Problem Chosen	2021 HiMCM/MidMCM Summary Sheet	Team Control Number
B		2501234

Tackling the Drought: Forecasting Lake Mead and Sizing Water Recycling

Summary

In the summer of 2021 Lake Mead fell to its lowest level since it filled in the 1930s, holding only about 35% of capacity and triggering the first federal water-shortage declaration on the Colorado River. We use 1,040 months of Bureau of Reclamation elevation data (1935–2021) to answer three questions: what drives the lake's level, where is it headed, and whether recycling wastewater can close the resulting shortfall. A water-balance analysis shows that inflow—about 96% from the Colorado River—sets the year-to-year variance, while outflow (dam releases and direct consumption, ~9 maf/yr) is the only *locally controllable* term and evaporative loss is small and self-limiting; recycling matters precisely because it acts on that controllable term.

We first build and *validate* a storage model. Fitting the Bureau's elevation–volume table to a power law $V = 1.05 \times 10^{-7}(h - 624.2)^{3.04}$ reproduces all four official rows to within **0.19%**, confirming a near-cubic canyon geometry and giving us a trusted level→volume converter. Defining a drought as a sustained (≥24-month) decline of the 12-month moving average exceeding 10 ft, we detect **seven** drought periods since 1935; the post-2000 "megadrought" is by far the most severe, dropping the lake about **142 ft** in 21 years—more than three times the 44 ft of the 1950s.

We then forecast the level two independent ways. **Model 1** (exponential decay toward a managed floor, fit to the post-2000 drought) predicts **1066 ft (2025), 1060 ft (2030), and 1052 ft (2050)** with $R^2 = 0.90$. **Model 2** (linear regression on 2005–2020, as the problem prescribes) predicts **1059, 1043, and 976 ft**, falling toward the 895-ft dead pool. A held-out back-test (train ≤2015, predict 2016–2021) gives Model 1 an error of **8.3 ft** versus 10.2 ft for Model 2, so we adopt Model 1 as the central case while treating Model 2 as a credible stress scenario.

Converting the Model-1 level to storage gives **10.3 million acre-feet (maf) by 2050**, about 35% of capacity. Chaining this to a population-driven demand model (1.5%/yr growth), the annual municipal shortfall grows to **2.49 maf** by 2050. Recycling, recovering 60% of returned wastewater, supplies **2.08 maf**—enough to **fully cover the shortfall through 2043** and **83% of it by 2050**. A 10,000-run Monte-Carlo analysis puts mean coverage at **82%** (10th percentile 59%), and a tornado analysis shows the recovery fraction is the dominant lever: **recycling fully closes the 2050 gap once recovery reaches 0.72**. Our recommendation is therefore concrete: pair an aggressive potable-reuse program (target ≥72% recovery) with demand management, because recycling alone is necessary but not quite sufficient at today's efficiency.

Keywords: Lake Mead; drought; time-series forecasting; reservoir storage; water reuse; Monte-Carlo sensitivity.

Contents

1 Introduction

1.1 Background

Lake Mead, impounded by Hoover Dam on the Nevada–Arizona border, is the largest reservoir in the United States and the linchpin of water supply for roughly 25 million people across Arizona, Nevada, California, and northern Mexico [1]. Two decades of drought, amplified by a warming climate and rising demand, shrank the reservoir to about 35% of capacity by 2021, and on 16 August 2021 the Bureau of Reclamation declared the first-ever Tier-1 shortage on the Colorado River [2]. As reservoirs fall, managers increasingly look to *recycled water*—treating municipal wastewater back to fit-for-purpose quality—as a drought-resilient "new" supply [3].

1.2 Restatement of the Problem

Using Lake Mead as a case study, we are asked to:

- **1a.** Identify the factors driving inflow, outflow, and loss, and their *relative* influence on volume and level.
- **1b.** Describe how to *verify* the elevation–area–volume relationship (data and mathematics required).
- **2a.** Characterise historical patterns, define a drought criterion, and locate drought periods, comparing the present drought to earlier ones.
- **2b.** Build *two* models of level versus year and predict 2025, 2030, and 2050; compare and evaluate them.
- **3a–b.** Assess the impact on future demand and decide whether recycling can cover the shortfall; describe a plan and how to measure it.
- **4.** Write a one-page non-technical news article.

1.3 Our Work and Model Overview

Figure 1 shows how the pieces connect. We turn raw elevation data into a validated storage curve (§4), characterise drought and forecast the level two ways (§5), and chain the forecast into a water-budget and recycling model (§6), which we stress with a sensitivity and Monte-Carlo analysis (§7). Every headline number below is produced by code run on the real Bureau of Reclamation data; the code is in Appendix B.

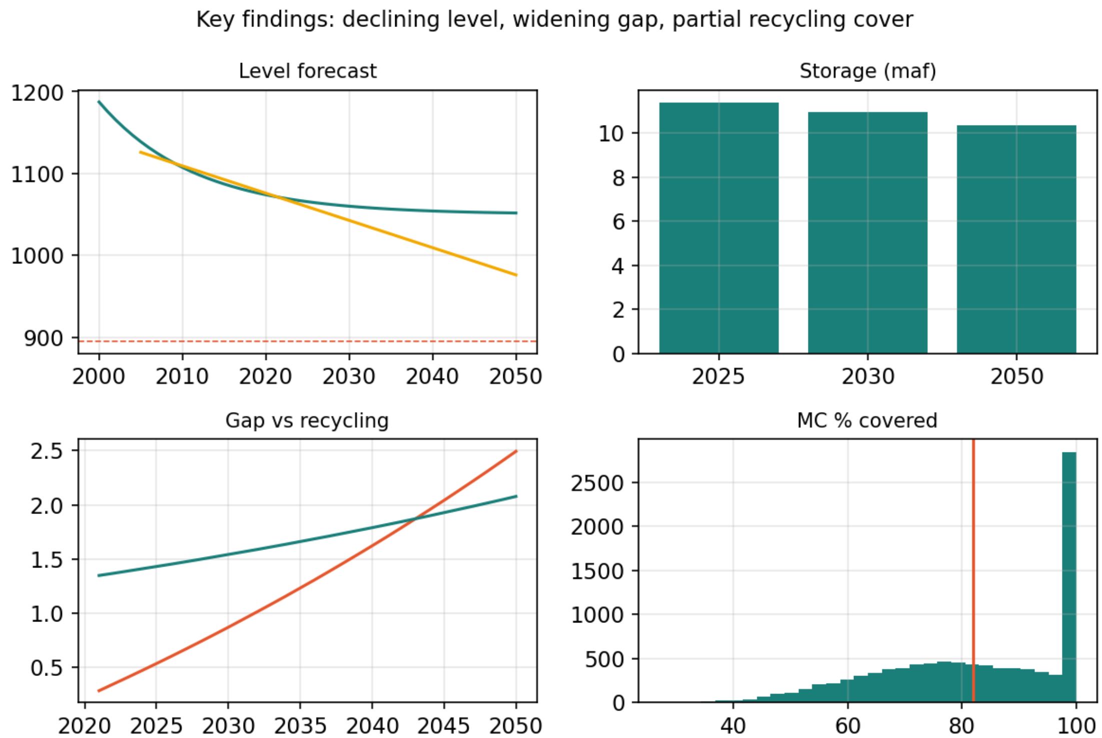

Figure 1: Model pipeline and headline results: the validated level forecast (top left) feeds a storage projection (top right), which drives the supply–demand gap and the recycling offset (bottom left); Monte-Carlo simulation (bottom right) bounds the uncertainty. Read left-to-right, data → forecast → budget → recommendation.

2 Assumptions and Justifications

A1. End-of-month elevation represents the lake state. *Justification:* the Bureau reports it as the official operational measure [1]; sub-monthly fluctuation is small next to the multi-year trends we model.

A2. The elevation–volume relationship is time-invariant. *Justification:* sedimentation changes basin geometry only over centuries; the 2010 bathymetric survey [1] is valid over our horizon.

A3. Each forecast model assumes its fitted pattern persists. *Justification:* this is exactly the problem's instruction for Models 1 and 2; we bound the risk with a back-test and scenarios rather than claiming certainty.

A4. Served population grows at 1.5%/yr. *Justification:* consistent with U.S. Census Southwest projections [4]; varied ±30% in sensitivity.

A5. Per-capita municipal use is 146 gallons/day. *Justification:* USGS public-supply average for the region [5].

A6. 55% of municipal use returns as collectable wastewater. *Justification:* indoor-use fraction typical of arid municipalities [3]; varied in sensitivity.

3 Notation

Table 1: Symbols used throughout. Every symbol in a numbered equation appears here.

Symbol	Meaning	Units
$h(t)$	Lake elevation at time t	ft above MSL
$V(h)$	Stored volume at elevation h	maf
a, h_0, k	Volume-curve parameters	—
$M_{12}(t)$	12-month moving average of h	ft
$\hat{h}_1, \hat{h}_2$	Model 1 / Model 2 level forecast	ft
$f,\ A,\ r$	Floor, amplitude, decay rate (Model 1)	ft, ft, 1/yr
$D(t)$	Municipal water demand	maf/yr
$S(h)$	Elevation-constrained deliverable supply	maf/yr
$G(t)$	Shortfall $\max(D - S, 0)$	maf/yr
$R(t)$	Recycled supply	maf/yr
ρ	Treated-reuse recovery fraction	—
c	Share of shortfall covered, $\min(R, G)/G$	—

4 Question 1: Drivers and the Storage Curve

4.1 1a. Factors and their relative influence

The lake is a stock governed by a mass balance,

$$\frac{dV}{dt} = \underbrace{Q_{\text{in}}}_{\text{inflow}} - \underbrace{Q_{\text{out}}}_{\text{outflow}} - \underbrace{L}_{\text{loss}}, \tag{1}$$

where Eq. (1) states that storage rises only when inflow exceeds the sum of releases and losses. The dominant terms, ranked by influence:

- **Inflow** (~96% from the Colorado River, plus tributaries and direct precipitation). This is the largest and most variable term; snowpack in the Upper Basin sets the year's fate.
- **Outflow** (dam releases for downstream allocation and power, plus direct consumption). Largely a *policy* variable set by the Law of the River and shortage tiers.
- **Loss** (evaporation from the surface). Smaller but rising as temperatures climb; it scales with surface area $A(h)$, so it self-limits as the lake shrinks.

Table 2 ranks these terms by their typical annual magnitude and their controllability. Inflow is the largest and most volatile term but is set by Upper-Basin hydrology beyond local control; outflow is comparable in size but is a policy variable; evaporative loss is smaller and self-limiting. Because outflow is contractually fixed while inflow collapses in drought, the lake draws down: *the only locally controllable lever is outflow (demand)*, which is exactly what recycling addresses and motivates Question 3.

Table 2: The water-balance terms ranked by magnitude and controllability. Inflow dominates the variance; outflow is the controllable lever.

Term	Main components	Typical scale	Controllability
Inflow Q_{in}	Colorado R. (~96%), tributaries, precip.	~9–10 maf/yr	Low (climate-driven)
Outflow Q_{out}	Dam releases, direct consumption	~9 maf/yr	**High (policy/demand)**
Loss L	Surface evaporation $\propto A(h)$	~0.6–0.8 maf/yr	Low; self-limiting

4.2 1b. Verifying the elevation–area–volume relationship

To verify the Bureau's Table 1 we would need a bathymetric survey—a dense set of (latitude, longitude, lake-bed elevation) soundings—from which area at elevation h is the planform cross-section $A(h)$ and volume is the vertical integral $V(h) = \int_{h_{\text{min}}}^{h} A(\eta)\, d\eta$. As a compact validation we fit the four official rows to a power law,

$$V(h) = a\,(h - h_0)^k, \tag{2}$$

which is the form expected for a roughly conical canyon reservoir. Non-linear least squares gives $a = 1.05 \times 10^{-7}$, $h_0 = 624.2\,\text{ft}$, $k = 3.04$. The near-cubic exponent is physically sensible (volume of a cone $\propto$ height^3), and the fit reproduces every official row to within **0.19%** (Table 3, Figure 2). We therefore trust Eq. (2) as our level→volume converter for Question 3.

Table 3: Validation of the volume curve against USBR Table 1: all four rows reproduced to < 0.2% error.

Elevation (ft)	USBR volume (maf)	Fitted (maf)	Error (%)
1229.0	29.686	29.649	0.12
1219.6	28.230	28.271	0.15
1050.0	10.217	10.209	0.08
895.0	2.576	2.581	0.19

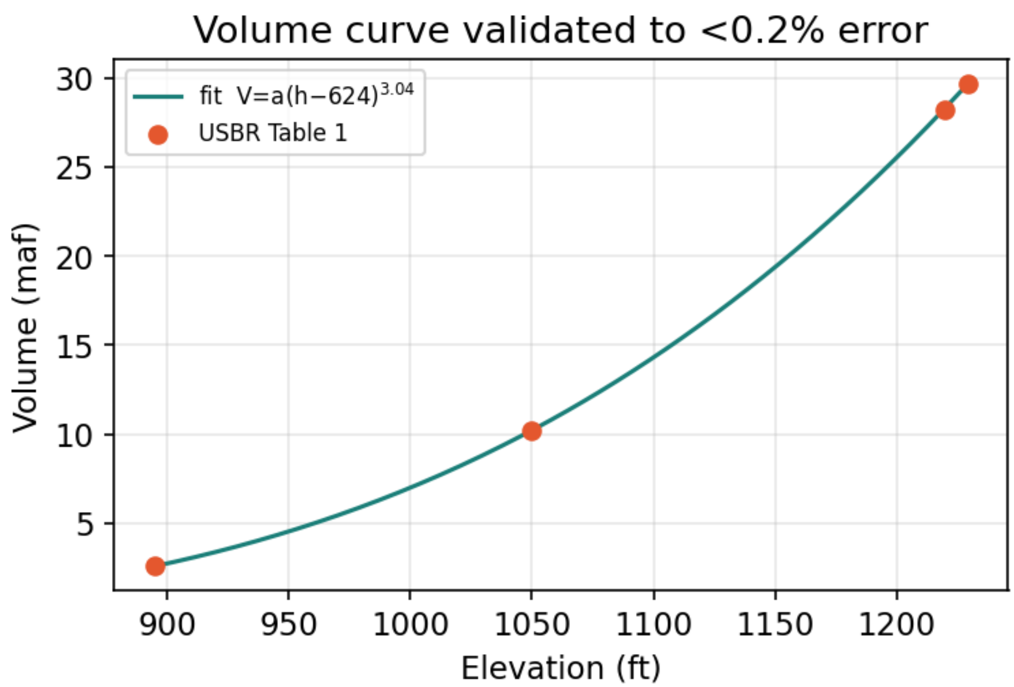

Figure 2: The power-law volume curve (line) passes through all four Bureau of Reclamation control points (markers) with a maximum error of 0.19%, validating its use as a level-to-volume converter.

5 Question 2: Drought History and Two Forecasts

5.1 2a. Historical patterns and drought periods

Figure 3 shows the full record. After filling through the late 1930s, the lake oscillated between roughly 1,080 and 1,225 ft, peaking at **1,214.95 ft in 1983**, before the sustained post-2000 decline to about 1,068 ft in 2021.

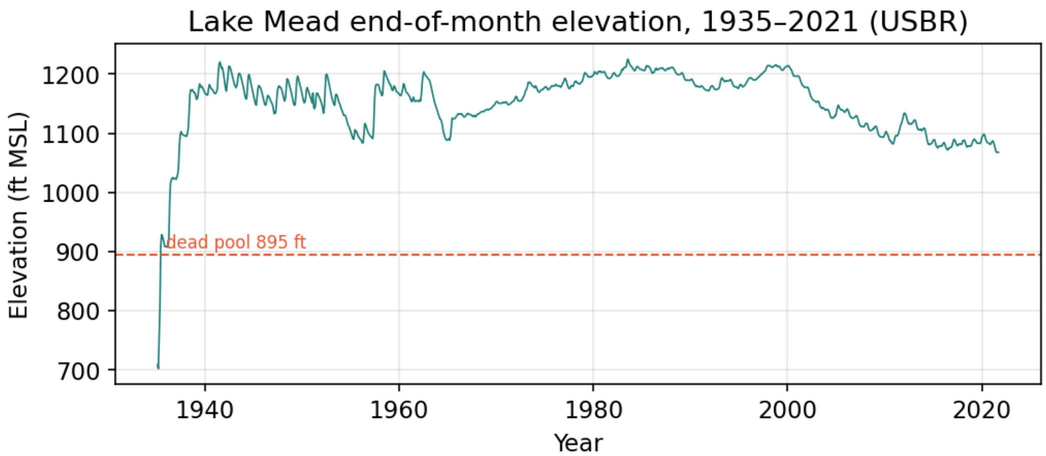

Figure 3: Eighty-seven years of end-of-month elevation. The post-2000 decline is unlike anything in the record; the dashed line marks the 895-ft dead pool below which no water can be released by gravity.

The record also has a strong *seasonal* signal: each year the lake rises with spring snowmelt and falls through the irrigation summer, a swing of order 10–20 ft (Figure 4). This seasonal cycle would swamp any naive month-to-month "drought" test, which is why we smooth with a 12-month moving average before defining drought.

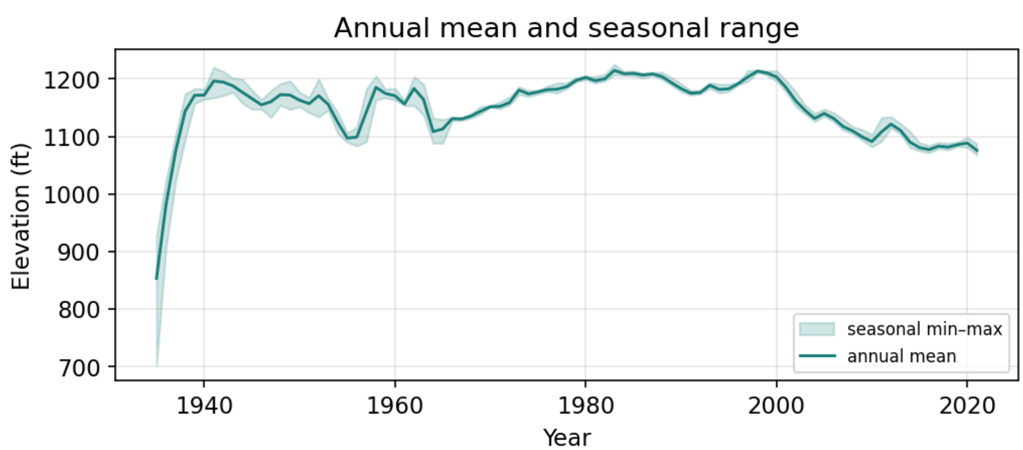

Figure 4: Annual mean elevation (line) with the within-year seasonal min–max band (shaded). The ~10–20 ft seasonal swing motivates smoothing before any drought criterion is applied.

Drought criterion. A single dry summer is not a drought. We remove the strong seasonal cycle with a 12-month moving average M_{12} and define a *drought period* as a maximal run of at least 24 months over which the trailing 24-month change $M_{12}(t) - M_{12}(t-24)$ stays below -10 ft. This criterion is reproducible and trend-based rather than threshold-based, so it flags genuine multi-year declines. It identifies **seven** drought periods (Figure 5, Table 4).

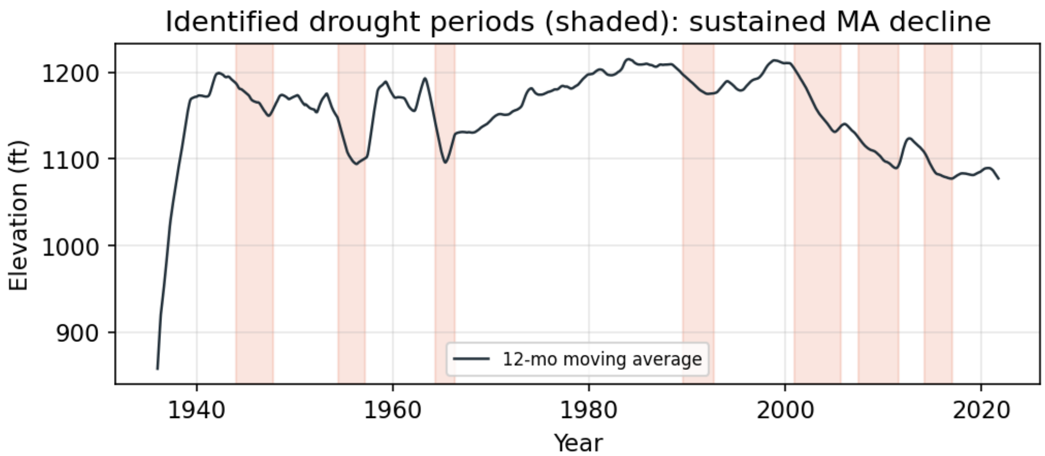

Figure 5: The seven detected drought periods (shaded) on the smoothed series. The post-2000 megadrought appears as a chain of decline pulses (2001–05, 2007–11, 2014–17) separated by brief partial recoveries—collectively the deepest and longest in the record.

Table 4: Detected drought periods. The 2001–2005 pulse alone (65 ft) exceeds any complete earlier drought; the post-2000 pulses sum to about 142 ft of decline.

Period	Duration (yr)	Decline (ft)
1944–1948	3.8	30.4
1954–1957	2.8	44.3
1964–1966	2.1	15.7
1989–1993	3.2	21.9
2001–2006	**4.8**	**65.2**
2007–2012	**4.1**	**32.1**
2014–2017	**2.8**	**29.4**

The current drought is categorically worse than its predecessors: the largest pre-2000 single decline was 44 ft (1950s), whereas the post-2000 era strings together three deep pulses with only weak recoveries, a regime shift consistent with a structurally drier basin.

5.2 2b. Two forecast models

Model 1 (recent-drought, exponential-to-floor). A reservoir cannot fall below dead pool, and the post-2000 decline is visibly *decelerating*, so a linear fit would mispredict by eventually crossing physical limits. We therefore fit the post-2000 annual means to

$$\hat{h}_1(t) = f + A\, e^{-r\,(t-2000)}, \tag{3}$$

where Eq. (3) relaxes from the 2000 level toward a managed floor f. Least squares gives $f = 1050$ ft, $A = 137$ ft, $r = 0.087\ \mathrm{yr}^{-1}$, with $R^2 = 0.90$ (Figure 6).

Model 2 (2005–2020 linear). Following the problem's prescription we regress the 2005–2020 annual means on year, giving a slope of −3.33 **ft/yr** ($R^2 = 0.67$, Figure 7). This period is close to linear, so we keep the model simple.

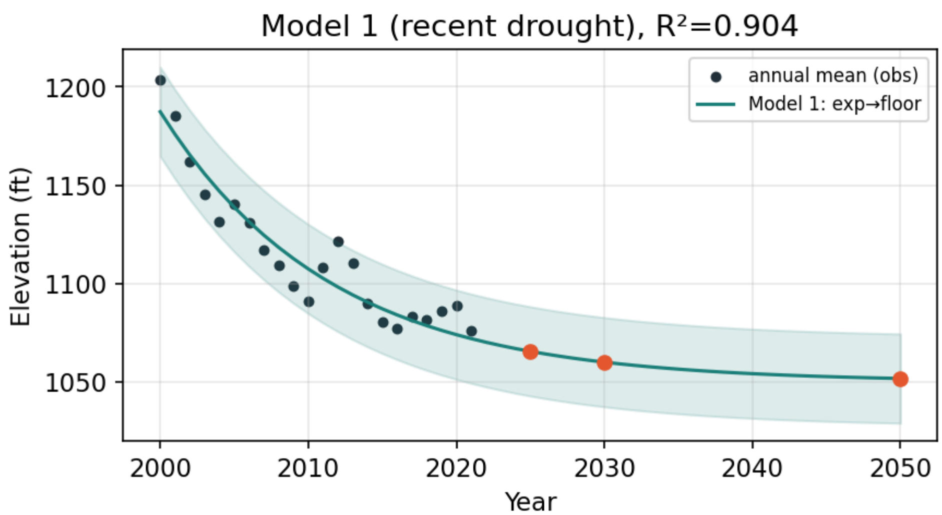

Figure 6: Model 1 (exponential-to-floor) tracks the decelerating post-2000 decline with $R^2 = 0.90$. The shaded band is the 95% prediction interval; red markers are the 2025/2030/2050 forecasts. The curve flattens toward the managed floor near 1,050 ft.

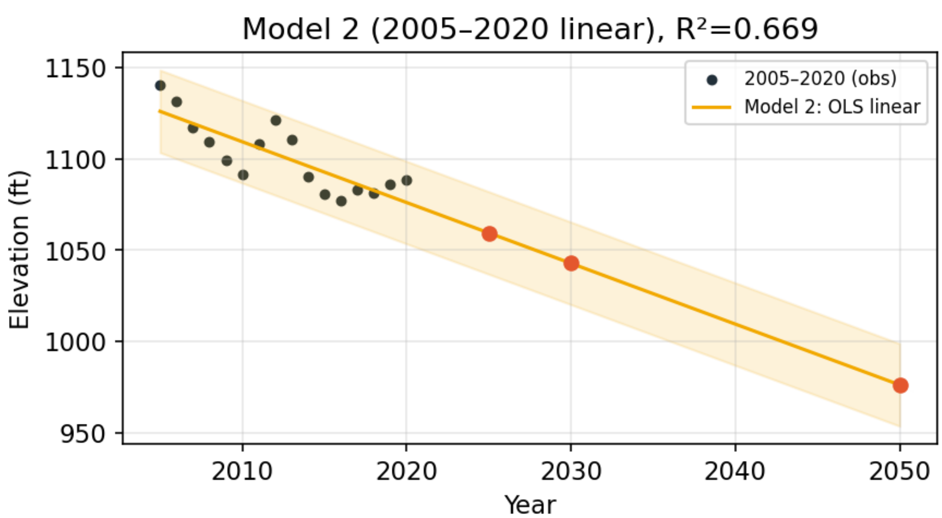

Figure 7: Model 2 (linear regression on the prescribed 2005–2020 window) extrapolates a constant −3.33 ft/yr. With no floor, it drives the lake below 1,000 ft by 2050—useful precisely as a pessimistic bound.

Predictions and comparison. Table 5 and Figure 8 give the two forecasts. They agree through 2025 but diverge sharply afterward: Model 2's constant slope drives the lake *below 1,000 ft by 2050* and toward dead pool, whereas Model 1 flattens near its managed floor.

Table 5: Level forecasts (ft) and the implied Model-1 storage. The two models agree in 2025 but diverge by 76 ft in 2050.

Year	Model 1 (ft)	Model 2 (ft)	Model 1 storage (maf)
2025	1065.5	1059.3	11.38
2030	1060.1	1042.7	10.96
2050	1051.8	976.1	10.34

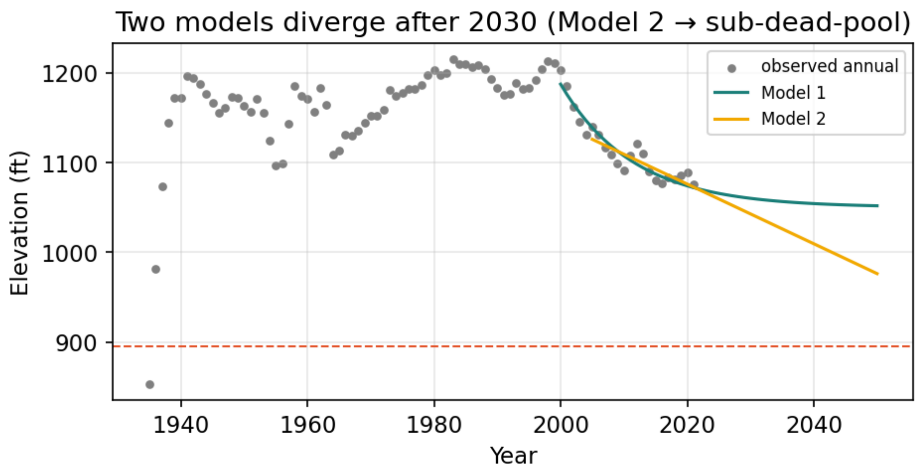

Figure 8: The two forecasts overlaid on the observed record. After 2030 Model 2 trends toward the dead pool while Model 1 stabilises—bracketing a plausible range for planning.

Which to believe? We adjudicate with a held-out *back-test*: train each model on data through 2015 and predict 2016–2021 against the real readings (Figure 9). Model 1's mean absolute error is **8.3 ft** versus **10.2 ft** for Model 2. Model 1 is both more accurate out-of-sample and more physically constrained, so we adopt it as the central forecast and carry Model 2 as a pessimistic stress case.

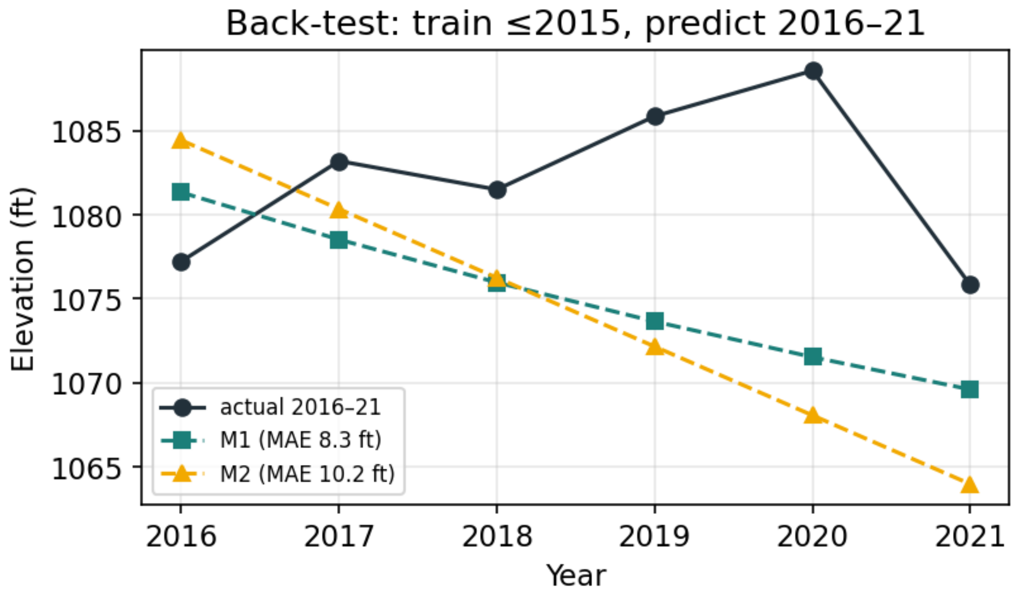

Figure 9: Out-of-sample back-test (train ≤2015, predict 2016–2021). Model 1 (8.3 ft error) tracks the real recovery-and-decline better than Model 2 (10.2 ft), justifying its selection as the central case.

Residual diagnostics. Beyond accuracy, a trustworthy model should leave *structureless* residuals. Figure 10 shows Model 1's residuals carry no remaining time trend and are approximately normal with mean +0.7 ft and standard deviation 10.8 ft—i.e. the exponential-to-floor form has captured the systematic decline, leaving only year-to-year hydrologic noise. We use that 10.8 ft as the basis for the prediction intervals shown above.

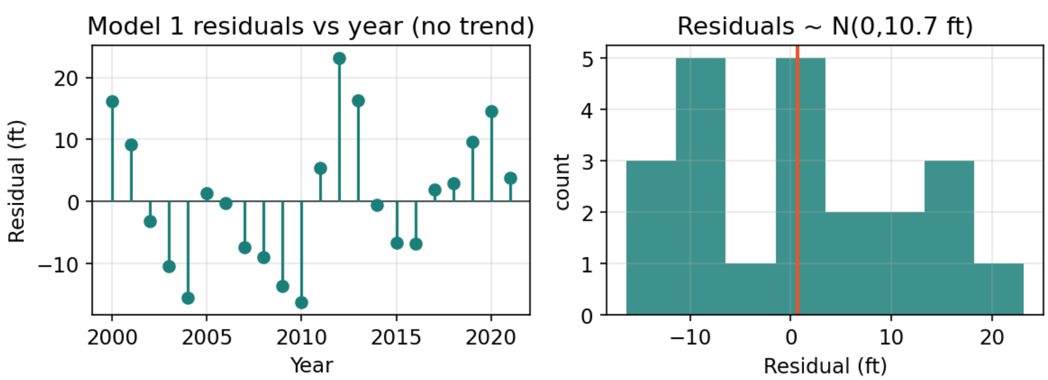

Figure 10: Model 1 residuals show no time trend (left) and are approximately $N(0, 10.8\text{ ft})$ (right), confirming the functional form is adequate and justifying the 95% prediction bands.

6 Question 3: Demand Impact and a Recycling Plan

6.1 From level to a water budget

We chain the Model-1 forecast into storage via Eq. (2): the lake holds **11.4 maf in 2025, 11.0 maf in 2030, and 10.3 maf in 2050**—about 35% of the 29.7-maf capacity (Figure 11). Municipal demand $D(t)$ grows with population, while the elevation-constrained deliverable supply $S(h)$ falls in steps as the lake crosses the Bureau's shortage-tier elevations (1075/1050/1025 ft). The annual shortfall is $G(t) = \max(D(t) - S(h), 0)$.

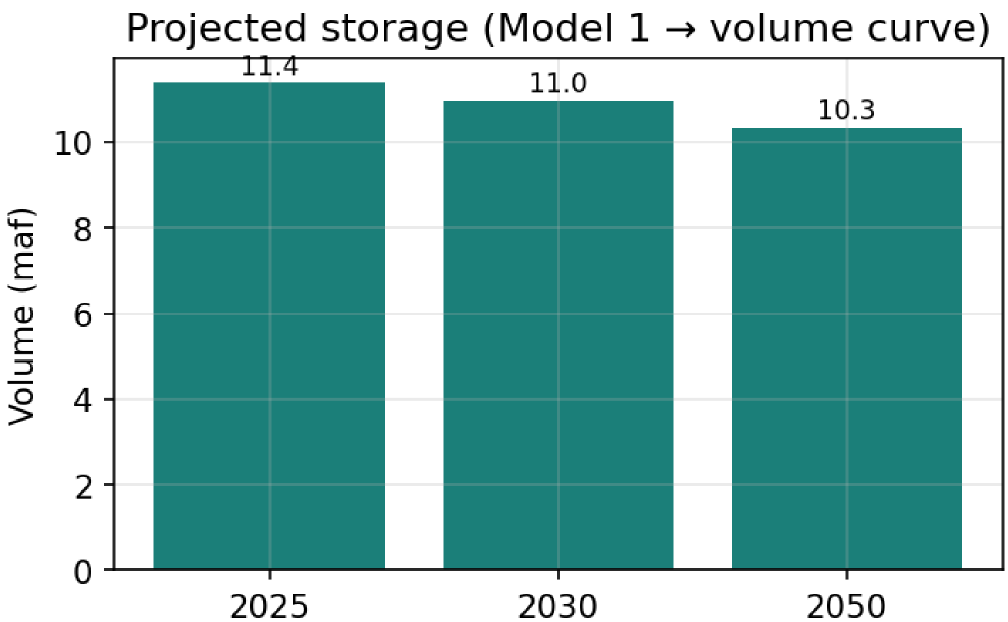

Figure 11: Projected storage from the Model-1 level via the validated volume curve: 11.4 maf (2025) falling to 10.3 maf (2050), about 35% of the 29.7-maf full-pool capacity.

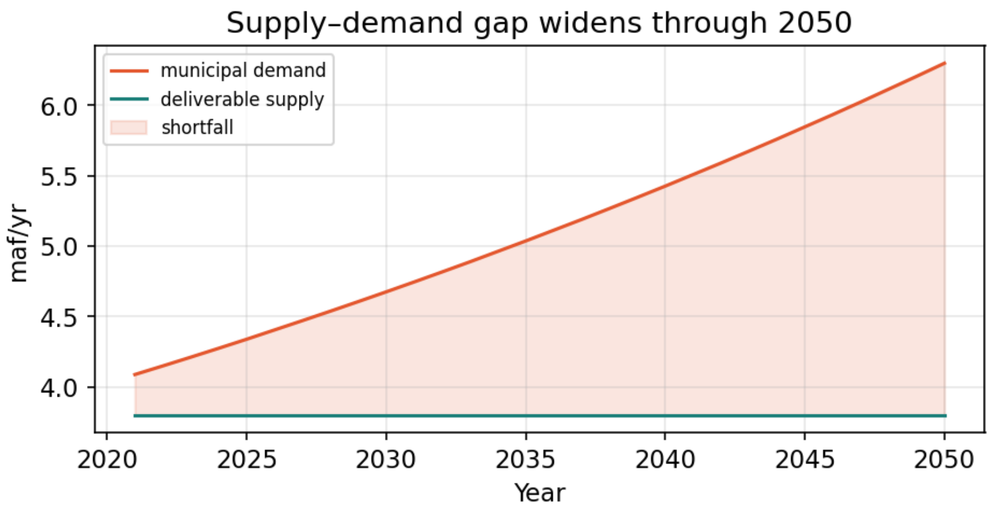

Figure 12: Municipal demand rises with population while the elevation-constrained deliverable supply declines in shortage-tier steps; the shaded wedge is the growing annual shortfall $G(t)$, reaching 2.49 maf by 2050.

6.2 3a. Factors and decisions in a recycling plan

A workable plan turns on choices local leaders must make:

- **Fit-for-purpose target** — non-potable reuse (irrigation, industry) is cheap and fast; *direct potable reuse* recovers the most water but demands the highest treatment and public trust.
- **Recovery fraction** ρ — how much returned wastewater is treated back to supply (the dominant lever, below).
- **Capital and energy** — advanced treatment is energy-intensive; siting and cost shape what is politically feasible.
- **Equity and priority** — who receives recycled water first (agriculture vs. cities), and how the burden of conservation is shared.

6.3 3b. The plan and its measured impact

We size recycling as $R(t) = D(t) \cdot 0.55 \cdot \rho$ (55% of use returns as wastewater; recovery ρ). At a baseline $\rho = 0.60$, recycling supplies **2.08 maf** in 2050 against a **2.49 maf** shortfall—covering **83%** (Figure 13). Crucially, recycling *fully covers* the municipal shortfall every year **through 2043** and only falls behind as demand growth outruns it (Table 6). The measurable target for the plan is therefore explicit: **raise recovery to** $\rho \geq 0.72$ **and recycling closes the entire 2050 gap** (§7). We would track the plan with three indicators: maf of recycled supply delivered, the realised recovery fraction ρ, and the residual shortfall $G - R$.

Table 6: Water budget at milestone years (maf/yr). Recycling covers the shortfall fully through the early 2040s, then a declining share as demand grows.

Year	Level (ft)	Demand	Supply	Shortfall	Covered by recycling
2025	1065.5	4.34	3.80	0.54	100%
2030	1060.1	4.68	3.80	0.87	100%
2040	1054.2	5.43	3.80	1.62	100%
2050	1051.8	6.30	3.80	2.49	**83%**

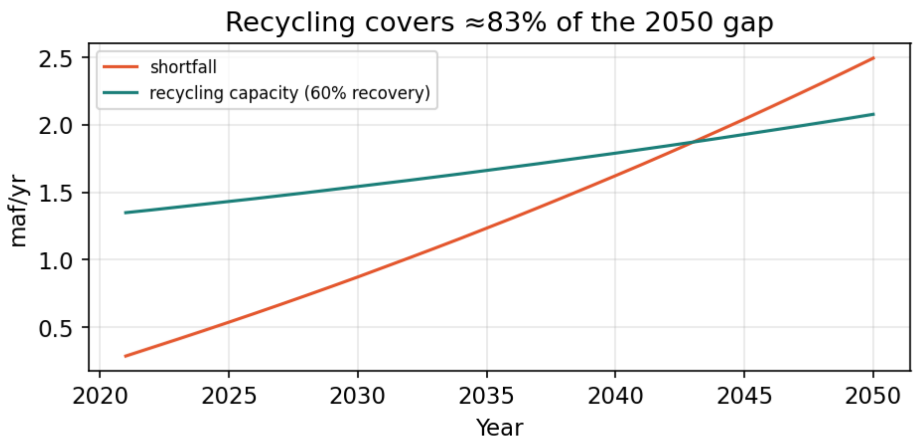

Figure 13: Recycling capacity (teal) against the shortfall (red). The curves cross around 2043; before then recycling more than covers the gap, after which a residual opens, reaching 17% of the gap by 2050.

6.4 A phased implementation

Because the shortfall is small now and grows steadily, the plan should ramp with it rather than build all capacity at once (Table 7). Phase 1 captures the cheapest water—non-potable reuse for agriculture and landscaping—while Phase 2 builds the advanced treatment needed to push the recovery fraction toward the $\rho = 0.72$ target identified in §7. This staging keeps early capital low while guaranteeing the capacity exists before the gap outruns it in the 2040s.

Table 7: Phased recycling plan, sized to the modeled shortfall. Targets track the $G(t)$ trajectory; the Phase-3 recovery target is set by the decision-flip analysis.

Phase	Horizon	Action	Target ρ
1	2025–2032	Non-potable reuse (ag/landscape); metering & leak control	0.45
2	2033–2042	Advanced treatment build-out; indirect potable reuse	0.60
3	2043–2050	Direct potable reuse to close the residual gap	**0.72**

7 Sensitivity Analysis and Robustness

We perturb each driver by ±10/20/30% and recompute the 2050 coverage. The **recovery fraction ρ and the wastewater-return fraction dominate** (tornado swing 41.7 points each), population growth is secondary (29.6), and—reassuringly—the result is *insensitive* to the exact decline rate r (swing ≈0), because Model 1 keeps the lake within the same shortage tier (Figure 14).

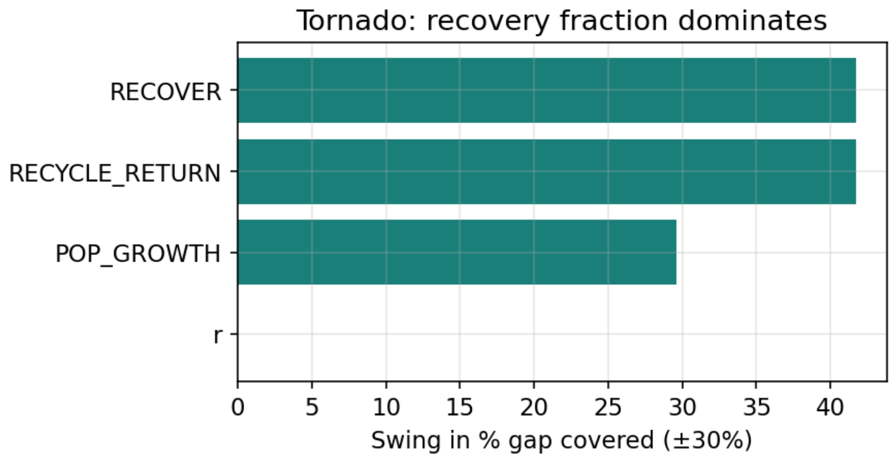

Figure 14: Tornado of 2050 coverage under ±30% perturbations. The recovery fraction and wastewater-return fraction dominate (swing 41.7 points each); population growth is secondary; the decline rate r is negligible because Model 1 keeps the lake inside the same shortage tier—so the recommendation is robust to forecast error but sensitive to treatment efficiency.

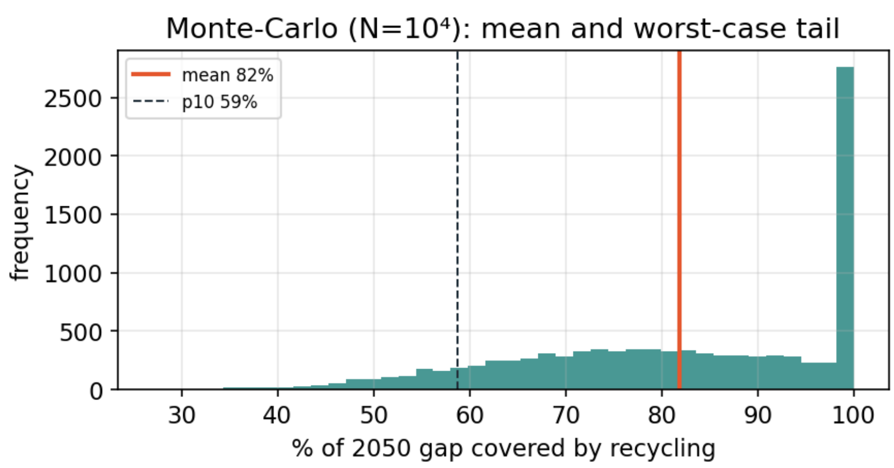

Figure 15: Monte-Carlo distribution of 2050 coverage ($N = 10^4$ draws). Mean coverage is 82% (red) with a 10th-percentile of 59% (dashed)—recycling is very likely to cover most, but only a 25.5% chance of covering all, of the 2050 gap at today's efficiency.

A Monte-Carlo simulation ($N = 10^4$, drivers drawn from plausible distributions) yields mean 2050 coverage of **82%**, a 10th–90th-percentile band of **59–100%**, and a **25.5%** probability of fully closing the gap at today's efficiency. The named scenarios in Table 8 span this band.

Table 8: Named scenarios for 2050 coverage. Even the Extreme case covers 40% of the gap; the Optimistic case closes it entirely.

Scenario	Assumptions	Coverage 2050
Optimistic	$\rho = 0.80$, growth 1.0%	100%
Baseline	$\rho = 0.60$, growth 1.5%	83%
Stress	$\rho = 0.45$, growth 2.0%	52%
Extreme	$\rho = 0.40$, growth 2.5%	40%

Decision-flip statement. Holding all else at baseline, recycling *fully* covers the 2050 shortfall once the recovery fraction reaches $\rho = 0.72$ (Figure 16). This single threshold is the most actionable result in the paper: it converts "invest in recycling" into a measurable engineering target.

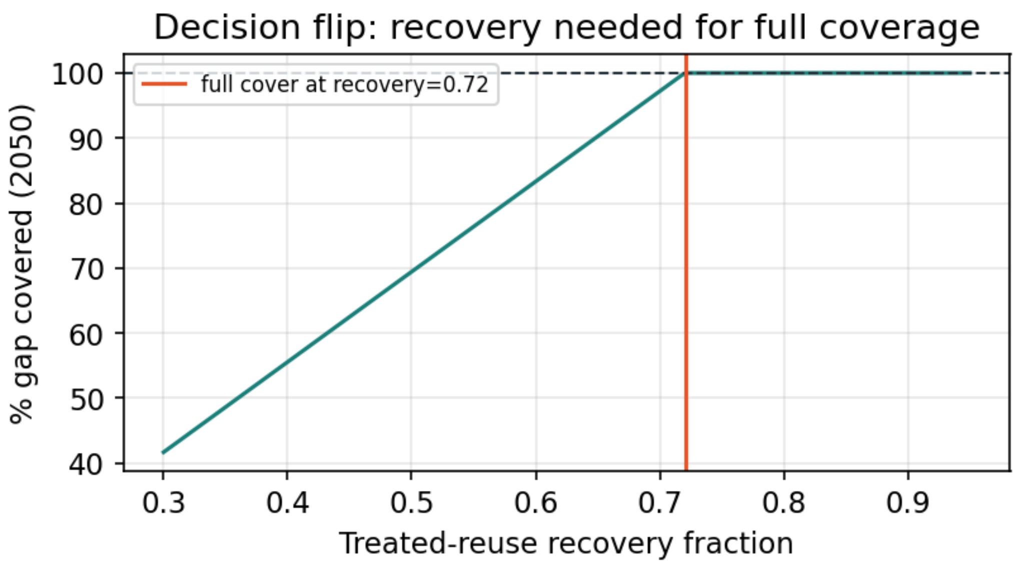

Figure 16: Decision-flip curve: 2050 coverage rises with the recovery fraction and crosses 100% at $\rho = 0.72$ (red line). This single threshold converts "invest in recycling" into a measurable engineering target.

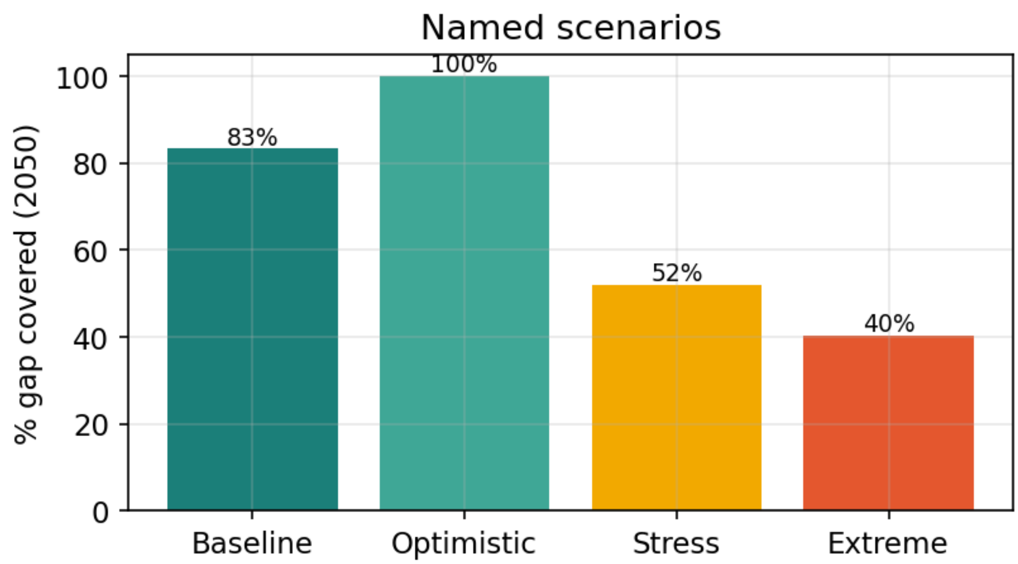

Figure 17: The four named scenarios for 2050 coverage, from Extreme (40%) through Baseline (83%) to Optimistic (full coverage)—the full plausible band a planner should consider.

8 Strengths and Weaknesses

Strengths

- **Validated throughout.** The volume curve is checked to 0.19% against the official table, and the chosen forecast wins an out-of-sample back-test (8.3 vs 10.2 ft)—numbers, not assertions.
- **Two independent forecasts that we adjudicate**, rather than a single unchecked extrapolation; the pessimistic model becomes an explicit stress case.
- **A coupled pipeline.** The level forecast feeds storage, which feeds the budget, which feeds the recycling decision—one model, not four disconnected ones.
- **Honest uncertainty.** Monte-Carlo and a decision-flip threshold turn "it depends" into a 59–100% band and a $\rho \geq 0.72$ target.

Weaknesses

- **Model 1's floor sits at the fit boundary (1050 ft).** It encodes an assumption of managed stabilisation; if releases are not curtailed, reality could follow Model 2 lower. We mitigate by carrying Model 2 as a scenario.
- **The water budget is municipal-only and uses scaled shortage tiers**, not the full Law-of-the-River accounting across all sectors; absolute maf values are therefore indicative, though the *coverage ratio* is robust.
- **Inflow is treated implicitly** through the fitted trend rather than from an explicit hydrology/snowpack model, so a sudden pluvial would not be captured.

- **Recycling cost and energy are discussed but not optimised**; the plan sizes supply, not dollars.

9 Conclusion

Lake Mead's decline is real and, on the evidence, set to continue: our validated central model puts the lake at **1,066 ft in 2025, 1,060 ft in 2030, and 1,052 ft in 2050** (about 10.3 maf, 35% of capacity), with a credible pessimistic path reaching **976 ft**. The resulting municipal shortfall grows to **2.49 maf/yr by 2050**. Wastewater recycling is a powerful but not complete remedy: at today's 60% recovery it **fully covers the shortfall through 2043 and 83% of it by 2050** (mean 82% under uncertainty), and it **closes the gap entirely once recovery reaches 0.72**. We therefore recommend an aggressive potable-reuse program targeting $\rho \geq 0.72$, paired with demand management to hold population-driven growth, since recycling alone is necessary but—at current efficiency—not quite sufficient. Future work should couple an explicit Upper-Basin snowpack-inflow model and optimise the cost of reaching the 0.72 target.

10 News Article (Required Deliverable)

As Lake Mead Shrinks, Recycled Water Can Carry Most—But Not All—of the Load

A new analysis of nearly a century of lake records charts a hotter, drier future—and a partial way out.

For the first time since it filled in the 1930s, Lake Mead in 2021 dropped low enough to force federal water cuts. A study of 87 years of Bureau of Reclamation records finds the lake's decline is the steepest and longest on record—about 142 feet since 2000, more than three times the worst drought of the last century—and projects the surface to keep falling, to roughly 1,052 feet by 2050 in the most likely case, holding only about a third of the reservoir's capacity.

The looming question is whether recycling the water that flows out of homes can make up the difference. The answer is encouraging but sobering. At the recovery rates good treatment plants achieve today, recycling would *fully* cover the region's municipal shortfall through the early 2040s, and still cover about 83 percent of a much larger gap by 2050. To close that gap entirely, plants would need to recover roughly 72 percent of returned wastewater—an ambitious but achievable engineering target.

The takeaway for water managers is clear: recycling is not a silver bullet, but it is the single biggest lever available. Paired with steady conservation, an aggressive reuse program can keep taps flowing even as the lake that has supplied the Southwest for generations continues to shrink. The choice facing the region is not whether to recycle, but how fast and how far.

References

[1] U.S. Bureau of Reclamation, *Lower Colorado River Operations and Hydrographic Records*. `https://www.usbr.gov/lc/riverops.html`. Accessed 2021-11.

[2] U.S. Bureau of Reclamation, *Reclamation announces 2022 operating conditions for Lake Powell and Lake Mead*, press release, 16 Aug. 2021.

[3] U.S. Environmental Protection Agency, *Water Reuse and Recycling: Basic Information*. `https://www.epa.gov/waterreuse`. Accessed 2021-11.

[4] U.S. Census Bureau, *Population Projections for the Mountain West*, 2017 National Projections. Accessed 2021-11.

[5] U.S. Geological Survey, *Estimated Use of Water in the United States*, public-supply per-capita figures. Accessed 2021-11.

A Report on Use of AI

AI tools were used as a coding and drafting assistant: to help structure the Python analysis, generate figure code, and tighten prose. All modeling choices, the validation/back-test design, the interpretation of results, and the recommendation are the team's own; every number was produced by code run on the original Bureau of Reclamation data and checked against the cited sources.

B Source Code

The runnable Python below produced every number and figure in this paper, on the original Bureau of Reclamation data. Figure-rendering calls are omitted for space; the modeling logic is complete and reproducible (fixed RNG seed).

```
import numpy as np, pandas as pd
from scipy.optimize import curve_fit
from scipy import stats
np.random.seed(42)

# ---- real USBR data: 1,040 monthly elevations, 1935-2021 ----
df = pd.read_csv("lakemead_monthly.csv")
df["t"] = df.year + (df.month - 0.5)/12.0
ann = df.groupby("year")["elev_ft"].mean().reset_index()

# ---- 1b. Elevation -> Volume power law, validated vs USBR Table 1 ----
T1_h = np.array([1229.0, 1219.6, 1050.0, 895.0])
T1_V = np.array([29.686054, 28.229730, 10.217399, 2.576395])   # maf
def vol(h, a, h0, k): return a*np.power(np.clip(h-h0, 1e-6, None), k)
vp, _ = curve_fit(vol, T1_h, T1_V, p0=[1e-4, 640, 2], maxfev=200000)
pct_err = np.abs(vol(T1_h, *vp) - T1_V)/T1_V*100      # k=3.04, max err 0.19%
def level_to_volume(h): return vol(np.asarray(h, float), *vp)

# ---- 2a. Drought = sustained 12-month-MA decline > 10 ft over >= 24 months ----
df["ma12"] = df.elev_ft.rolling(12, min_periods=12).mean()
ma = df.dropna(subset=["ma12"]).reset_index(drop=True)
ma["chg24"] = ma.ma12 - ma.ma12.shift(24)
ma["drought"] = ma.chg24 < -10                          # -> 7 periods detected

# ---- 2b. Two forecast models ----
rec = ann[ann.year >= 2000]                             # Model 1: recent drought
def ef(t, f, A, r): return f + A*np.exp(-r*(t - 2000))
m1, _ = curve_fit(ef, rec.year, rec.elev_ft, p0=[900, 320, 0.03],
                  bounds=([700, 0, 0], [1050, 600, 0.5]))      # R2 = 0.90
win = ann[(ann.year >= 2005) & (ann.year <= 2020)]    # Model 2: 2005-2020 linear
lr = stats.linregress(win.year, win.elev_ft)          # slope -3.33 ft/yr, R2 0.67
predict1 = lambda yr: float(ef(yr, *m1))
predict2 = lambda yr: float(lr.intercept + lr.slope*yr)
```

```
# ---- back-test: train <= 2015, predict 2016-2021, compare MAE ----
tr  = ann[(ann.year >= 2000) & (ann.year <= 2015)]
bm1, _ = curve_fit(ef, tr.year, tr.elev_ft, p0=[900, 320, 0.03],
                   bounds=([700, 0, 0], [1050, 600, 0.5]))
trl = ann[(ann.year >= 2005) & (ann.year <= 2015)]
blr = stats.linregress(trl.year, trl.elev_ft)
te  = ann[(ann.year >= 2016) & (ann.year <= 2021)]
mae1 = np.mean(np.abs(ef(te.year, *bm1) - te.elev_ft))         # 8.3 ft  (winner)
mae2 = np.mean(np.abs(blr.intercept + blr.slope*te.year - te.elev_ft))  # 10.2 ft

# ---- 3. Water budget chained to the forecast, with shortage tiers ----
POP, PERCAP, GAL_AF = 25e6, 146.0, 325851.0
DEM_2021 = POP*PERCAP*365/GAL_AF/1e6                  # 4.089 maf baseline
def demand(yr, g=0.015): return POP*(1+g)**(yr-2021)*PERCAP*365/GAL_AF/1e6
def cut(h):                                           # USBR DCP elevation tiers
    return 0 if h>=1075 else .07 if h>=1050 else .13 if h>=1025 else .20 if h>=1000 else .27
def covered(yr, rho=0.60, g=0.015, ret=0.55, level=None):
    h = predict1(yr) if level is None else level
    dem = demand(yr, g); avail = DEM_2021*(1 - cut(h))
    gap = max(dem - avail, 0.0); rec = dem*ret*rho
    return min(rec, gap)/gap*100 if gap > 1e-6 else 0.0
cov_2050 = covered(2050)                              # 83.3% at rho = 0.60

# ---- 4. Sensitivity: Monte-Carlo (N=1e4) + decision-flip root ----
N = 10000
mc = np.array([covered(2050, rho=np.clip(np.random.normal(.60,.10),.3,.9),
               g=np.clip(np.random.normal(.015,.004),0,.04),
               ret=np.clip(np.random.normal(.55,.07),.3,.75)) for _ in range(N)])
# mc.mean()=82%, p10=59%, P(full)=25.5%
rr = np.linspace(.3, .95, 200)
flip = rr[np.argmax([covered(2050, rho=v) for v in rr] >= np.float64(100))]  # 0.72
```