HiMCM 2021: Problem B

HiMCM Team 11617

Summary

Climate change and increasing demand for water in the American Southwest have led to immense drought conditions in Lake Mead, the largest reservoir in the country. In order to prevent the consequences of water shortage from disturbing the lives of the 25 million people Lake Mead serves, we seek to restore balance between water storage, consumption, and reuse. In this paper, we present multiple predictive models of water elevation at Lake Mead. Furthermore, we propose a general model that fits well with our presented approaches and allows for the modelling of complex volumetric relationships, permitting a greater degree of complexity and allowing for the selection of any time-based and statistical analyses.

Prior to developing our model, we delineate and compare the effects of several volumetric factors that contribute to changes in water levels. These factors are categorized as naturalized inflow, outflow, and loss and have varying contributions to Lake Mead's fluctuations. Our exposition on these items direct us toward the objective of developing an accurate and adaptable model.

In our model development, we first explore the relationship between reservoir water elevation and storage volume, and we consider criteria for identifying drought periods based on consumption, streamflow, and elevation. We employ a Seasonal Auto Regressive Integrated Moving Average (SARIMA) guided by a rolling forecast algorithm on elevation data from 2005-2020 as our robust predictive model, which we compare to a linear regression model. SARIMA accounts for the periodic nature of water elevation and projects several decades of future elevations in Lake Mead.

Along with empirical evaluation of our developed models, we propose an improved approach that fits well with the aforementioned approaches and considers the complex relationship between climate change and other volumetric factors. Combining statistical analysis with time-series forecasting, our approach is modular and allows for increasingly accurate predictions when paired with additional data.

Finally, we provide a priority order for addressing drought conditions on human activity, and we outline two phases of a provisional water reuse plan in addition to suggestions for water conservation practices. A miniature case study is included to demonstrate that effectiveness of the proposed plan with respect to success conditions and metrics.

Keywords: Water Conservation and Recycling, Climate Change and Hydrological Drought, Curve Fitting by Method of Least Squares, Time-Series Forecasting on Discrete-Time Data, Seasonal Auto Regressive Integrated Moving Average (SARIMA)

Contents

I Introduction

1.1 Background

Lake Mead, located on the Colorado River on the Nevada-Arizona border, is the largest water reservoir in the United States. In the summer of 2021, drought, coupled with increasing demand for water, caused the water level of Lake Mead to reach a historic low. If left unaddressed, this water shortage may have significant adverse effects on surrounding industries and communities.

A proposed solution to this problem is to recycle wastewater with fit-for-purpose specifications. Evaluating the effects of the drought on Lake Mead will help to determine the extent to which a water recycling plan must be implemented.

1.2 Problem Interpretation

Our objective is to create predictive models for the water level of Lake Mead over time. Based on conclusions drawn from these models, we develop a plan for ameliorating the impact of the drought on human activity in the area. We consider the following elements:

- Specific factors that to contribute the inflow, outflow, and loss of water in Lake Mead and their relative impacts.
- The relationship between elevation, surface area, and volume of water in Lake Mead.
- Criteria for identifying drought periods.
- The efficacy of wastewater recycling in ameliorating water usage limitations specifically in the region about Lake Mead.

1.3 Assumptions

1. Colorado River Lower Basin:
   The model will be developed based on data gathered for the Lower Basin of the Colorado River.

   Justification: Lake Mead is located in the Lower Basin of the Colorado River and serves the states Nevada, Arizona, New Mexico, and California, which are downriver. Data on natural flow in the Lower Basin are inherently representative of upriver activity. Therefore, reliable sources of data on the Lower Basin are sufficient for modeling water level in Lake Mead.

2. Groundwater Discharge:
   The groundwater discharge of the Colorado River is included in the inflow of Lake Mead.

   Justification: Groundwater enters streams as baseflow which is absorbed into the streamflow of the Colorado River, which contributes to the inflow of Lake Mead. Similarly, groundwater recharge is already accounted for. Therefore data collected on the surface water and precipitation of Lake Mead is sufficient for analyzing the water level.

3. Wildlife:
   In determining relationships between volume and water level, wildlife activity is not considered as a major factor.

   Justification: Although wildlife does reside in the area about Lake Mead, it is a mathematically valid option to construct the model based on flow frequencies.

4. Lower Colorado River Operations:
   Current operational procedures of Lake Mead will stay consistent throughout the implementation of wastewater recycling in the region.

   Justification: It is expected that river and dam management do not see major changes and that the Drought Contingency Plan (DCP) remains in effect until a solution to the water shortage is implemented.

II Preliminary Analysis

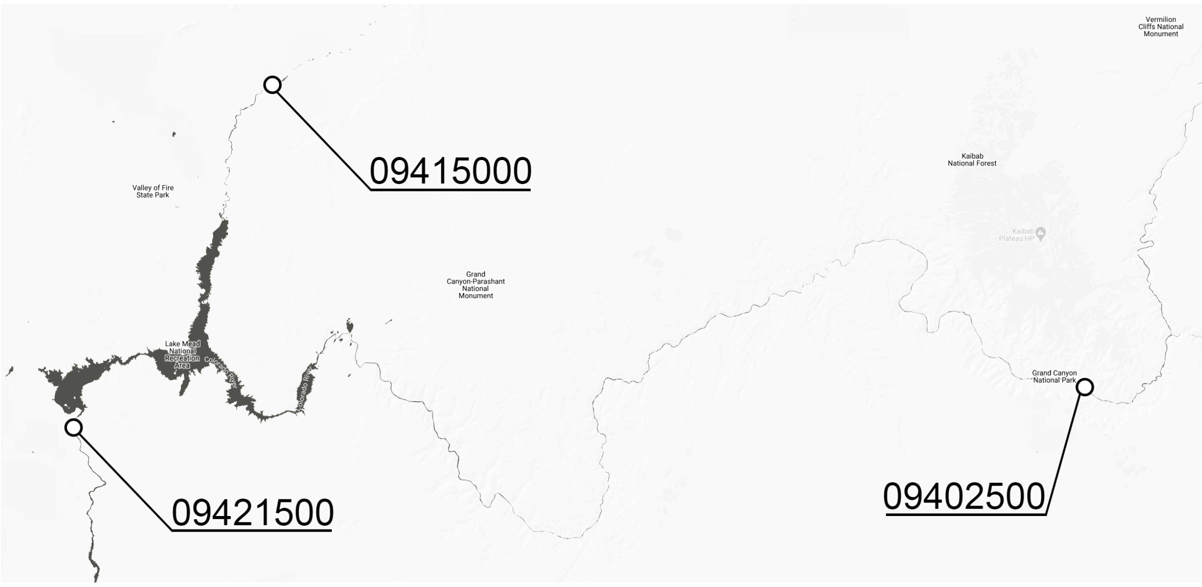


Figure 1: Map of region surrounding the Lake Mead reservoir. Three stream flow gauges are marked, with inflow arriving into the reservoir from the East and North [11].

2.1 Identifying Volumetric Factors

The volume of Lake Mead fluctuates heavily and has shown to be highly variable over the course of the past century. Here, we describe the primary factors that impact the volume of Lake Mead, considering the reservoir as a closed system. Figure 1 maps the region and isolates the area that we consider to be representative of this system. This approach allows us to analyze historical patterns and effectively model the lakes volume as a function of the identified factors. We show that this interpretation is consistent with historical patterns and propose external factors that allow for the consideration of future changes in regional climate.

2.1.1 Volume from Elevation

Due to the irregular shape of the Lake Mead reservoir, identifying an accurate conversion between water level and water volume can allow for more convenient calculations. Utilizing the provided data from 2010, along with additional records from The Boulder Canyon Operations Office's annual Water Accounting Reports, we outline one straightforward method for such a conversion [11].

Figure 1 considers measured volumes of Lake Mead at varying points throughout the period from 2007 to 2020. Paired with the corresponding elevation at these times, we find that a strong linear relationship can be drawn,

$$V(E) = 0.0866727E - 83.6569$$

where V denotes the volume associated with some elevation E.

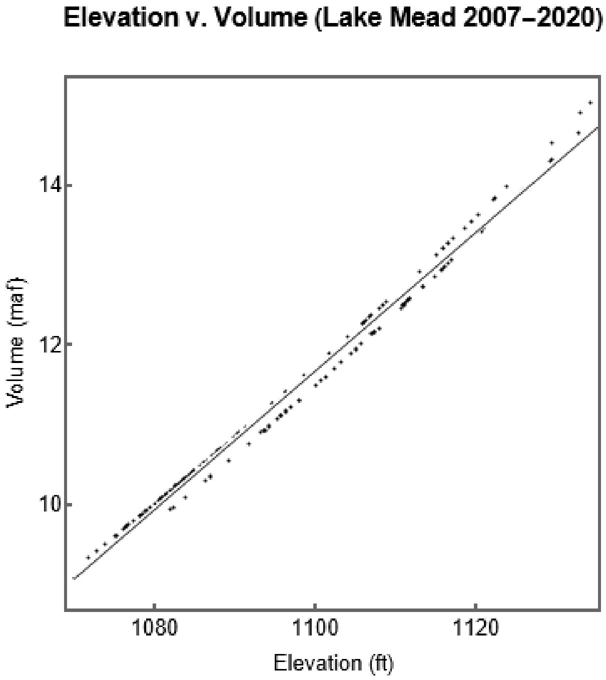


Figure 2: Elevation vs Volume for Lake Mead, plotted monthly from 2007 to 2020 [11]

It is important to note that this linear relationship models this conversion *well* only for the observed elevations (between approximately 1070 and 1140 ft). While this interval aligns with the typical loads stored in the reservoir, the extremes of lake elevation (i.e. when the lake is either close to full capacity, or close to being fully empty) are not adequately modelled.

This is best illustrated by comparing the conversion for the provided elevations from the Bureau of Reclamation's designations for Storage Capacity and Dead Storage, which are 1229.0 ft and 895.0 ft respectively [13]. At these extremes, the linear model is approximately 22,863,848 AF and -6,084,834 respectively.

While largely inaccurate for the extremes, we do not consider such conversions in our calculations and thus do not suffer from the consequences of such inaccuracies. Drawing more complete data, particularly at the extremes, may allow for a more precise conversion between volume and elevation, where a non-linear relationship can be utilized.

2.1.2 Inflow

The Colorado River, flowing into Lake Mead from the East, provides the reservoir with around 96% of its net inflow. The Colorado River basin splits into two independent basins: the lower and the upper basin. We consider the total natural flow of the surrounding region, primarily focusing on the flow of the Colorado River as it leads into the lower basin, where Lake Mead is situated. Inflow to Lake Mead results from the below factors:

- **Natural flow** of the Colorado River and its tributaries
- **Precipitation**
- **Surface Runoff**
- **Treated Wastewater** which may be released upriver

We primarily focus on the effects of **natural flow**, which accounts for nearly all of the collected inflow and is concretely measurable. Utilizing naturalized flow data documented by the Bureau of Reclamation [11], we consider nearby stream gauges that measure naturalized water flow at their respective sites. Figure 1 depicts the location of 3 such gauges, located on the Colorado River as it flows into Lake Mead, on the Hoover Dam as water is released, and within one of Lake Mead's tributaries.

We consider data gauge 09402500 located at the Colorado River Near the Grand Canyon due to its proximity to Lake Mead. Taking inflow measurements from this gauge allows us to accurately analyze the extent of the inflow fluctuations that occur annually. Figure 3 displays this region's inflow over the course of the past century, and indicates a consistent decline in naturalized flow.

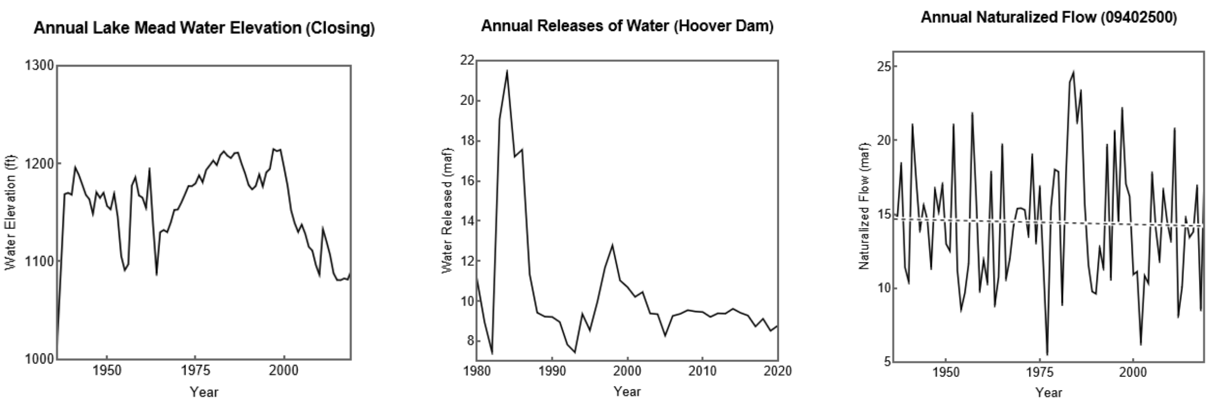


Figure 3: Annual data collected from the Bureau of Reclamation regarding elevation, releases, and inflow of water surrounding Lake Mead.

2.1.3 Outflow

Outflow from the reservoir expresses the total volume of water that is released or consumed. More concretely, we consider the following factors that influence the total volume of outflow:

- **Water released** from the Hoover Dam
- **Water consumed** directly from the reservoir and the surrounding Colorado River

Figure 3 illustrates the fluctuations of such factors, including the released water volume, based on extensive data reports from the Bureau of Reclamation [11]. These reports suggest that the annual volume released by the Hoover dam through regular operations has steadily gone down over the past 40 years from its peaks, but has remained stable around an average of around 10 MAF (million acre-feet).

Outflow incurred by consumptive uses is cited extensively by the Bureau of Reclamation, which notes the requirement for all parties to report their total consumption. Between the point of measured inflow (gauge 09402500) and the Hoover dam, consumption is reported to come from multiple parties, including, but not limited to:

- **The Boulder Canyon Project**
- **The Robert B. Griffith Water Project**
- **Pacific Coast Building Products Inc.**

Due to the self-reported nature of consumptive uses, the reported values are highly variable. For our models, we consider calculating the total consumption explicitly based on the outflow and inflow suffered by other factors and considering annual changes in volume. While this is not directly required, we note this fact when considering the availability of reliable data.

2.1.4 Loss

Loss from the reservoir refers to unintentional removal of water. Losses can occur due to various causes, of which we consider

- **Seepage**
- **Surface evaporation**

Seepage in dams is reduced ordinarily, but losses by evaporation can be extensive. Annual evaporation at Lake Mead from 2011-2015 was found to be an average of 1818 mm, or approximately 5.96 ft [3][4].

2.2 Classifying Droughts

2.2.1 Drought Contingency Plan

The Drought Contingency Plan (DCP) has been in effect since 2019 to mitigate drought effects and protect the water resources of the Colorado River. Under this agreement, Lower Basin states must make water resource contributions when Lake Mead is reduced to an elevation of 1,090 ft.

To illustrate this policy, we provide an example with the state of Arizona. DCP Tier 1 shortage conditions were triggered in summer of 2021 when Lake Mead fell below 1,075 ft, and the Colorado River water supply to Arizona was cut by 320,000 AF. If Tier 2 is triggered at 1,050 ft, then Arizona must cut 400,000 AF. Finally, if Tier 3 is triggered at 1,025 ft., then Arizona must cut 480,000 AF.

2.2.2 Identifying Drought Periods

For the purposes of our investigation, we develop identification procedures for droughts that considers both the data we have available and proposes potential metrics for more informed identifications.

For the available data, primarily consisting of the changing elevation levels, as well as the inflow and outflow metrics, we consider a period as one of drought if water levels decrease in excess of 50 feet in a period of 5 years.

This simple model allows us to identify potential droughts early and to plan accordingly. Water level decreases indicate increasing demand for consumptive use and release, which are typically expected during an extensive period of drought. Additionally, naturalized flow may be considered as a means of predicting whether the issue will be exacerbated in the coming years.

Since a decrease in naturalized flow will typically result in higher demand for consumption, we anticipate the two factors correlate to describe a period of drought within the region, though they may not necessarily coincide.

2.2.3 Alternate Identification

Since there already exist DCP elevation benchmarks for classifying the drought conditions of Lake Mead, we take water level into consideration when generating criteria for determining when drought periods occur. We must also factor in the rising demand for water over time to better understand the water shortage. The opposite trends of these two factors is the primary cause of drought. Therefore, we define a drought to be a period of at least two consecutive years for which

$$DroughtIndex = \frac{\sum W}{S} \geq 20\%$$

where W includes withdrawals for domestic, industrial, and agricultural uses that are not immediately replenishable, and S stands for the quantity of available water supply. We choose 20% [2] as the threshold in order to allocate a minimum time span of five years for reform before the supply is depleted.

2.2.4 Another form of Identification

Furthermore, to confirm and identify drought periods, we can use a pre-existing scale for measuring droughts. Given that these droughts are only characterized by their change in water, we only need to look for scales measuring hydrological droughts. There are four such scales that suffice: the Palmer Hydrologic Drought Index (PHDI), the Surface Water Supply Index (SWSI), the Reclamation Drought Index (RDI), and the Streamflow Drought Index (SDI). Due to the fact that the other indices relied heavily on precipitation, something negligible in the region, SDI was the most appropriate index to identify a drought in the lake. Based on this index, drought periods are defined to be times at which the $SDI \leq 0$. Applying this criteria to our SDI calculations for the region we identify drought periods: $1936-1941$, $1946-1947$, $1950-1952$, $1953-1957$, $1959-1962$, $1963-1965$, $1966-1968$, $1972-1973$, $1974-1975$, $1976-1978$, $1981-1982$, $1988-1993$,$1994-1995$, $2000-2005$, $2007-2008$, $2010-2011$, $2012-2014$, $2015-$.

2.3 Verifying Measurements

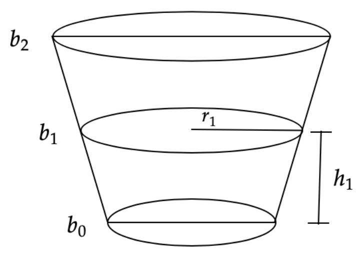
One way to verify the elevation, area and volume relationships shown in Table 1 is through crude geometric estimations by simplifying the irregular shape of the reservoir.

Elevation (feet)	Area of Lake (acres)	Volume of Lake (acre-feet)
1229.0	159,866	29,686,054
1219.6	152,828	28,229,730
1050.0	73,615	10,217,399
895.0	30,084	2,576,395

Table 1. Area and Volume of Lake Mead by Elevation Level

Provided a contour map of some precision, we consider condensing the volume at each contour level into a frustum maintaining the given surface area at that elevation in the top base, and the surface area of the next lower contour level as the bottom base. We then stack the frustums into a single complete frustum of a cone and find its associated volume. This can be found by applying

$$V = \frac{h_n}{2}(b_n + b_0)$$



where h is the absolute depth from the top base that has area b_h of a given layer of the combined frustum to the absolute bottom base with area b_0.

We utilize the provided surface area and elevation values to set up a frustum approximation of Lake Mead.

First, we set to the absolute bottom base to elevation 900 ft, so the area of the absolute bottom base is around $30,000$. This is consistent with the provided Dead Storage value.

Then we substitute the surface areas at the elevations of 1050 ft and 1200 ft, at about $73,500$ and $155,500$ respectively, which returns the volumes $27,750,000$ and $7,762,500$ respectively.

Given that both volumes are within 25% error of the volumes provided and many intermediate estimations were made, this crude method is within reason.

Given enough data on elevation and corresponding surface area, disk integration using a function $R(x)$ for radius at various depths in the frustum could be used to determine volume more accurately.

$$V = \pi \int_0^h R(x)^2 \, dx$$

Alternate Method For a more precise way to compute the area and volume of the lake without access to surface area, we suggest breaking up a satellite image of the lake into smaller parts such that the outer edge of the lake could model a mathematical function.

This can be achieved with Lagrange Interpolation, in order to obtain polynomial functions that represents the outer edge of the lake along each set of contour lines.

Integrating each of these functions yields the surface area of each contour, which can be utilized with the previous model in a frustum-based volume approximation.

III Development of General Models

3.1 Model 1

Based on our proactive drought criteria, we establish 1999–2020 as the most recent drought period and form assumptions based on data from only this period.

For this first preliminary model, we utilize linear regression to predict the continuation of the volumetric data provided. We run a linear regression on the provided elevation data on the interval from 1999 to 2020 and find the relationship can be modelled by

$$e(t) = 11590.5 - 5.21072t$$

for elevation e and year t. **Under this model, we predict the elevation within Lake Mead in 2025, 2030, and 2050 to be 1038.79, 1012.73, and 908.52 ft, respectively.**

Strengths This model benefits from an incredibly simple relation that fits the provided drought period well, as reflected in fig. 4. It is not unreasonable to assume that this

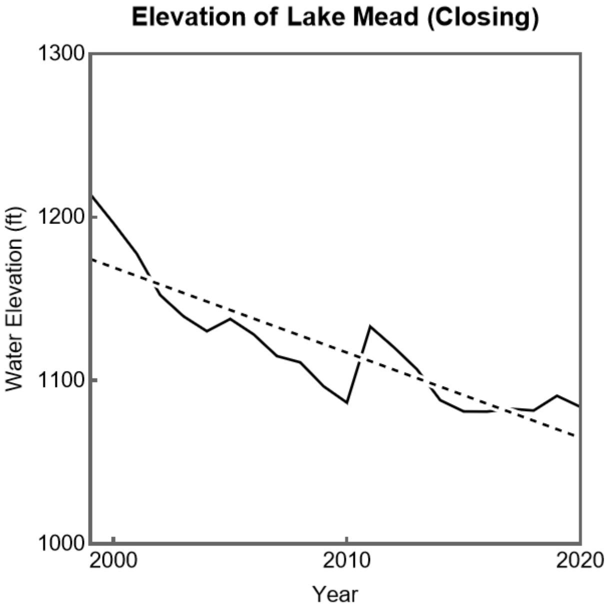


Figure 4: Collected elevation data during identified drought period 1999–2020, along with linear regression.

model will be accurate for years closer to 2020.

Weaknesses This method is computationally efficient but oversimplifies the problem at hand; it is not sensible to assume a linear relationship between elevation and time. While it may provide accurate predictions for proximate years, future predictions will be largely inaccurate. Furthermore, this model does not consider accommodations made by management to alleviate drought conditions, as well fluctuations in inflow and outflow at the lake.

3.2 Model 2

Our second model considers the interval from 2005-2020 only. We employ a Seasonal Autoregressive Integrated Moving Average (SARIMA) in conjunction with a rolling forecast algorithm as the basis of this alternate model. SARIMA models are used to make forecasts on time series and are parameterized by trend and seasonal autoregression, difference, and moving average orders of the given data. Parameters of the model were determined using the augmented Dickey-Fuller test and a grid search algorithm. Refer to Appendix A for a detailed explanation of the training of this model.

In comparison to its predecessor, ARIMA (Autoregressive Integrated Moving Average), SARIMA is better suited to handle seasonal trends. We find this to be appropriate for our problem given the yearly elevation spikes observed in the data. By formulating Lake Mead elevation as a time series, a SARIMA model can be used to predict future elevation values.

Furthermore, to improve the robustness of the model, we introduce a rolling forecast algorithm to iteratively make predictions. While the SARIMA model is capable of forecasting an entire range of values in one computational step, each value is based entirely on the training window and does not account for the predictions immediately adjacent

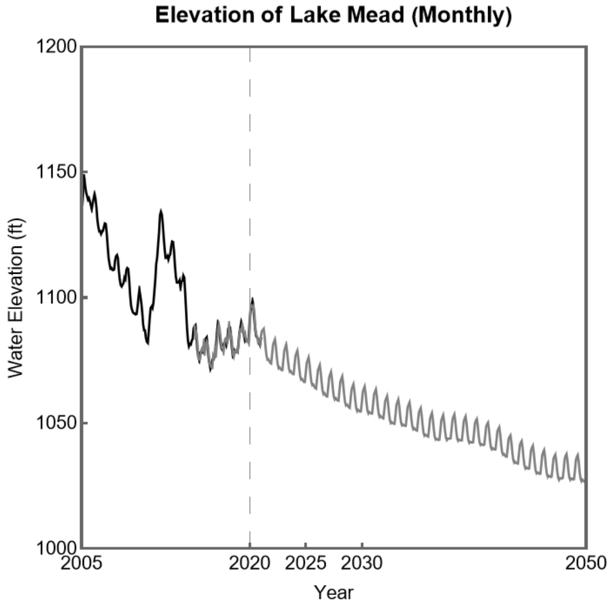


Figure 5: Monthly water elevation at Lake Mead plotted against SARIMA model forecast. Trained on data from 2005-2020, with initialization of rolling forecast at 2015.

to itself. This results in loss of accuracy over long periods of time. The rolling forecast algorithm prompts the SARIMA model to make a prediction at one time step in the future (1 month) and re-calibrates the model, adding the predicted value to the train set. Doing so improves the accuracy of the model and is a better estimate for real time.

The model predicts the average elevation in 2025, 2030, and 2050 to be 1068.71, 1057.71, and 1029.36 ft, respectively.

Strengths As seen in Figure 5, the SARIMA model effectively accounts for seasonal trends in the data; the yearly peaks observed in the data from 2005-2020 are mimicked in the model's forecast up to 2050. Accounting for seasonality allows for more precise forecasting, as the model provides reasonable predictions with monthly precision.

Weaknesses Like the first model, Model 2 is unable to predict complex environmental fluctuations; the trend of the forecast is still largely linear. However, given the nature of the data provided, this is to be expected. As shown earlier, Lake Mead elevation is a function of numerous environmental factors, each affecting elevation in different proportions. Without greater insight into these factors, any model for elevation will not yield reliable results.

IV Analysis

4.1 Improving Model Accuracy

As suggested within our Preliminary Analysis, we consider the Lake Mead Reservoir as a system, with various volumetric factors influencing the amount of water stored. We consider the following relation for the yearly change in volume of the reservoir:

$$\Delta V = I - O - L$$

Where I expresses the total inflow, O expresses the total outflow, and L expresses the total losses. This simple and intuitive model can allow for a modular approach where each volumetric factor can be considered independently and whose precision can be increased provided more data and or different modelling approaches.

4.1.1 Limitations of the Previous Approaches

As previously noted, the previous two models fail to account for fluctuations in the various volumetric factors we outline in our preliminary analysis. Here, we consider these factors and propose an improved model for predicting the elevation of Lake Mead in a given year. This model deals with the shortcomings of the previous models and can be improved to better fit the complex trend of Lake Mead's volume.

4.1.2 Inflow: Considering Climate Change

Experts agree that Lake Mead's current period of drought is greatly influenced by the rapidly changing climate [8]. Climate change caused increasing temperatures and decreasing precipitation and has subsequently decreased the average inflow into Lake Mead steadily over the course of the past decades. This trend will only amplify as the population grows and technology moves forward. In the period from 2011-2015 alone, the average daily temperature at Lake Mead had increased by 1.7 degrees C, from 21.0 C to 22.7 C.

An accurate model must consider the declining rate in temperature when considering fluctuations, particularly for inflow and precipitation. These factors are both decreased when temperatures increase, and thus are directly influenced by climate change (consider the decreased amount of melted snow flowing into the Colorado river as a result of higher temperatures). More concretely, the model considers flow f as a function of the average temperature T and precipitation P at time point t

$$f(t) = F(T(t), P(t))$$

4.1.3 Outflow: Characteristics of Consumption

While the Hoover Dam must continue to provide water to surrounding regions throughout its operation, the total released amount depends highly on seasonal demand and the current supply of water within the reservoir. During drought periods, a delicate balance forms between the increasing demand for water and an increase in stringent management policies designed to limit the lake from falling below the operational minimum/dead cap space.

While modelling annual releases of water can be accomplished with a time-series based approach, it is naive to reduce this complex balance to one based on this relationship. More accurate models should consider both the consumptive uses during this period, as well as the volume of the lake at that time. More concretely, a model could consider total outflow o as a function of the consumption C and the volume v, which can be modelled as a function of time:

$$o(t) = O(c(t) + v(t))$$

4.1.4 Losses: Evaluating Evaporation

While losses come in various forms, it is fair to claim that the primary source of loss will occur through evaporation at the surface of the reservoir. The rate of evaporation, considered largely constant in previous analyses, is in fact highly dependent on the surface area at any current time, as well as the temperature. Other factors such as wind speed may be considered.

We can model the fluctuating evaporation with the surface area and temperature. With e denoting the evaporation at any time, we note that it is a function of surface area S_a and temperature T, which are both modelled by the volume v and time t respectively:

$$e(t) = E(S_a(v(t)) + T(t))$$

4.2 General Model Construction

We thus find that an adequate model for volume could be constructed as follows:

$$\Delta v(t) = F(T(t), P(t)) - (O(c(t) + v(t)) + E(S_a(v(t)) + T(t)))$$

Since our fundamental goal was to model changes in volume as a function of time, we can form independent predictions for each function that is solely dependent on time itself. Each associated relationship, such as inflow F as a function of temperature T and precipitation P can be modelled independent of time with statistical analysis. The associated temperature and precipitation are predicted as a function of time.

This allows any arbitrary selection for the time-based approximations, such as linear regressions or the previously noted SARIMA models.

Strengths This model allows for more robust considerations in fluctuations of inflow, outflow, and evaporation on Lake Mead. By indirectly modelling change in volume as a function of time, this model escapes the fundamental issues in the previously noted flawed approaches. Additionally, its approach can be repeated to further parameterize factors such as temperature and precipitation. As more layers are added and complexity increases, we achieve a more realistic and interpretable model that is still fundamentally a function of time.

Weaknesses While this model successfully captures the complex relationship between Lake Mead's elevation and its associated factors, it suffers greatly from accumulation of error. Since each component is modeled independently, the inaccuracies in each model may sum up to result in an inaccurate forecast. Given the performance of current time series forecasting models, this weakness is inevitable. However, we expect that it will improve exponentially with the development of more concrete forecasting models, making this model the most scalable approach.

As a result, while we expect that increasing the depth of the expression may improve accuracy for proximal years, the model's accuracy will falter under larger time intervals. This is inherent to any model that attempts to predict the future: time is not the only independent variable.

V Future Outlook

5.1 Recycling as a solution

Based on all of the proposed models, Lake Mead is clearly under high risk of future drought that will impact the region immensely. While inflow and losses are largely out of the control of any plans of action, minimizing the outflow by repurposing consumptive uses appears to be a viable method of alleviating the present issues. We consider the recycling of waste water from such consumptive uses in order to aid with such effort.

5.1.1 Recycling Outflow

Elements of water reuse include: screening, pumping, aerating, sludge digestion, scum removal, disinfection, odor control [9]. The renewed water, known as effluent, then exits the treatment facility to local waterways or to fit-for-purpose plants to be further processed using primary, secondary, tertiary, and advanced treatment techniques.

Factors that influence a large-scale water recycling plan include:

- Environmental and Operational Factors
  - Management Framework and Storage Practices
  - Water Quality: Polluted vs. treated; potable vs. non-potable and other public health implications
  - Upstream and Downstream Environmental Conditions
- Government Factors
  - Regulations and Policy Coordination: Coordinate federal, state, tribal, and local water reuse programs and Riparian rights
  - Financial Support and Cost-Efficiency
  - Additional Infrastructure
  - Domestic Use
- Business Interests
  - Treatment Facilities: Meet fit-for-purpose requirements
  - Agricultural Use
  - Industrial and Commercial Use: Private sector

5.1.2 Proposed Plan

The first priority that policy makers should make is to ensure that there is enough water to meet the local survival need. In addition, they should strive to maintain compliance with the Clean Water Act (CWA), which regulates the discharge of certain categories of pollutants in waterways. Hence, another priority is to agree upon an analytic method of control: enforceable numeric limits on the release of toxins near Lake Mead would depend on the availability of specific technologies to treat corresponding wastes.

Evidently, it is the responsibility of public sector organizations to spearhead such a plan, which requires detailed planning, conference, and reevaluation between policymakers and community stakeholders to mitigate drought conditions in a timely fashion. We reiterate policy priorities for managing drought:

1. Municipal Water Supply and Rural Groundwater Security
    - Water Price: generate a tiered rate structure within reason to increase revenue while limiting sales
    - Urban Landscape Irrigation: introduce financial incentives for efficient water use, e.g. for watering lawns and gardens
2. Positive Trends in Reservoir Water Levels
    - Water Recycling: see below
    - Model Reevaluation: incorporate new factors into existing reservoir models as conditions change
3. Efficiency in Agricultural Irrigation Systems
4. Sustainable Industrial Water Use Practices
    - Clean Water Act: enforce compliance with existing federal and state environmental mandates
5. Ecosystem Stability

We consider outflow as the primary target of any water recycling plan. More specifically, we consider total consumptive uses from Lake Mead, as those most accurately reflect impacts on the water supply on the lake itself. We localize our changes to Lake Mead to remain consistent with our previous analysis.

We aim to minimize the consistent decreases in water volume at Lake Mead by reverting as much of the consumptive uses as possible. This provides us a concrete metric to measure the success of our plan and provide insight into future modifications that may aid with alleviating the drought problem.

From fig. 6 we see that the majority of consumptive use withdrawals in the Lower Basin are from the irrigation and public categories. We reason that, since Lake Mead provides the Lower Basin with much of the water supply (as well as being central to the water flow), its relative consumptive uses model a similar distribution.

As a first step to integrate recycling into the drought management plan, since water used for irrigation does not generate treatable wastewater, we focus our attention on reducing waste generated by public supply and domestic use. The core concept of this

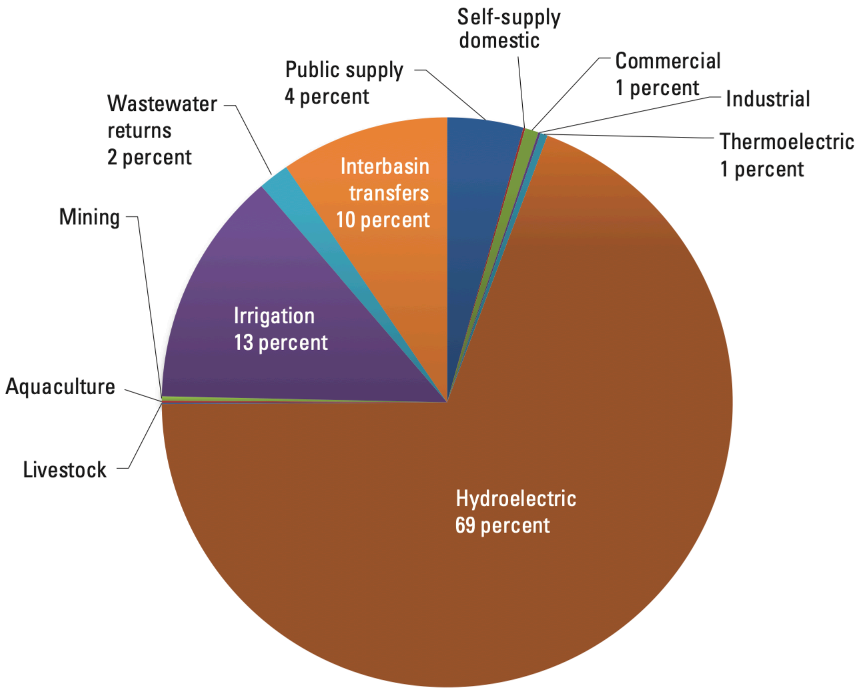


Figure 6: Percentage of total estimated water withdrawals by category, hydroelectric power use, and wastewater returns for the Lower Colorado River Basin, Southwestern United States, 2010. [5]

phase of our water reclamation sketch is to *treat* wastewater generated from potable domestic use, and industrial use if possible, to *fit* non-potable purposes, namely agricultural and industrial.

This proves an effort to slow the rate at which food production for the growing population depletes freshwater resources. Moreover, the wastewater only needs to undergo primary and secondary treatments in order to be serviceable for irrigation. This treatment process is less complex and therefore less expensive than for advanced potable purposes. Once we see a degree of agricultural stability, measured by stress factors and gross domestic product, and suitable improvements in reservoir water levels, we may continue to the next step explained below.

5.1.3 Expected Results

For the second phase of our sketch, we first consider a *highly optimistic* recycling plan, which maximizes the total available water that is recycled. In other words, we assume all of the consumptive uses that are able to be recycled (Self-supply domestic, commercial, industrial, public supply) will be returned as surface water or for reuse. Therefore, in the best case scenario, there should be an additional 5% wastewater return for a total of 7% returned. Success indicators include continued positive trends in reservoir water levels with respect to water demand, diminution of toxins present in surface water, and consumer approval.

We propose several avenues of diversification that, along with recycling, could ameliorate the water shortage:

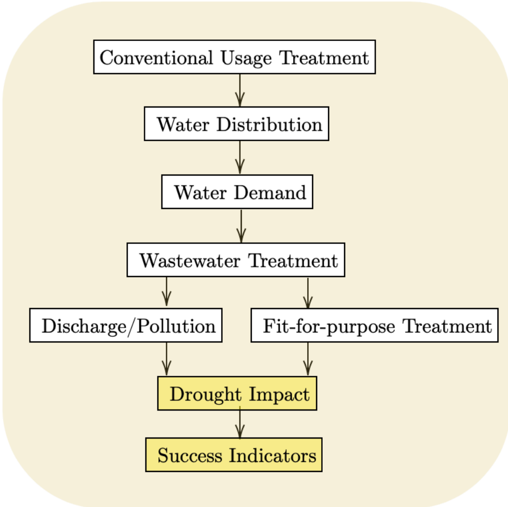


Figure 7: A simplified system for reducing consumptive water use.

- Optimization of Treatment Facility Operations
- Optimization of Groundwater Resource Allocation
- Stormwater Harvesting
- Metal-Organic Frameworks
- Desalination of Saltwater

While we concretely achieve around 7% reduction in consumptive uses, we can further utilize our time-series models to predict the total amount conserved. This can be achieved quite intuitively with our previous model structure. Total consumption can be calculated by subtracting the total inflow and other outflow causes from the net change in volume.

We show in our Preliminary Analysis, as well as in our models, that both of these factors are able to be estimated using extensive data collected by the Bureau of Reclamation. As a sample, we consider 2015 as a case study for our recycling plan. Lake Mead was reported to have had 10.087 MAF of water at the end of December 2015. This was a net decrease of 0.580 MAF from the previous year [10]. Along with this change, there was a net inflow of 13.437 MAF with 9.414 MAF released through the Hoover Dam. Without considering other factors, such as losses and precipitation, this would account for around 4.603 MAF of total consumption. With our 7% return strategy, this would send around 0.322 MAF of water back into Lake Mead– a 55% decrease in net volumetric losses.

VI Conclusion

In this paper, we developed two models to predict water elevation in Lake Mead 2020-2050. The models describe factors influencing inflow, outflow, and loss as functions of time. We began our exploration with a basic linear regression model that predicts water elevation well in the coming decade. We then progressed to a stronger model with the application of time-series analysis. In particular, the SARIMAX model accounted for the periodicity of elevation in Lake Mead and demonstrated the steady decline that is to come.

To relate our model back to the central problem, the drought in Lake Mead, we proposed political priorities and a water recycling plan to manage consumption. We posit that our plan, if coupled with various other conservation initiatives, could restore stable reservoir levels. In addition, we believe that in-depth analysis of patterns in population and business would give greater insight into proper policy and allocation of government funding. Another factor worth considering is local groundwater resources. These were disregarded in the current model since their extraction involve operations that are largely separate from Colorado River flow and Lake Mead elevation, but their inclusion would make for a more realistic model when examining water reuse.

VII Article

Recycling to Refill: Addressing the Drought Crisis in Lake Mead, Nevada

Lake Mead

Located on the western side of the United States, the Lake Mead reservoir is responsible for providing water to nearly 30 million Americans in the region. This makes it one of the most important infrastructural centers in the country, as its provision of usable water is extensively utilized.

Despite its size and influence in the region, Lake Mead has found itself particularly susceptible to recent regional fluctuations that have left its water levels at dangerous lows.

Early Warning Signs

Over the past two decades, Lake Mead has experienced steadily decreasing water elevations, which can be attributed to a slow decline in flowing water in the Colorado River that supplies it with its water supply. Climate change has been identified as the primary culprit for this decline, and the bustling population of the western United States has only placed more pressure to deliver water from the reservoir.

It is clear that with the rise of technology and fossil fuels there is less water flowing into Lake Mead and more water being taken out to feed the growing population. While there is still time, the population must come together and urge policymakers to enact change– new and novel solutions must be considered in order to preserve this important center of American infrastructure.

Tackling the Drought

While various models have shown the steady decline in future water levels, it is through change that the decrease can be corrected. In order to deal with the looming threats on the region, local governments must consider the recycling of wastewater as a potential solution. While much of the water taken out of the Lake Mead Reservoir is not used directly by consumers, recycling provides a concrete method of placing water back into the reservoir.

Of the consumed water being removed from Lake Mead annually, about 7% is capable of being recycled by means of traditional treatment and fitting. Building sufficient infrastructure to treat as much of this water as possible should be emphasized– after all, the monetary impacts of inaction are far greater than those of change.

By successfully harnessing recycling and returning this 7% of consumed water, Lake Mead can be placed on the right track to re-stabilizing. It turns out that this 7% can account for massive (sometimes even majority) of total losses of water. In other words, while 7% may seem insignificant, it can turn annual net losses of water levels into net gains! Clearly, there is hope for a safer, more aqueous future in the arid western United States, and it starts with recycling.

References

[1] Tim P. Barnett and David W. Pierce. "When will Lake Mead go dry?: WHEN WILL LAKE MEAD GO DRY?" In: *Water Resources Research* 44.3 (Mar. 2008). ISSN: 00431397. DOI: 10.1029/2007WR006704. URL: http://doi.wiley.com/10.1029/2007WR006704 (visited on 11/11/2021).

[2] "Water Futures: Assessment of Long-range Patterns and Problems". In: (Oct. 30, 2008). URL: https://www.sei.org/publications/water-futures-assessment-long-range-patterns-problems-2/ (visited on 11/13/2021).

[3] Michael T Moreo. *Evaporation Data from Lakes Mead and Mohave, Nevada and Arizona, March 2010 through April 2015*. Type: dataset. 2015. DOI: 10.5066/F79C6VG3. URL: https://www.sciencebase.gov/catalog/item/55f6fba8e4b0477df11bff2b (visited on 11/12/2021).

[4] Michael T. Moreo. *Meteorological Data for Lake Mead and Lake Mohave, Nevada and Arizona, April 2013 to April 2017*. Type: dataset. 2018. DOI: 10.5066/F7G44PJ9. URL: https://www.sciencebase.gov/catalog/item/5a74e8c1e4b00f54eb1c82df (visited on 11/11/2021).

[5] *Scientific Investigations Report*. Scientific Investigations Report. Series: Scientific Investigations Report. 2018.

[6] *Scientific Investigations Report*. Scientific Investigations Report. Series: Scientific Investigations Report. 2020.

[7] *Scientific Investigations Report*. Scientific Investigations Report. Series: Scientific Investigations Report. 2020.

[8] Lauren Sommer. "The Drought In The Western U.S. Is Getting Bad. Climate Change Is Making It Worse". In: *NPR* (June 9, 2021). URL: https://www.npr.org/2021/06/09/1003424717/the-drought-in-the-western-u-s-is-getting-bad-climate-change-is-making-it-worse (visited on 11/12/2021).

[9] *A Visit to a Wastewater Treatment Plant*. URL: https://www.usgs.gov/special-topic/water-science-school/science/a-visit-a-wastewater-treatment-plant?qt-science_center_objects=0#qt-science_center_objects (visited on 11/14/2021).

[10] *Boulder Canyon Operations Office | Lower Colorado Region | Bureau of Reclamation*. URL: https://www.usbr.gov/lc/region/g4000/wtracct.html (visited on 11/11/2021).

[11] *Boulder Canyon Operations Office | Lower Colorado Region | Bureau of Reclamation*. URL: https://www.usbr.gov/lc/region/g4000/NaturalFlow/ (visited on 11/11/2021).

[12] *Bureau of Reclamation: Lower Colorado Region - Colorado River Water Uses Since 1906*. URL: https://www.usbr.gov/lc/region/g4000/uses.html (visited on 11/11/2021).

[13] Mailing Address: 601 Nevada Way Boulder City and NV 89005 Phone: 702 293-8990 Contact Us. *Storage Capacity of Lake Mead - Lake Mead National Recreation Area (U.S. National Park Service)*. URL: https://www.nps.gov/lake/learn/nature/storage-capacity-of-lake-mead.htm (visited on 11/12/2021).

[14] *Overview of Lake Mead - Lake Mead National Recreation Area (U.S. National Park Service)*. URL: https://www.nps.gov/lake/learn/nature/overview-of-lake-mead.htm (visited on 11/11/2021).

[15] *Programs and Activities | Lower Colorado Region | Bureau of Reclamation.* URL: https://www.usbr.gov/lc/region/programs/strategies/FEIS/index.html (visited on 11/11/2021).

[16] *Trends in recent historical and projected climate data for the Colorado River Basin and potential effects on groundwater availability.* URL: https://pubs.er.usgs.gov/publication/sir20205107 (visited on 11/11/2021).

Appendices

A Formulation of the SARIMA Model

The SARIMA model and all mentioned statistical analysis tools were implemented using the Python module statsmodels https://www.statsmodels.org/stable/index.html.

The SARIMA model can be expressed in the following notation:

$$SARIMA(p, d, q)(P, D, Q)m$$

where p, d, and q are the autoregression, difference, and moving average orders of the time series trend, respectively; P, D, and Q are the seasonal autoregression, difference, and moving average orders; and m is the length of a single period. The algorithm itself is rather complex to include but can be further explored here: https://online.stat.psu.edu/stat510/lesson/4/4.1.

The objective of formulating a SARIMA model is to determine optimal parameters for p, d, q, P, D, Q, and m:

At first glance, m can be defaulted to 12, as we are considering monthly data with a yearly cycle.

$$m = 12$$

D and d, the integration orders, are determined with the augmented Dickey-Fuller test for stationarity. Running Dickey-Fuller on the elevation data yields a p-value of 0.386, which is not enough to reject the null hypothesis. By taking the first difference of the series and running the test again, we obtain a p-value of 0.004: sufficient to accept the series as stationary.

$$d = D = 1$$

P and Q are traditionally determined via the partial autocorrelation and autocorrelation plots of the time series, respectively. However, as shown in Figure 8, there is no surface-level interpretation of the plots. Instead, we employ a grid search algorithm to determine optimal values of p, P, q, and Q. Each parameter was assigned a value ranging from 0-3, and every possible configuration of values was evaluated using Akaike information criterion as the loss estimator. At an AIC score of 679.85, the following optimal parameters were determined:

$$p = 2, P = 1, q = 0, Q = 1$$

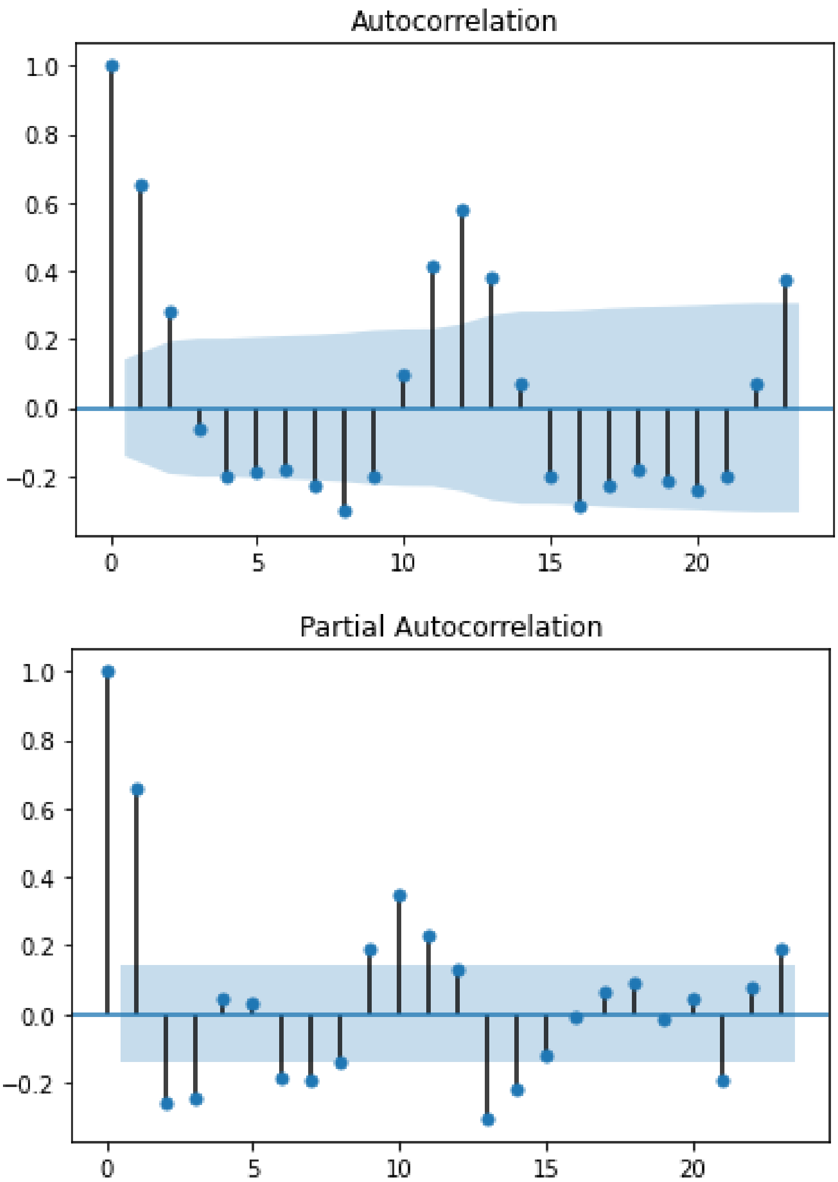


Figure 8: ACF and PACF plots of the first-differenced data. There does not appear to be a surface-level interpretation of these plots to determine p and q.

Therefore, we use SARIMA(2, 1, 0)(1, 1, 1)12 as our model to forecast Lake Mead elevation.

B SDI Drought Index Values at Lake Mead

See this link for the table of the SDIs at Lake Mead 1936-2019.
`https://docs.google.com/spreadsheets/d/1tnPym4MxL-WUew7thkEY1JdOwsvbXpIFzPA8BTygRuI/edit?usp=sharing`

C SARIMA Model Code

This code utilizes the provided elevation data.

```python
import statsmodels.api as sm

split_date = '2015-01-01'
train = data['Elevation'].loc[:split_date]
test = data['Elevation'].loc[split_date:]

def rolling_forecast(train, test, order, season):
    history = [x for x in train]
    model = sm.tsa.statespace.SARIMAX(history, order=order, seasonal_order=season)
    model_fit = model.fit(disp=False)
    predictions = []
    yhat = model_fit.forecast()[0]
    predictions.append(yhat)
    history.append(test[0])

    for i in range(1, index_to_num('2050-12-01')-index_to_num('2015-01-01')+1):
        model = sm.tsa.statespace.SARIMAX(history, order=order, seasonal_order=season)
        model_fit = model.fit(disp=False)
        yhat = model_fit.forecast()[0]
        predictions.append(yhat)
        try:
          obs = test[i]
          history.append(obs)
        except:
          history.append(yhat)

    predictions = pd.Series(predictions, index=generate_index(2015, 2051))
    return predictions

rolling_fcast = rolling_forecast(train, test, (2, 1, 0), (1, 1, 1, 12))
predictions

def index_to_num(index):
  y = index[:4]
  m = index[5:7]
  return (int(y)-1936)*12 + int(m)


def generate_index(start, end):
  indices = []
  for i in range(start, end):
    for month in ["JAN", "FEB", "MAR", "APR", "MAY", "JUN", "JUL", "AUG", "SEP", "OCT", "NOV", "DEC"]:
      indices.append(str(i) + '-' + month)
  return pd.DatetimeIndex(indices)
```