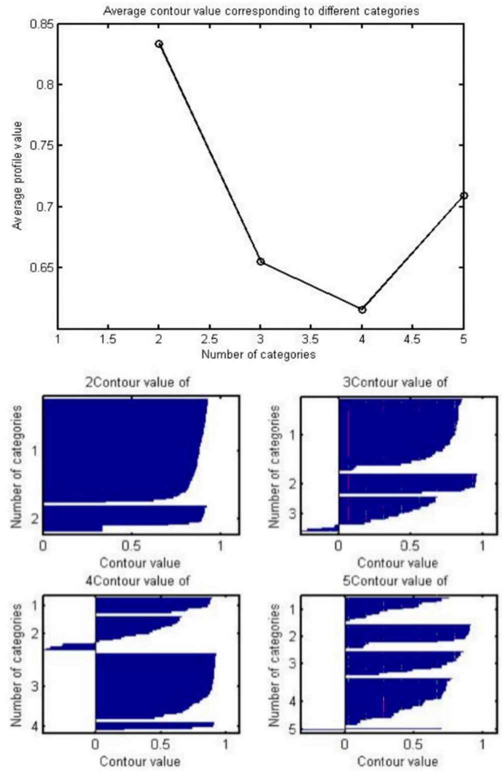


**Figure 2:** Clustering Analysis for the Original Data & Number of Categories Vs. i-th Contour Value of Clustering

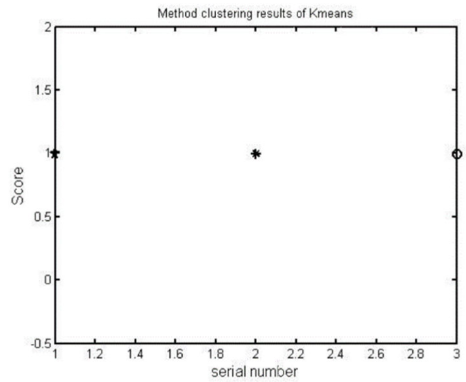


**Figure 3:** Result Graph

【Clustering result】
Type l: Center point: 0.002 This kind of sample number: 2.
Type 2: Center point: 0.002 This kind of sample number: 3.
Type 3: Center point: 0.996 This kind of sample number: 1.
According to the analysis, it can be determined that the main factor of influence is time.

### 2.4 Establish the Link between the Variables

The Lagrangian model is used to describe a multi-element geographic environment system in which numerous (more than two) factors impact and correlate with one another. As a result, the model of multiple geographic regression is more universal.

## 3. The Algorithm for Volume Estimation Based on Discretization of Surface and Idealization of Shape

### 3.1 Overall Pattern of Historical Data of Lake Mead Water Level

The Mead Lake lacks a regular formed surface or a regular shaped body beneath the surface. We cannot utilize pre-existing formulas to compute the volume of Mead Lake. As a result, we must rely on estimation to determine the approximation value of Mead Lake's volume. There are numerous approaches that can be used, as demonstrated below.

### 3.2 Method 1: Discretization of Surface

#### *3.2.1 Derivation of Principle*

We can divide the volume of Mead Lake into tens of thousands of little cuboids with a height change of $dh$. Mead Lake's volume is equal to the total of the volumes of all these little cuboids combined that corresponds to a precise area value at each level of height. Additionally, because the Mead Lake should be shrinking in areas according to its unique height. As a result, we can conclude that the surface area of the Mead Lake is the greatest among the areas at every height.

#### *3.2.2 Function for Volume Calculation*

We will refer to Mead Lake's surface area as $A$ cause the values of the areas are constantly decreasing, we can utilize the collected data to determine the area's average rate of change at its specified height. This average rate of change is denoted by the constant $k$.

The average rate of change in area can be used to determine the volume of Mead Lake by deriving equation:

$$V = \int_{bottom}^{top} (A - kx)dx$$

Where $A - kx$ denotes the area at each discrete height. Each time the height of Mead Lake is reduced by one unit, the surface area of the lake decreases by $kx$.

According to the formula we derived, the constants A and k will have the most influence on the outcome of our estimation. That is, when the surface area of Mead Lake increases, the area beneath the surface increases proportionately. However, the most significant element affecting the outcome is the constant k. If the average rate of change is greater, our estimation result will be smaller, since each level of area will fall in a greater multiple.

### 3.3 Method 2: Idealization of Shape

#### *3.3.1 Theoretical Model of Lake Bottom Shape*

Because Mead Lake is unevenly shaped, we can assume that its shape is regular. We can initially assume that the shape of Mead Lake approaches a half sphere due to the buildup and precipitation of silt under the water, which makes the lake's bottom spherical. As a result, we may estimate the volume of Mead Lake using the formula for computing the half sphere:

$$V = \frac{2}{3}\pi r^3$$

The radius of this half sphere, is equal to the distance between the surface of Mead Lake and the bottom of Mead Lake. Also, based on this calculation, we can assume that the Mead Lake's radius is directly proportionate to its volume and is the only thing that will influence the Mead Lake's volume. The radius is equal to the elevation of Mead Lake. As a result, the greater the height of Mead Lake, the greater the volume of Mead Lake.

#### *3.3.2 Refinement of volume*

Furthermore, we can suppose that the shape of Mead Lake is like a truncated cone, which is a reasonable assumption given that the lake's bottom precipitated a lot of silt. As a result, we may estimate the volume of Mead Lake using the volume formula for a truncated cone, such as:

$$V = \frac{\pi h(R^2 + Rr + r^2)}{3}$$

Where The radius $R$ is defined as the distance between the center of Mead Lake and the lake's margin. The height $h$ shall be defined as the distance from the top of Mead Lake to its bottom.

As a result, we can conclude that there are two elements that will influence our computation of the volume of Mead Lake. We can deduct from the formula that the height of Mead Lake has a direct influence on the volume of Mead Lake. The volume of Mead Lake increases as the distance between the top and bottom of the lake increases. If the radius of the Mead Lake is increased, the volume of the Mead Lake will increase as well.

### 3.4 Summary of Methodology

Our hypothesis of possible methods for determining the volume of the Mead Lake are all conceivable based on the data provided, which indicates that the higher the elevation of the Mead Lake, the larger its volume.

The lake's area is proportional to the increase and decrease in the volume of Mead Lake. Because the area and elevation of Mead Lake are not proportionate to the volume of the lake, the relationship between volume, area, and elevation of Mead Lake is not proportional.

 https://doi.org/10.1051/e3sconf/202235203001

# 4. Drought in Lake Mead

## 4.1 Definition of a Dry Period

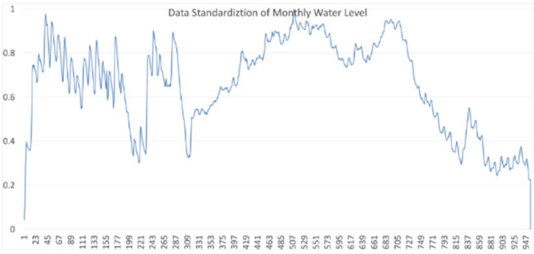


**Figure 4:** The Standardized Elevation of Lake Mead Vs. Time

We standardize the water level data by converting it to a range of 0 to 1, ensuring that the data is not too huge or too small for analysis. To facilitate study, we use a number between 0 and 1 to denote the water level of Mead Lake. As seen in the graph above, the data is highly concentrated between 0.2 and 1. As a result, we define the number 0.6 as the boundary line—any value less than 0.6 reflects the drought period.

## 4.2 Prediction of the Water Level

In 1936, the sea level was the lowest in nearly a century. However, as time passes, the water level continues to increase directly.
In the decade after 1936, the value of Sea level fluctuated between increases and decreases. However, the numbers have been extremely fluctuating over the last decade, but you can see that the total number has been falling.
Even the sea level began to plummet over a cliff in 1952, albeit it also increased dramatically during the next five years. This state of insecurity persisted only until 1961.
In 1961, the rate of sea level rise began to slowly increase. By 1986, the SEA level's value had begun a steady fall.
Sea level began a continual drop in 2002. This is the first time in nearly a century that such a protracted drop has occurred.
In the recent decade, there has been a reversal, although only temporarily.
The sea level has been decreasing in the years leading up to 2020.
As a result, I expect that the SEA level will continue to decrease in the future, up to and including 2030, unless something is done to reverse the trend.

## 4.3 Comparison of Different Drought Periods

According to these data, the earlier dry period was more severe, but within five years, the dry period ended abruptly, and the sea level swiftly surged. Although the sea level fluctuated, it was not until 1952 that a second dry spell occurred. This dry spell is two points longer than the previous one. To begin, the sea level is not as low as it was during the initial dry spell. The second argument is that the time span between the second and third droughts is quite short. The third dry spell was quite like the second. Sea level then began a steady, sluggish rise that lasted only until 2002, when the first dry period in modern times occurred. In recent years, the current drought has steadily eclipsed 1936 in terms of sea level, and its duration is the longest on record, while sea level has been rapidly falling. In 2015, there was only a modest reversal. Thus, the present dry spell does not just outlast the preceding one in terms of duration. Simultaneously, the sea level decline eventually outpaced that of the early drought era.
The water level was extremely low over the first few months. This is most likely because the reservoir was recently completed and water accumulated in the reservoir over time. As water accumulated in the reservoir, the water level continuously increased during the next two decades, from 0.042 to 0.73. Between the 21st month and the 1960s, the water level fluctuated significantly and was frequently greater than 0.6. (Even though there were once two temporary drought periods, and the water level fell under 0.4). Since the early 1960s, the total water level has increased significantly, and the fluctuation has become too minor to alter the trend. The water level peaked between approximately the 500th and 700th months, which corresponds to the 1980s and 1990s. This is sometimes referred to as the "full pool" level.
The explanation for the rise in water level and "full-pool" level could be the growth of the US economy and population. In the 1950s, following World War II, the baby boom occurred in the United States, resulting in a significant increase in population and an increase in demand for food and supplies. This tendency explains the expansion of manufacturing in America, as more factories opened and increased carbon dioxide production. As carbon dioxide accumulated in the atmosphere, the greenhouse effect intensified and resulted in an increase in temperature. As a result, the snow on the Rocky Mountains, the source of the river and lake water, melted more quickly and converged into the river and lake.
However, after the mid-1990s, despite the slight fluctuation, the overall water level continued to decline drastically, eventually falling below 0.6 in the late 1990s. Even if the water level increased from 0.2 to 0.5 a few years later, the downward trend persisted, and the water level dipped below 0.4 once more. According to credible internet sources, the warmer environment appeared to increase the ease with which water evaporated into the atmosphere and significantly reduced the required water supply after the 1990s, but the temperature increase did not halt due to the greenhouse effect. While some scientists are attempting to determine the state of precipitation in the Rocky Mountains, we can be certain that rain will be insufficient to compensate for the significant water loss. This significantly contributes to the lake's subsequent drought, and the growing population also consumes a significant amount of water.

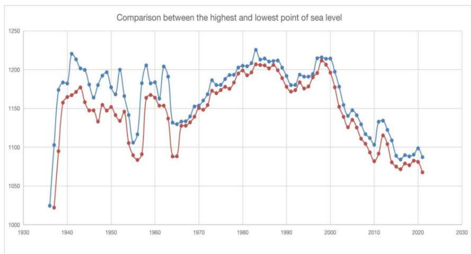


**Figure 5:** High Elevation and Low Elevation of Lake Mead

Since the greenhouse effect is still present and unresolved, the reservoir's water supply is likely to continue diminishing, and thus the lake's water level may continue to decline until there is nearly no water left.

The water level is at its lowest point, as seen by the graph above. This is possibly because the reservoir was still under construction at the time and the water was still pouring into the lake. However, both the lowest and highest water levels subsequently increased considerably. We can see from the graph that the maximum water levels occurred in the 1940s, 1980s, and 1990s. Americans consumed a significant amount of fossil fuels. This produced a large amount of carbon dioxide, resulting in the greenhouse effect and, thus, an increase in temperature. When the temperature increased, the icebergs melted more easily, resulting in the rise in the sea level. More seawater entered the rivers and entered the lake. As a result, the lake's greatest water level was relatively high in 1940. Carbon dioxide and other greenhouse gases absorb some of the earth's infrared radiation, trapping it in the atmosphere.

As a result, more heat is absorbed than is lost, and the temperature continues to rise. This is the mechanism by which the greenhouse effect operates. When these greenhouse gases occur naturally on the earth, the earth reaches its ideal temperature for life. However, because human activities create an excessive amount of greenhouse gases, global warming occurs, temperatures rise rapidly, the Antarctic melts, and sea levels rise. Thus, lowering sea levels requires reducing carbon emissions and collecting surplus carbon dioxide in the atmosphere. In comparison, carbon dioxide levels in the atmosphere have increased dramatically as a result of the global industrial revolution and the combustion of fossil fuels such as coal and oil. Carbon dioxide in excess functions similarly to the glass or plastic sheeting in a greenhouse, a phenomenon dubbed the "greenhouse effect." This results in an annual increase in world temperature, which causes the sea's temperature to rise, which causes the water to expand as it heats up. The warming of the oceans and atmosphere at the poles is melting ice sheets in the polar regions and seas around areas such as Greenland, contributing to increasing sea levels.

# 5. Wastewater Recycling Project

## 5.1 The Plan's Critical Components

Appropriate Legal and Supervisory Framework

New regulations are needed to regulate the reuse of water resources and the use of wastewater by-products. Regulations governing the quality of connected items are currently inadequate or nonexistent, and market insecurity may deter investment. Investment and regulatory incentives can help to foster and grow relevant product markets.

Cost Recovery and Appropriate Financing Mechanism

Wastewater management and sanitation are frequently viewed as high-cost, capital-intensive enterprises. This is especially true for large-scale centralized processing systems, which frequently need a considerable initial investment. Additionally, it demands rather significant operating and maintenance costs in the medium and long term to avoid rapid system degradation. If long-term investment in mechanism development and human resource development is not made, this problem will deteriorate. However, if wastewater management is not adequately funded, the costs will be higher, particularly in terms of direct and indirect harm to human health, social and economic development, and the environment.

Knowledge and Capacity Building

For policymakers, researchers, practitioners, and public institutions, data and information on wastewater generation, treatment, and utilization are critical. Only with data and knowledge can national and local action plans be developed that safeguard the environment and ensure the safe production of wastewater Sexual use. The amount of wastewater produced and the type of waste water are critical pieces of information for ensuring the health and safety of humans and the environment. To improve wastewater management, it is vital to ensure that the necessary capacity building occurs. The wastewater management industry's organizational and institutional capacities are frequently lacking, which makes investment in large-scale centralized wastewater management systems or small-scale on-site treatment systems challenging.

Public Awareness and Social Recognition

To obtain public acceptance, maximize the benefits of wastewater reuse, and avoid negative impacts, we must assess, manage, monitor, and report on the health impact of water reuse. In terms of drinking water (reuse of tap water), we need to conduct extensive information distribution initiatives to help the people develop trust in the system and overcome their disgust with it.

## 5.2 Influence of the Leader's Decisions

1) Increase support for agricultural cooperative organizations, especially water associations and cooperatives, and scientifically guide the healthy and sustainable development of agricultural cooperative organizations. However, the investigation found that many cooperative organizations had no real name, and the problem of fake and shoddy goods was serious, which made it difficult to play its role. Therefore, in view of the above problems, the government should establish a cooperation mechanism, strengthen supervision and inspection, and carry out centralized rectification of the problem cooperative organizations, to ensure the healthy and sustainable development of cooperative organizations.

2) Grassroots organizations can organize training, exchange meeting or get-together from time to time, so that farmers can have the opportunity to improve their social skills, broaden their social network, increase social capital, and deepen the trust and cooperation among farmers. Strengthen cooperation with rural cooperative organizations, establish rural multimedia information network, broaden access to disaster information, timely and accurate transmission of seasonal drought and other climate change information and related adaptive technology to farmers, reduce the cost of adaptive behavior, reduce the risk.

3) The government should increase investment in rural basic education, widely carry out publicity and training on seasonal drought and other climate change knowledge, improve the ability of farmers to acquire technology and information, and urge them to take the initiative to adapt measures.
4) strengthen the artificial weather control ability and emergency response capacity building, improve the meteorological information transmission services, improve service quality, broaden the service network, standard service, improve the early warning and forecasting ability, transmit the relevant meteorological information in a timely manner to the hands of farmers, improve farmers sensitivity to climate change, increase their understanding of the science of climate change. Enhance farmers' sensitivity to climate change and enhance their scientific understanding of climate change. Let farmers take the initiative to take different adaptation measures before and after climate change to avoid and reduce the destructive impact of climate change on agricultural production.
5) the government shall be based on local area type and hydrological conditions, the investment to strengthen water conservancy irrigation facilities, for the maintenance and construction of rural small water conservancy facilities, improve local water conservancy irrigation condition, strengthen the research, demonstration, and promotion of water-saving irrigation technology, to reduce the loss of seasonal drought, climate change brings to the farmers.

### 5.3 Recycling Plan of Wastewater

Conduct wastewater recycling surveys to ascertain local policies, public understanding of the importance of wastewater recycling, and Lake Mead's use.
Educate the public about wastewater recycling while bolstering research and development efforts to lower costs. Ensure comprehensive waste water recycling and strive to increase water use in daily life following waste water recycling. Simultaneously, the government enacted pertinent legislation to increase awareness; businesses participated competitively and successfully executed the plan.

## References

1. https://earthobservatory.nasa.gov/images/148758/lake-mead-drops-to-a-record-low
2. https://www.nps.gov/lake/learn/nature/overview-of-lake-mead.htm
3. https://wap.cnki.net/touch/web/guide?ivk_sa=1024320u