Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

How Many Days of Water? Linking Storage Adequacy, Tank Integrity, and Inspection Assurance in Strategic Potable-Water Reserves

Ayman Ramadan

ayman_as55@yahoo.com

International Business Academy of London https://orcid.org/0009-0002-0317-7867

Research Article

Keywords: strategic water reserves, desalination, usable storage, dead storage, Monte Carlo, expert elicitation, radiography, public–private partnership, water security, tank integrity, storage adequacy and inspection assurance

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11732966/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

How Many Days of Water? Linking Storage Adequacy, Tank Integrity, and Inspection Assurance in Strategic Potable-Water Reserves

Dr. Ayman Salah Abdelaziz Ramadan

Independent Researcher

Kingdom of Saudi Arabia

Email: ayman_as55@yahoo.com

ORCID: https://orcid.org/0009-0002-0317-7867

Abstract

Background. Strategic potable-water reserves in desalination-dependent countries are announced in nominal days of demand. The volume a city can actually draw during a disruption depends on dead storage, operating level and the availability of the storage units, and these are rarely reported next to the nominal figure.

Methods. We built a unit-level Monte Carlo model of a strategic reserve (20 tanks in four construction batches, clustered defects, four radiography regimes, three demand tiers) and parameterised it with a two-round structured consultation of practitioners (10 experts in Round 1 on 10 November 2026; 13 in Round 2 on 11 November 2026). In Round 2 all 13 reported having discussed their answers with colleagues, so the pooled values are a consensus view and not independent judgements.

Results. With the panel's dead storage (about 17.5% of design volume) and operating level, seven nominal days equal about five usable days under total loss of supply; about 9.8–10 nominal days are needed to deliver seven usable days. For events of the ordinary size described by the panel (1–3 days) no shortfall occurred. For rare events (median 8.6 days, about 47% loss of supply capacity) the reserve was exhausted in 28–32% of runs at normal demand and in almost none at essential-level demand, although the panel judged that restrictions can be tolerated for only about four days (median). Extending radiography changed usable capacity by 0.2–1.3% of nominal volume depending on assumptions; whether this offsets its cost depends on defect severity and fragility assumptions.

Conclusions. Reserve targets should state their basis and report usable days. Operating policy and event severity matter more than inspection regime for capacity, while inspection assurance mainly narrows the tail. Findings rest on a consensus panel and on parameters that remain to be verified.

Keywords: strategic water reserves; desalination; usable storage; dead storage; Monte Carlo; expert elicitation; radiography; public–private partnership; water security; tank integrity; storage adequacy and inspection assurance.

1. Introduction

Desalination supplies a large share of potable water in water-scarce regions , and cities that depend on it hold strategic storage as insurance against interruptions of supply. Saudi Arabia relies on desalination for roughly 70% of its municipal supply according to one source , and its water strategy sets a target of seven days of strategic storage . A 2023 report put existing storage at about 21 million m³, equal to 2.2 days of municipal demand, and a 2026 source reports 21.8 million m³ . The water transmission and storage company reports more than 120 strategic storage projects, over 70 million m³ and about US$10 billion of investment .

Reserve targets are expressed in days, but a day-count is a ratio of two quantities, and each can be defined in more than one way. The numerator may be the volume to the overflow level or to the design liquid level, with or without the water that lies below the pump suction and cannot be drawn. The denominator may be normal demand, restricted demand, or demand net of the supply that survives a disruption. In practice, announcements of storage capacity do not always distinguish between nominal and usable storage. The dead volume is often not removed when days are announced. Published estimates for the Gulf range from about two to five days , and press sources disagree on how long reserves last.

Three gaps follow. First, we found no compact method that reconciles nominal and usable days. Second, reserve adequacy is usually judged by a single number rather than by the probability of shortfall and by tail risk, as reliability, resilience and vulnerability criteria would suggest . Third, the integrity of the tanks themselves, which depends on construction quality and inspection, is rarely connected to reserve capacity or to the contracts under which reserves are built. This statement of the gap is relative to the sources we consulted; it is not the result of a systematic search.

This paper makes four contributions: (1) definitions of nominal volume, dead storage, operating level and usable days; (2) a probabilistic model that reports shortfall probability and tail risk for ordinary and rare events; (3) a structured consultation of practitioners that supplies parameters usually missing from the literature; and (4) contract and financing implications, including a test of whether extending radiography pays for itself in usable volume.

2. Background

Strategic reserves in the Gulf. Earlier work estimated emergency capacity in the Gulf Cooperation Council states at roughly two to five days of demand . Aquifer storage and recovery has received most attention: Abu Dhabi stored about 26 million m³, described as 90 days of supply . A GIS-based framework for artificial aquifer recharge in Qatar estimated that up to 182.8 million m³ could be added to the strategic reserve . Above-ground tanks are a significant component of the current Saudi strategic-storage programme .

Tank types in strategic reserves. Above-ground strategic storage relies on three main technologies: welded steel tanks built to API 650 and AWWA D100 ; reinforced or prestressed concrete tanks built to standards such as AWWA D110 and AWWA D115 ; and glass-fused-to-steel (GFS, also called glass-lined steel or GLS) bolted tanks covered by AWWA D103 and EN 28765 . Each type has different dead-storage geometry, failure modes and inspection requirements. The present study models welded steel tanks, for which the radiography data and inspection regimes are best documented; extending the framework to concrete and GFS tanks is identified as future work.

Integrity of welded steel tanks. Welded steel tanks for potable water are built to standards such as API 650 and the AWWA standard for welded carbon steel water-storage tanks and are accepted on a package of examinations, including visual, vacuum-box, penetrant, radiographic and hydrostatic examinations . Inspection, repair, alteration and reconstruction of welded storage tanks in service are addressed by API Standard 653 .

In a field case study of two nominally identical tanks, 67 of 86 listed locations were radiographed and 6 were rejected for porosity (9.0%; exact 95% confidence interval 3.4–18.5%), all in one tank and one radiography batch. Confidence limits were calculated using the Clopper-Pearson method .

A 25% random sample would find an isolated defective seam in only a quarter of cases but would often find a cluster; simulation suggested that extension to the whole batch could identify defective seams throughout the affected batch . In the fourth shell course, rejection was 5 of 12 in one tank against 0 of 12 in the other (risk difference 41.7 percentage points, 95% interval 8.7 to 68.0). Confidence intervals were calculated using the method described by .

That study motivates the inspection regimes modelled here. Probabilistic work on weld-inspection sampling plans and on probability of detection supplies the logic of the regimes, but we found no work that links them to reserve capacity.

Contracting for storage. Saudi Arabia has extended public–private partnerships from desalination to transmission and storage . Performance-based contracts need measurable indicators. A nominal-days target does not say how much water is available when needed, which is the gap this paper addresses.

Table 1. Positioning of this study against the approaches reviewed. Statements about other works reflect the sources consulted, not a systematic review.

Approach	Typical output	Usable volume (net of dead storage)	Asset reliability	Link to contracts
National targets (seven-day target) [S]	Nominal days	Not explicitly reported	Not addressed	Storage PPP programme
GCC emergency-capacity estimates [1]	Days of capacity by country	Not explicitly reported	Not addressed	Not addressed
Aquifer storage and recovery studies [2, 3]	Volume, days of supply, recharge potential	Not applicable (subsurface storage)	Not addressed	Not addressed
Radiography coverage case study [5]	Rejection rates, detection probability	Not addressed	Inspection coverage and detection	Proposes batch-based extension
This study	Usable days, shortfall probability, tail risk	Explicit	Unit-level, clustered defects, four inspection regimes	Indicator and clause candidates

3. Methods

3.1 Definitions

- **Nominal volume:** volume at the design liquid level. The design level was described by practitioners as about 80% of the geometric volume of the tank, the remainder lying above the design level.
- **Dead storage:** water below the pump suction or outlet, which cannot be drawn.
- **Operating level (fill):** the share of design volume held at a random moment.
- **Usable volume:** nominal volume × (fill − dead storage) × share of units available.
- **Days of autonomy:** usable volume divided by daily net draw-down, under (i) total loss of supply, equal to normal daily demand, and (ii) partial loss, where the surviving supply covers part of demand.

3.2 Reserve model

The reserve consists of K = 20 equal units arranged in B = 4 construction batches. Usable storage is

$$U = N \times max\,(0, F - D) \times R$$

where N is nominal storage (days), F the operating fill fraction, D the dead-storage fraction, and R the proportion of units available during the event.

Each unit may contain a defect with marginal probability q. In the clustered-defect variant, a batch is classified as bad with probability

$$P_b = min(1, q \,/\, q_u)$$

and units within a bad batch are defective with probability $\boldsymbol{q_u = 0.67}$, preserving the marginal defect probability q.

Four inspection regimes, in the spirit of probabilistic sampling-plan studies , act on each defective unit: A, a 25% sample; B, a 35% sample; C, a 25% sample with batch-triggered extension, in which finding a defect in any unit examines the whole batch; and D, full radiography. Detection probability is calculated as the inspection coverage multiplied by the single-exposure probability of detection, p_d .

An undetected defective unit is out of service when needed with probability φ; any unit may also be unavailable for service reasons [A].

An event is defined by the share L of supply capacity lost and its duration T. The net daily draw-down is $r = max\,(0, D_r - (1 - L) \times e)$, where D_r is the demand tier relative to normal demand and e is delivery effectiveness [A]. A shortfall occurs when $U < r \times T$. Demand tiers are normal (1.0), essential and survival, both expressed relative to normal demand. Results are conditional on an event; no annual frequency is modelled. Tail risk is the probability that fewer than 85% of units are available. Shortfall probability and tail risk follow the reliability and vulnerability criteria of Hashimoto et al. (1982).

A stress-test scenario was also examined in which the defect fragility parameter (φ) was multiplied by three. This scenario was introduced after all Round-2 participants stated that hidden defects were more likely to become evident under emergency loading. The multiplier itself was not elicited from experts and should not be interpreted as an estimate of the true increase in fragility.

3.3 Structured expert consultation

Parameters that are not published were obtained from practitioners using a structured, Delphi-type questionnaire , in which quantities were elicited as a low, best and high value with a confidence rating, following the logic of the IDEA protocol , and agreement as ratings from 1 to 5. Round 1 was run on 10 November 2026 with 10 experts; Round 2 on 11 November 2026 with 13 (the 10 returning experts and 3 new ones). Round 2 gave the group medians and ranges from Round 1 and asked four clarifying questions. The consultation is reported with reference to the CREDES recommendations .

The panel worked in management, operating and technical roles: project, PPP and general management (8), operations and maintenance (1), hydraulic design (1), health, safety and emergency management (1), and non-destructive-testing or welding specialists (2). We recorded the employer and not the client, and corrections to Round-1 profile entries are documented in a change log. Table 2 gives the composition; employer types are aggregated to keep answers anonymous.

Table 2. Composition of the panel (aggregated).

Characteristic	Round 1 (n = 10)	Added in Round 2 (n = 3)	Round 2 total (n = 13)
Employer type			
Utility or operator	6	0	6
Consulting office	3	0	3
Government ministry	0	1	1
Steel fabricator	0	2	2
Contractor	2	2	5
NDT	1	1	2
Hydraulic design	1	0	1
Role			
Project, PPP or general management	6	2	8
Operations and maintenance	1	0	1
Hydraulic design	1	0	1
Health, safety and emergency management	1	0	1
NDT or welding specialist	1	1	2
Experience			
More than 15 years	4	0	4
5 to 15 years	6	3	9
Extensive direct experience with strategic tanks	10	3	13

Independence. In Round 2 all 13 experts answered that they had discussed their answers with colleagues before submitting, and several Yes/No answers were identical across the whole panel. The pooled values are therefore reported as the output of a consensus consultation. We did not treat the answers as independent draws, and we do not report statistical tests on them. Independent first-round judgements are a core feature of Delphi and IDEA-type designs , so this is a departure from them.

Pooling used the equal-weight average of each expert's low, best and high values, and triangular distributions were fitted to the pooled triples . Sensitivity to disagreement is shown through the ranges and through alternative variants rather than through expert-level resampling.

3.4 Inputs

Table 3. Inputs of the model. Tags: [S] sourced; [E] expert-derived (consensus panel); [A] assumption. Triangular (low, mode, high) values are the pooled expert triples.

Parameter	Value used	Tag	Source
Dead storage (fraction of design volume)	Triangular (0.135, 0.175, 0.222)	E	C1
Operating level (fraction of design volume)	Triangular (0.74, 1.00, 1.00)	E	C2
Defect prevalence q, any defect	Triangular (0.19, 0.515, 1.00)	E	D1
Defect prevalence q, major defects	Triangular (0.011, 0.021, 0.034)	E	Q4
Single-exposure probability of detection	Triangular (0.62, 0.745, 0.865)	E	D6
Fragility φ, defect only	Triangular (0.0096, 0.0169, 0.0296); variant ×3	E; factor 3 is A	Q1d; Q1c
Ordinary event: duration (days)	Triangular (1, 2.5, 3)	E	E1
Ordinary event: share of capacity lost	Triangular (0.25, 0.385, 0.565)	E	E2
Rare event: duration (days)	Triangular (4.15, 8.62, 13.08)	E	Q3
Rare event: share of capacity lost	Triangular (0.30, 0.47, 0.62)	E	Q3L
Essential demand (litres per person per day)	Triangular (74, 169, 197.5)	E	E3b
Survival demand (litres per person per day)	Triangular (48, 66.5, 80.5)	E	E3a
Normal demand (litres per person per day)	284	S	Saudi Water Authority

Parameter	Value used	Tag	Source
Delivery effectiveness e	Triangular (0.85, 0.92, 1.00)	A	none
Service unavailability of a unit	Triangular (0.005, 0.02, 0.05)	A	none
Clustering parameter q_u; layout	0.67; 20 units in 4 batches	A	none
Storage cost (US dollars per m^3)	157, 257, 473 (pooled)	E	F4 (discussion only)
Extra cost of extending radiography (fraction of tank cost)	0.009, 0.010, 0.021 (pooled)	E	D8 (benefit–cost ratios)

3.5 Computation

Each scenario was simulated with 200,000 draws and common random numbers across inspection regimes, following standard Monte Carlo practice . Outputs are usable days (mean and 5th percentile), shortfall probability by demand tier, tail risk, and the gain in usable fraction relative to regime A. The benefit–cost ratio of a regime is the gain in usable fraction divided by the extra cost as a fraction of tank cost. The code is provided as supplementary material.

Sensitivity analysis. A one-way sensitivity analysis was performed around the base case by varying four influential inputs individually while holding all other parameters constant: dead-storage fraction, operating level, defect prevalence q, and defect fragility φ. Each parameter was varied by ±25% around its baseline value and the resulting change in mean usable days was recorded. Results are presented as a tornado chart (Figure S1).

4. Results

4.1 What the panel said

The panel’s dead storage and operating level agree closely with the author’s field figures. Median agreement was 4 or 5 of 5 for a usable-days performance indicator, a minimum operating level, disclosure of dead storage and records of consumables, but only 2 for batch-triggered extension of radiography.

When asked whether extension would be acceptable if conditional, all 13 said yes, but they did not state the conditions. Two potential contractual conditions are suggested for future evaluation, based on the consultation and not elicited from the panel: prior approval by the consultant or owner, and a clear definition of the batch (the same welder, period or materials). The reasons given against extension were cost (4), contractual disputes (3), schedule (2), difficulty defining a batch (2), shortage of radiography crews (1) and absence from the code (1). All 13 allocated shortfalls of operating level and events beyond the design basis to the authority, undetected construction defects to the developer, and reductions of inspection scope to shared responsibility. Experts judged that clustering of rejected welds by batch, period or crew is likely (median 4).

The panel considered that demand can be held at the essential level for a median of four days (range two to seven).

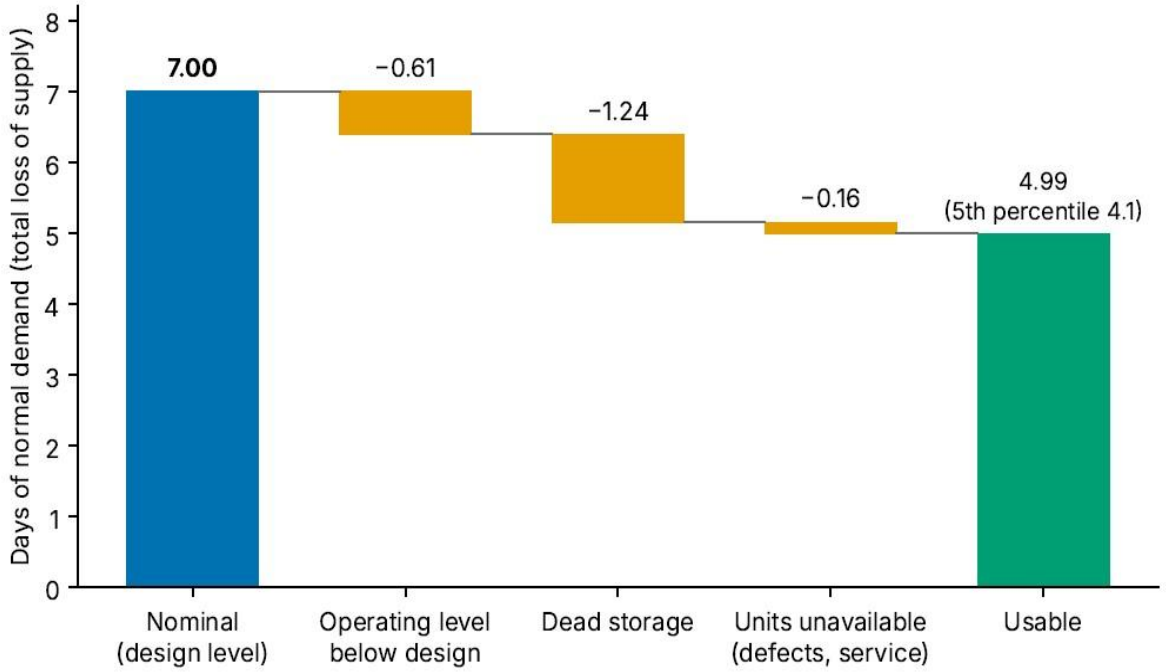
Experts interpreted D1 as any repairable defect, including minor cosmetic or workmanship-related defects, rather than safety-critical flaws.

4.2 Nominal versus usable days

With the panel’s inputs, seven nominal days gave 4.98–5.01 usable days on average under total loss of supply (5th percentile about 4.1–4.2). About 9.8–10 nominal days are needed to deliver seven usable days. Dead storage (1.24 days) and operating level below design (0.61 days) account for most of the 2.0-day gap; unavailable units account for 0.16 days.

Figure 1. Seven nominal days become about five usable days under total loss of supply. Bridge from nominal volume to usable volume: operating level below design, dead storage and unavailable units (mean of 200,000 runs; expert-derived inputs; file Fig1_nominal_to_usable).

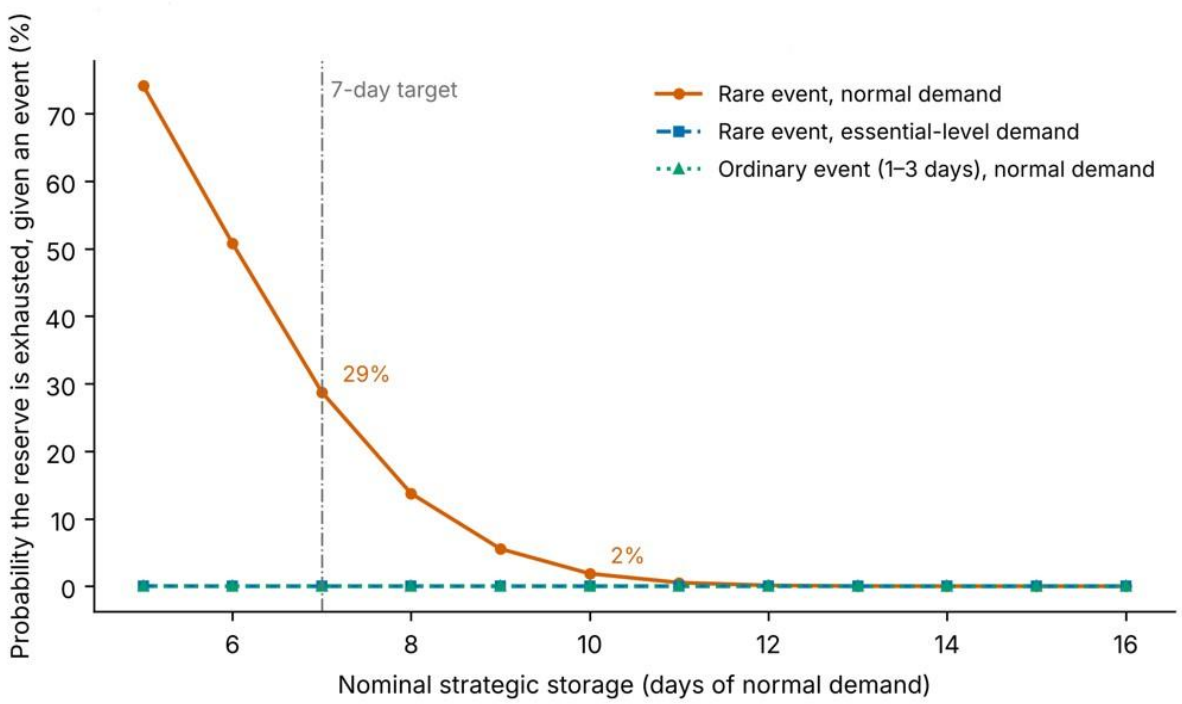
4.3 Adequacy under ordinary and rare events



of

to 10 days lowers the probability of exhaustion at normal demand from approximately 29% to approximately 2% under the baseline assumptions.

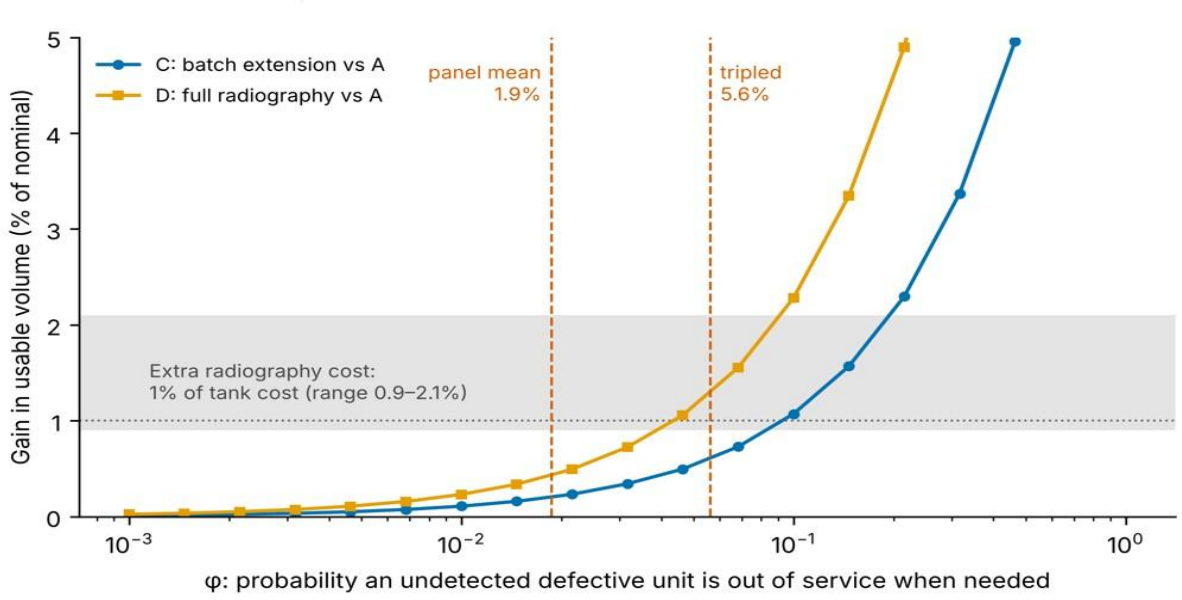
Figure 2. Probability that the reserve is exhausted, given an event, as a function of nominal storage, for ordinary events (1 to 3 days) and rare events, at normal and essential-level demand (inspection regime C; file Fig2_shortfall_vs_nominal).



4.4 Does inspection assurance change usable capacity?

Under the base inputs, batch-triggered extension raised the usable fraction by 0.20% of nominal volume and full radiography by 0.43%, compared with a 25% sample. Against an extension cost of about 1% of tank cost these give benefit–cost ratios of about 0.2 and 0.4 in terms of usable volume alone. If φ is tripled, the gains are 0.60% and 1.28% (ratios 0.6 and 1.3). If only major defects are considered, the gain is below 0.05%. Tail risk fell from 0.56% (regime A) to 0.43% (C) and 0.31% (D) in the base case, and from 2.1% to 1.3% and 0.5% with φ tripled. The cost of full radiography was not elicited, so its ratio is indicative. Under the base inputs, full radiography reaches a benefit-cost ratio of approximately 1 at φ of about 4%, and batch extension at about 9%.

Figure 3. Gain in usable volume from batch-triggered extension (C) and full radiography (D) relative to a 25% sample (A), as a function of φ, with the extra radiography cost of 1% of tank cost and the panel range of 0.9 to 2.1% (any-defect prevalence; file Fig3_inspection_breakeven).



4.5 Sensitivity analysis

For usable days (base 4.99), operating level was the dominant input: −25% and +25% changed it by −1.55 and +0.58 days, the latter truncated at 100%. Dead storage followed (∓0.30 days), while q and φ changed usable days by less than 0.01 days. The probability of exhaustion in a rare event (base 28.8%) was most sensitive to operating level (+48.8 and −14.5 percentage points), rare-event duration (−24.8 and +31.6) and capacity lost (−22.6 and +27.4), then to dead storage (−7.5 and +8.8); q and φ moved it by about 0.2 percentage points. The gain from batch-triggered extension (base 0.20% of nominal volume) was the only outcome in which q and φ led, ranging from 0.14 to 0.27% with q and from 0.15 to 0.25% with φ; in no case did it approach the 1% extension cost. A ±25% range does not reach the values at which extension breaks even (φ of about 9%, Figure 3), so the two analyses should be read together. (Figure S1, Supplementary material)

These findings support the robustness of the principal conclusion that the nominal-to-usable storage gap is governed primarily by operating level and dead storage rather than inspection assumptions.

5. Discussion

5.1 Implications for target setting

A target of seven nominal days delivers about five usable days when dead storage is about one-sixth to one-fifth of design volume and tanks are not always full. Authorities should state the basis of their targets and report usable days. A seven-day target specified in nominal terms and a seven-day target specified in usable terms are not operationally equivalent and may imply substantially different infrastructure investments.

Because operating level is a decision of operators, holding strategic tanks close to the design level is a low-cost lever, subject to water-quality and hydraulic constraints that the panel identified.

5.2 Contract and financing implications

The panel's support suggests that a usable-days performance indicator, a minimum operating level, disclosure of dead storage and mandatory records can be written into availability-based storage contracts; in-service inspection records could follow API 653 (2020). The panel did not endorse batch-triggered extension as a clause, even though it considered clustering likely and the extra cost small; our model finds that extension pays for itself in usable volume only if undetected defects are both common and likely to put units out of service when needed. The experts did not state the conditions under which they would accept extension. Two potential contractual conditions are suggested for future evaluation: (1) prior approval by the

consultant or owner before extension is initiated, and (2) a contractual definition of the batch (e.g., the same welder, period, or materials).

These proposals have not been tested with the panel and should be considered hypotheses for future evaluation. Testing them with practitioners is a subject for future work. The unanimous risk allocation is consistent with current practice but may reflect the common background of the panel.

5.3 Limitations

- **The panel was a consensus group.** All Round-2 respondents discussed their answers, the employers were concentrated in one utility and its supervising organisations, and some questions had identical answers from all experts. The pooled values may understate real disagreement. No tests of significance are reported.
- **Rare-event values come from the same panel and were obtained in one session.** They describe plausible events and not forecasts. No annual frequency is modelled, so shortfall probabilities are conditional.
- **Wide interpretation of some questions.** The panel understood a serious disruption as an operational outage of one to three days; rare events were asked separately. The defect-prevalence question was answered for any defect, mostly minor, and a separate question was needed for major defects.
- **Simplified structure.** Units are either available or not; there is no network hydraulics, no repair time, no partial leakage and no interaction with water quality. The model applies to welded steel tanks and does not address the different dead-storage fractions, failure modes or inspection regimes of reinforced concrete or glass-fused-to-steel tanks. Several parameters remain assumptions.
- **Inputs from press and investor sources** need verification against official documents before submission.
- **Costs** from the panel include pumps, piping and control systems and are not comparable to figures derived from public portfolios.

6. Conclusions

In the scenarios examined, seven nominal days translated into approximately five usable days, implying that about 9.8–10 nominal days were needed to reliably provide seven usable days, and the gap depends mainly on dead storage and operating level, not on weld inspection. Rare events, not ordinary ones, test a seven-day reserve, and restrictions on demand may matter as much as stored volume. Inspection assurance has a small effect on mean capacity and a larger effect on the tail.

Future work should test the proposed conditions for batch-triggered extension with practitioners, validate the inputs with an independent panel, extend the framework to reinforced concrete and glass-fused-to-steel tanks, and add a metric that accounts for the time rationing that can be tolerated.

DECLARATIONS

Ethics Approval

This study was conducted as independent research without institutional sponsorship or external funding. No formal ethical approval or exemption was obtained from an Institutional Review Board (IRB) or Research Ethics Committee prior to data collection. The study involved voluntary participation through a two-stage questionnaire, and participant confidentiality was maintained throughout the research process.

Informed Consent to Participate

All participants were individually contacted via email or WhatsApp and voluntarily agreed to participate before completing the questionnaires. Participation was voluntary, and the collected survey responses were handled confidentially. Participants' identities are not disclosed in the research findings.

Consent for Publication

The manuscript does not contain personally identifiable information about individual participants. Accordingly, separate consent for the publication of identifiable personal information is not applicable.

Conflict of Interest

The author declares no conflict of interest.

Funding

This research was conducted independently and received no external funding.

Data Availability Statement

The data supporting the findings of this study were derived from project inspection records that contain confidential project information and are therefore not publicly available. Anonymized data may be made available by the author upon reasonable request, subject to project-owner approval.

AI Usage Statement

Artificial intelligence tools were used to support language refinement, editing assistance, and manuscript development. All technical concepts, methodologies, interpretations, and conclusions remain the responsibility of the author.

Author Affiliation

AYMAN RAMADAN

Independent Researcher

Kingdom of Saudi Arabia

Author Contributions

AYMAN RAMADAN: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Framework Development, Writing – Original Draft, Writing – Review and Editing.

The author has reviewed and approved the final version of the manuscript.

References

American Water Works Association (AWWA), 2021. *Welded Carbon Steel Tanks for Water Storage,* Denver, Colorado, USA: American Water Works Association.

API 650, 2018. *API Standard 650: Welded Tanks for Oil Storage,* Washington, DC, United States: American Petroleum Institute.

API 653, 2020. *API Standard 653: Tank Inspection, Repair, Alteration, and Reconstruction,* Washington, DC, United States: American Petroleum Institute.

AWWA, 2013. *AWWA D110-13 (R18): Wire- and Strand-Wound, Circular, Prestressed Concrete Water Tanks,* Denver, Colorado: American Water Works Association.

AWWA, 2019. *AWWA D103-19: Factory-Coated Bolted Carbon Steel Tanks for Water Storage,* Denver, Colorado: American Water Works Association.

AWWA, 2020. *AWWA D115-20: Tendon-Prestressed Concrete Water Tanks,* Denver, Colorado: American Water Works Association.

Berens, A. P., 1989. NDE Reliability Data Analysis. In: *Metals Handbook: Nondestructive Evaluation and Quality Control.* 9th ed. Materials Park(Ohio): ASM International, pp. 689-701.

Clopper, C. J. & Pearson, E. S., 1934. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika,* 26(4), pp. 404-413.

Dalkey, N. & Helmer, O., 1963. An Experimental Application of the Delphi Method to the Use of Experts. *Management Science,* 9(3), p. 458–467.

Dar Al-Handasah, n.d. *Liwa, Al Dhafra – Strategic Water Storage and Recovery System.* Abu Dhabi, United Arab Emirates: Dar Al-Handasah.

Dawoud, M. A., 2012. *Strategic Water Reserve: New Approach for Old Concept in GCC Countries.* Abu Dhabi, United Arab Emirate, Environment Agency Abu Dhabi (EAD).

Greenberg, M. W., 2013. *A Step-Wise Approach to Elicit Triangular Distributions.* New Orleans, Louisiana, USA, NASA Cost Analysis Division, p. 16.

Hashimoto, T., Stedinger, J. R. & Loucks, D. P., 1982. Reliability, Resiliency, and Vulnerability Criteria for Water Resource System Performance Evaluation. *Water Resources Research,* 18(1), p. 14–20.

Heasler, P. G. & Simonen, F. A., 1990. *A Probabilistic Evaluation of Alternative Weld Inspection Sampling Plans.* College Park, Maryland, USA, IEEE.

Hemming, V. et al., 2018. A Practical Guide to Structured Expert Elicitation Using the IDEA Protocol. *Methods in Ecology and Evolution,* 9(1), pp. 169-180.

ISO 28765, 2022. *ISO 28765:2022. Vitreous and Porcelain Enamels — Design of Bolted Steel Tanks for the Storage or Treatment of Water or Municipal or Industrial Effluents and Sludges,* Geneva: International Organization for Standardization (ISO).

Jones, E. et al., 2019. The State of Desalination and Brine Production: A Global Outlook. *Science of the Total Environment,* Volume 657, pp. 1343-1356.

Jünger, S. et al., 2017. Guidance on Conducting and REporting DElphi Studies (CREDES) in Palliative Care: Recommendations Based on a Methodological Systematic Review. *Palliative Medicine,* 31(8), p. 684–706.

Mano, A., Katsuyama, J. & Li, Y., 2020. Influence Evaluation of Sampling Methods of the Nondestructive Examination on Failure Probability of Piping Based on Probabilistic Fracture Mechanics Analysis. *Mechanical Engineering Journal,* 7(3), p. 00567.

MEED, 2026. *Saudi's Strategic Surge in Water Transmission and Storage PPPs,* Dubai Media City, Dubai, United Arab Emirates: MEED Media FZ LLC.

Mohieldeen, Y. E., Elobaid, E. A. & Abdalla, R., 2021. GIS-based Framework for Artificial Aquifer Recharge to Secure Sustainable Strategic Water Reserves in Qatar Arid Environment Peninsula. *Scientific Reports,* 11(1), p. 1–13.

Newcombe, R. G., 1998. Interval Estimation for the Difference Between Independent Proportions: Comparison of Eleven Methods. *Statistics in Medicine,* 17(8), pp. 873-890.

Ramadan, A., 2026. *How Much Radiography Is Enough? Weld Porosity, Inspection Coverage and Confidence in Strategic API 650 Potable-Water Storage Tanks,* s.l.: Research Square Preprint.

Sotoodeh, K., 2024. Testing and Examination of Welding on Storage Tanks. In: 1, ed. *Storage Tanks Selection, Design, Testing, Inspection, and Maintenance: Emission Management and Environmental Protection.* s.l.:Elsevier, pp. 227-251.

SWA, 2026. *Saudi Water Authority (SWA). Saudi Water Authority and Vision 2030,* Riyadh, Saudi Arabia: Saudi Water Authority.

United States Department of Defense (DoD), 2009. *MIL-HDBK-1823A: Nondestructive Evaluation System Reliability Assessment.* Revision A ed. Washington, DC, USA: Department of Defense (DoD).

Vose, D., 2008. *Risk Analysis: A Quantitative Guide.* 3rd ed. Chichester, England; Hoboken, New Jersey, USA: John Wiley & Sons.

WaterWorld, 2018. *World’s Largest Man-Made Desalinated Water Reserve Completed in Abu Dhabi.* Tulsa, OK, USA: WaterWorld.

WTCO, 2025. *Project Development and Strategic Storage Programme,* Riyadh, Saudi Arabia: Water Transmission Company.

Supplementary material

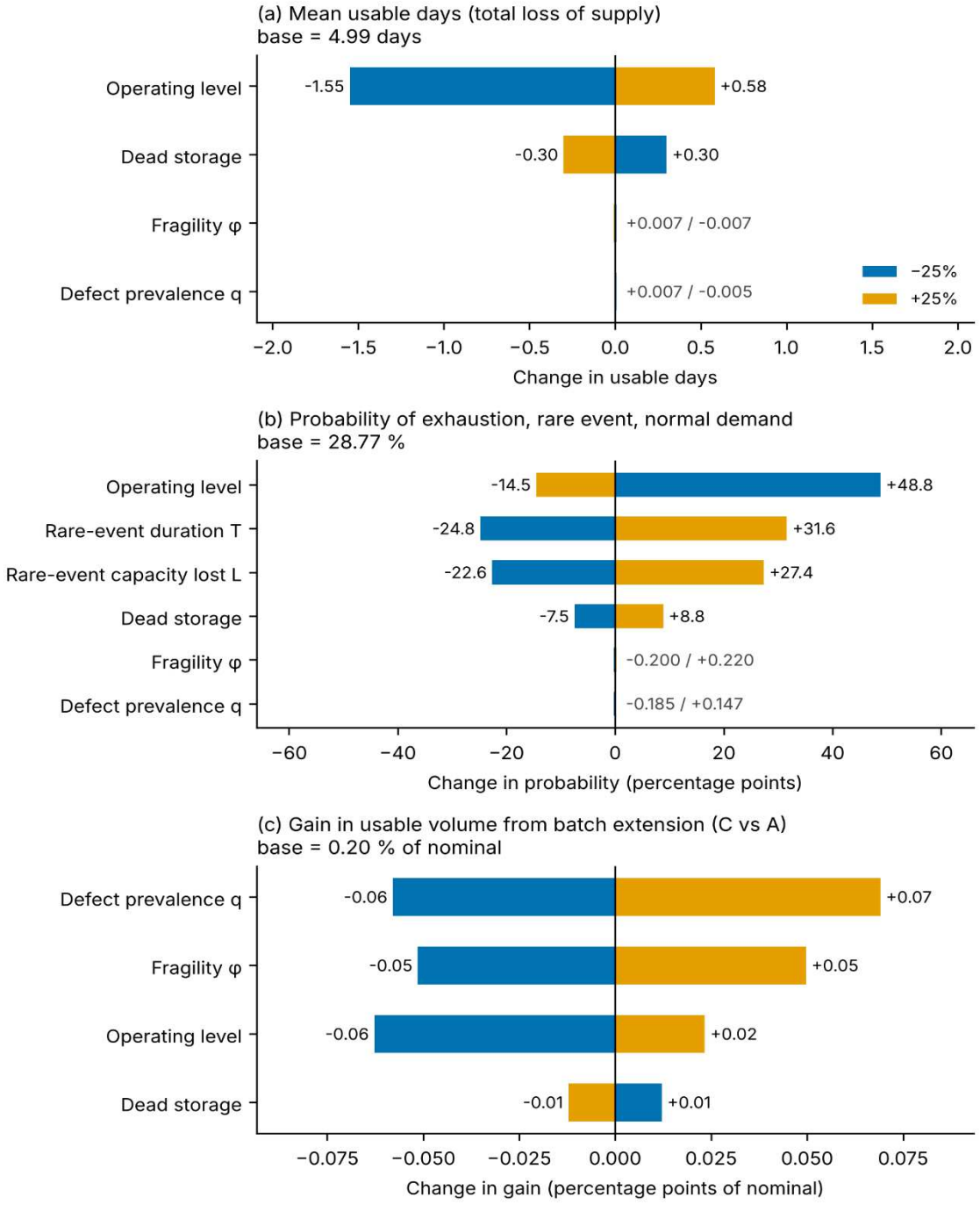


Figure S1. One-at-a-time sensitivity (±25%) of (a) mean usable days, (b) the probability of exhaustion in a rare event at normal demand, and (c) the gain in usable volume from batch-triggered extension, to dead storage, operating level, defect prevalence q and fragility φ (rare-event duration and capacity lost are added in panel b). Operating level is capped at 100%. Blue: −25%; orange: +25%.

Table S1. Numerical results behind Figure S1 (outcome values at −25% and +25% of each input; base case: seven nominal days, regime C).

Outcome	Parameter	Base	−25%	+25%
Mean usable days	Operating level	4.993	3.444	5.573
Mean usable days	Dead storage	4.993	5.293	4.693
Mean usable days	Fragility φ	4.993	5.000	4.986
Mean usable days	Defect prevalence q	4.993	5.000	4.988
Mean usable days	Rare-event duration T	4.993	4.993	4.993
Mean usable days	Rare-event capacity lost L	4.993	4.993	4.993
P(exhaustion), rare event (percentage points)	Operating level	28.765	77.591	14.241
P(exhaustion), rare event (percentage points)	Rare-event duration T	28.765	3.959	60.355
P(exhaustion), rare event (percentage points)	Rare-event capacity lost L	28.765	6.171	56.124
P(exhaustion), rare event (percentage points)	Dead storage	28.765	21.285	37.611
P(exhaustion), rare event (percentage points)	Fragility φ	28.765	28.565	28.985
P(exhaustion), rare event (percentage points)	Defect prevalence q	28.765	28.580	28.913
Gain from batch extension (% of nominal)	Defect prevalence q	0.202	0.144	0.271
Gain from batch extension (% of nominal)	Fragility φ	0.202	0.151	0.252
Gain from batch extension (% of nominal)	Operating level	0.202	0.139	0.225
Gain from batch extension (% of nominal)	Dead storage	0.202	0.214	0.190
Gain from batch extension (% of nominal)	Rare-event duration T	0.202	0.202	0.202
Gain from batch extension (% of nominal)	Rare-event capacity lost L	0.202	0.202	0.202