Global Academic Journal of Economics and Business

Available online at https://www.gajrc.com
DOI: https://doi.org/10.36348/gajeb.2026.v08i06.001

OPEN ACCESS



ISSN: 2706-9001 (P)
ISSN: 2707-2584 (O)

Original Research Article

Safety Assurance and Operational Resilience in Water Treatment Infrastructure: A Risk-Based Framework for Saudi Arabia's Vision 2030

Mohammad Anwar Hossain Mojumder[1]*
[1]HSE Manager (HSE Department), Fisia Itlimpianti S.p.A: Olaya District, Eastern Province, SA

***Corresponding Author**
Mohammad Anwar Hossain Mojumder
HSE Manager (HSE Department), Fisia Itlimpianti S.p.A: Olaya District, Eastern Province, SA

Article History
Received: 09.09.2026
Accepted: 02.11.2026
Published: 03.11.2026

Abstract: Water treatment infrastructure in Saudi Arabia is of course a vital public health function and is equally important in enabling industrial development, urban growth and the diversification of water supplies. You can't judge the reliability of that infrastructure on its capacity alone. An apparently continuous supply of water of poor quality could be a far more serious failure than a short-lived, isolated interruption. This is a critical integrative review of how safety assurance can be combined with operational resilience across the full range of water treatment infrastructure: desalination, municipal treatment, wastewater reuse and the transmission systems that link them. It draws on 30 peer-reviewed papers from 2020 to 2025 and uses a risk pathway approach to look at water quality, occupational exposure, cyber-physical interdependencies, energy supply, climate stress and governance. It makes a clear distinction between preventative barriers and recovery capabilities and identifies the weak links between technical monitoring, organisational accountability and capital decisions. A six domain assurance framework is proposed which links hazard characterisation, barrier verification, scenario based resilience planning, operational response, independent challenge and continuous learning. The framework is based on the evidence of barrier effectiveness rather than simply the existence of procedures and allows for a differentiated response to failures which could have irreversible health consequences. Its application in Saudi Arabia will require integration across the operators, the regulators, the contractors and the various infrastructure interfaces with clear service level thresholds and tested contingency arrangements. The outcome is a practical, auditable basis for aligning infrastructure investment and water sector transformation with the protection of public health and the provision of reliable service. This is a conceptual synthesis and not an empirical validation of plant performance.

Keywords: Water Treatment Infrastructure, Water Safety Assurance, Operational Resilience, Desalination, Municipal Water Treatment, Wastewater Reuse, Water Quality Management.

Copyright © 2026 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

Citation: Mojumder, M. A. H. (2026). Safety Assurance and Operational Resilience in Water Treatment Infrastructure: A Risk-Based Framework for Saudi Arabia's Vision 2030. *Glob Acad J Econ Buss, 8*(6), 1949-1959.
https://doi.org/10.36348/gajeb.2026.v08i06.001

1 INTRODUCTION

Saudi Arabia's water infrastructure is in a very peculiar position in terms of national development. Water security is predicated on a combination of sophisticated treatment technology, significant energy inputs, vast conveyance networks and an ever more important network of reuse. Simply increasing the volume of water available is therefore essential but not enough. The infrastructure has to be able to maintain treatment quality under a range of conditions including variable source water quality, mechanical deterioration, worker exposure, cyber incidents and the potential for disruption to propagate across the network. Recent Saudi research has highlighted the ongoing problems with wastewater collection and reuse, and drinking water assessments have drawn attention to the issue of quality control during desalination, blending and over long distances of transmission [1–3]. The focus of analysis has to shift from the individual plant to the reliability of the whole water service chain.

An operationally resilient system is one that can absorb a disturbance, limit the damage, adapt its configuration and restore essential functions to within acceptable limits. Restoration of function alone is not necessarily the definition of acceptable performance in the context of drinking water or wastewater systems. A treatment plant could recover very quickly and yet fail to demonstrate microbiological, chemical or process integrity and so could be transferring risk to the customer, the workforce or the receiving environment. On the other hand a protective shutdown may be an example of successful safety performance even if it has a negative impact on the availability of service. Work on the resilience of water distribution systems has shown that the indices which are commonly used to assess resilience tend to describe the characteristics of the network rather than its resilience to all possible disturbances [4–7]. Safety assurance therefore has to define what is an acceptable level of recovery and how quickly that recovery must occur.

The National Transformation Agenda adds another layer of complexity. The desalination projects are evolving in their technology, the use of reclaimed water is expected to become increasingly significant and new facilities are increasingly being designed with connected instrumentation and automated control. This leads to improved efficiency but also introduces a number of dependencies on specialist maintenance, validated sensors, secure communications and reliable power. The operating environment is also one of heat, coastal exposure and occasional heavy rain which can all have an impact on site access, the electrical distribution and the characteristics of the source water. These are all plausible risk pathways which will need to be assessed on an asset by asset basis rather than the suggestion that every facility in Saudi Arabia is subject to the same exposure or will suffer at the same frequency [2-10]. General guidance will therefore need to be translated into specific, documented engineering decisions.

The existing literature is of considerable value but very fragmented. Water safety planning is all about risk based preventative management. Desalination research looks at treatment performance and energy optimisation. Cyber security research looks at threats to the control system. Resilience research looks at continuity and network robustness. The integration is poor at the very points where a risk threshold should lead to a change in operating permission, change in maintenance priority, change in procurement criteria or change in emergency response. The failure of formally established safety plans is a clear illustration of the limitations of having paper processes without effective implementation [8, 9]. The issue we're trying to address here is how technical and organisational evidence can be brought together into an assurance regime which is credible before, during and after a disruption.

This is a risk based architecture rather than a single numerical score. The core of the argument is that public health protection, workforce safety and continuity of supply all require different but related decision rules. The architecture that results makes barrier ownership, performance evidence, degraded mode operation and recovery authorisation all explicit. It was developed in the context of the Saudi water infrastructure but the underlying logic could be applied to other highly interdependent, safety critical utilities where the successful restoration of service is clearly distinct from a return to service which is unsafe.

2 Aim, Objectives and Analytical Questions

The study is aimed at developing a defensible assurance architecture for water treatment infrastructure that balances the need to prevent unacceptable harm with the need to be able to maintain and restore critical water services in the face of disturbance. The first objective is to identify the key safety and continuity risks at the treatment and transmission interfaces. The second is to assess the strengths and weaknesses of current risk assessment, digital monitoring, engineering redundancy and emergency management. The third is to use that evidence to develop a staged governance model which is appropriate to the Saudi institutional and environmental context. The fourth is to define the indicators and validation priorities that will allow us to distinguish between documented compliance and actual resilience.

Three analytical questions form the basis of the synthesis: Which combinations of hazard and dependency failure are most likely to lead to a failure of safe service? What evidence would be required to demonstrate that the critical barriers are in fact performing as required under both normal and degraded conditions? How should operators and the regulators decide when it is appropriate for an affected facility to either continue to operate or to return to service? These are questions which deliberately separate out the research aim from the operational implementation objectives and thus prevent the framework from becoming a simple list of technologies.

3 Review design and Evidence Synthesis

A critical integrative review was chosen as the research question spans a number of technical disciplines whose individual studies report on very different outcomes. Desalination papers will typically report on energy intensity, membrane performance or cost; occupational papers on exposure; water safety papers on hazard controls; network papers on loss of service; and cybersecurity papers on attack detection. To pool those metrics would imply a level of methodological comparability that the body of literature simply does not allow for. The integrative review instead looks at the mechanisms, the assumptions and the transferability, while recognising the differences between what is established evidence and what is proposed as governance.

The literature base consists of 30 identified scholarly sources from 2020 to 2025. The search was focused on the records of scholarly publishers and bibliographic records using a combination of: water treatment, drinking water safety planning, desalination, wastewater reuse, water infrastructure resilience, risk assessment, occupational exposure, automation, cybersecurity, Saudi Arabia and energy continuity. Titles, dates and DOI's were cross checked against the published records. The water security paper was the basis for the treatment of storage and continuity and the other supplied papers informed the level of detail that was included without them being considered as out of period scientific evidence [4]. This is a purposive and thematic selection rather than an exhaustive list.

Eligibility was based on relevance to at least one safety or resilience mechanism, the availability of publication metadata, a publication date within the required time frame and a persistent DOI. Any research which was solely about unrelated products or general sustainability claims was not included in the evidential basis. The evidence was grouped into six analytical domains: source and process integrity; human exposure; physical and energy dependencies; digital control; distribution and reuse; and governance and verification. Each contribution was assessed in terms of the hazard in question, the observed or modelled result, the scope conditions, the limitations and the potential assurance implications. Conceptual recommendations were only retained where the mechanism behind them could be clearly explained.

Interpretation was deliberately conservative. Work done elsewhere was treated as purely mechanistic or methodological and not as a direct estimate of the incidence in Saudi Arabia. The pilot simulations and testbed results were not presented as evidence of the outcome in a full scale scenario. There were no interviews, no operator audits, no access to the confidential incident data, no statistical meta-analysis and no field validation. The framework proposed is an analytic synthesis to be tested against operational records and against controlled exercises. This level of transparency precludes any overclaiming and makes the subsequent evaluation more reproducible [6-12].

4 Saudi Water-System Context and the Assurance Boundary

A Saudi assurance framework will require a well defined boundary. The system in question would start with intake of seawater, groundwater or wastewater and would then proceed through pretreatment, primary separation, disinfection, storage, pumping, blending, transmission, distribution and receiving or reuse. That is the functional boundary and it is distinct from the plant's contractual boundary. A problem may well be generated upstream of the plant and the effects may only be apparent downstream. Drinking water quality research has demonstrated the importance of mixing and conveyance in the operational context and wastewater research has similarly shown that the treatment investment must be considered in conjunction with collection access, reuse demand and distribution arrangements [1–3]. The risk register therefore needs to reflect the interfaces rather than just the machinery within the fence.

Desalination is a textbook example of the need for whole system thinking. The performance of a membrane is a function of the feed water variability, the quality of the pretreatment, the operating pressure, the cleaning regime and the post-treatment conditioning. A desalination train may well produce perfectly acceptable permeate with the upstream biological fouling leading to a massive increase in maintenance exposure, or a single fault in the post-treatment dosing could have a destabilising effect on the water later in the network. There are a multitude of options available in the way of pretreatment and desalination technology as evidenced by the

numerous reviews of the subject, but the performance of any given system is highly conditional upon the feed composition, the energy available and the operating regime [13, 14]. Procurement should therefore be based on validated operating envelopes and demonstrable failure management rather than on nominal rated capacity.

Energy is not just an operating cost in this scenario. A power failure can stop the pumps, upset the pressure balance, disable the instrumentation and generate restart transients. An emergency generator will not automatically provide treatment continuity in the event of a power failure if transfer switching doesn't work, there's no fuel available, the process isn't stable or there's no residual disinfection. Reviews of renewable integration and desalination optimisation have clearly identified the potential for significant efficiency gains and diversification, albeit with an appreciation of the variability and storage issues [15–20]. For the purposes of safety, the energy transition and resilience planning must be co-designed so that the move to lower emissions doesn't in any way compromise the minimum level of service that is required in the event of a critical contingency.

Wastewater reuse brings with it a whole new set of issues relating to responsibility. A perfectly functional treatment plant can still result in exposure to workers or end users if the sampling, storage segregation, pipework, access controls and end use restrictions are not adequate. The adequacy of a barrier is therefore dependent on its downstream purpose. The quality objectives and monitoring requirements for non-potable irrigation, industrial substitution and potentially high exposure uses will all be different. Sector studies in Saudi Arabia do point to the potential for increased reuse but do not in any way suggest that all recovered water can be treated as being of equal quality [1-3]. There needs to be a clear cut end use and a documented response to off-spec effluent prior to any significant expansion of reuse.



Figure 1: System boundary for safety assurance across water treatment and delivery. Original conceptual synthesis

5 Safety Assurance as a System of Verified Barriers

Safety assurance is much more than just checking that you have policies, training records and operating instructions. It's a structured process of demonstrating that the controls in place against a specified hazard are appropriate, are in fact being implemented and are effective over time. In the case of drinking water infrastructure a defensible safety claim should state the hazard, the control mechanism, the evidence of the control's current state and the action to be taken if that evidence is absent. Water safety research has shown that poorly constructed risk matrices can actually hide the very uncertainty they are meant to address and that documented implementation can fail even in the face of apparent compliance [8, 9]. Assurance should therefore start with explicit safety claims and verifiable acceptance criteria rather than with some unqualified 'low/medium/high' risk rating.

A barrier model separates the preventive, detective, mitigating and restorative controls. Pretreatment and membrane integrity can prevent transfer of the contaminant; calibrated water quality measurement can detect a departure; isolation and diversion can mitigate customer exposure; and validated restart protocols can restore compliant service. You don't have one category of control replacing the others. A late alarm won't prevent exposure in the past and a backup pump won't fix a dosing error which has been replicated across the parallel treatment trains. The value of climate sensitive source water screening and natural hazard analysis is in demonstrating how an external disturbance can circumvent controls which would otherwise be very effective [10, 11].

The quality of the evidence is important. Routine readings must be traceable to calibration, sampling, lab verification and change controlled instrumentation. Alarm performance must be assessed in terms of the actual consequences of false positives and missed events. Critical valves, bypass arrangements and isolation interlocks should be tested to prove they work where it is technically feasible. Where a safeguard is not available then operators should employ pre-defined operational limitations, enhanced supervision or a controlled shutdown rather than just quietly accepting a lower safety margin. The key point is the difference between a safeguard that is assumed to be in place and one that is in fact supported by the current, relevant evidence.

Occupational protection needs the same rigour but the exposure pathways are different. Wastewater work can give rise to intermittent H2S releases as well as mechanical, chemical and biological hazards. Work in sewers has been the subject of a number of studies which have highlighted the potential for episodic exposures to be missed by conventional sampling [21–23]. The requirements for assurance in this context are therefore very much the same as for any other type of work: confined space permits, atmospheric testing, ventilation, contractor competence, emergency rescue arrangements and verification of isolation prior to any maintenance work. Odour control should not be taken as evidence that the atmosphere is safe. The maintenance and recovery phases are of particular importance as the non-routine nature of the work may well invalidate the assumptions which underpin the normal process risk assessment.

Institutional independence will enhance these technical controls. The plant team will still be responsible for safe operation and a separate assurance or audit function will be required to verify that the evidence supports the claims and that the corrective actions are actually closed. Independent challenge doesn't mean some remote compliance check which has no connection to the engineering reality. The reviewers will need access to plant records, deviation histories and to competent technical advice and any unresolved high consequence exceptions will have to get to the people who have the authority to commit resources and to grant or withhold operating permissions. This is particularly important in a situation where there may be a pressure to minimise outages in order to meet production targets.

6 Operational Resilience and Compound Disturbances

Operational resilience should be defined in terms of a particular essential function, duration and consequence threshold. Production volumes, distribution pressure and time to restore flow are all useful metrics but none of these in isolation will tell you that the water you are delivering is safe. A meaningful assessment of resilience will need to consider at least four aspects: the ability to anticipate stress, the ability to maintain a safe minimum service, the ability to contain the escalation of failure and the ability to regain verified performance. Reviews of distribution metrics have highlighted the danger of assuming that a single engineering metric will capture all of these aspects [6, 7]. For water infrastructure in particular, a resilience assessment which ignores health based constraints will be one which can reward unsafe continuity.

Physical hazards often compound through shared dependencies. Heavy rain can restrict site access, flood electrical rooms and degrade raw water quality. Heat can put equipment under stress and increase demand. A prolonged external power failure could in turn deplete the local backup. There is international case study evidence of the relevance of flood protection and adaptive wastewater infrastructure, but that should not be simply transplanted to the Saudi climate and institutional context [24]. The Saudi operators will need site specific hazard maps, flood levels, critical equipment elevations, supply chain assumptions and recovery scenarios which take into account both the rare high impact events and the more chronic degradation.

Cyber-physical disruption provides another, less obvious, pathway of dependency. Today's plants are highly dependent on sensors, PLCs, historian databases and supervisory systems to coordinate chemical dosing, pumps, membranes and distribution. An attack or configuration error can corrupt the measurements or commands and yet leave significant parts of the physical plant still appearing to be operating. Reviews of machine learning monitoring have demonstrated some useful potential for anomaly detection, though the predictive capability is still entirely dependent on the quality of the training data and the quality of the sensors under the conditions in which they are operating [25, 26]. A digital dashboard should therefore be used to confirm, not to replace, the operator's understanding of the process and the independently verified critical measurements.

Cybersecurity and process safety need to be linked at the incident command level. Work on cyber-physical systems and critical infrastructure in the water sector has shown that an alarm could be due to instrument failure, adversarial manipulation or a real process excursion [27, 28]. The appropriate initial response is to protect the physical process, maintain trustworthy observability and prevent any unsafe

commands; the root cause can then be investigated. Detection studies and experimental water treatment testbeds have shown the potential of model-based and data-driven approaches but that doesn't mean we can assume perfect identification in an actual operating plant in Saudi Arabia [29]. Manual fallbacks must be used, documented and restricted to those configurations which will maintain the essential protective functions.

Compound events are more important than individual piece of equipment failures. The combination of poor feed water quality, loss of a treatment train and a delay in the arrival of spare parts could well exceed the assumptions that underpin the concept of nominal redundancy. A power cut and subsequent hurried restart is another scenario which can lead to water quality problems at precisely the time when you're under the most pressure to get things back to normal. Scenario analysis should be used to expose common cause failures, dependence on a single specialist contractor, the existence of inaccessible valves and the uncertainty around the time it will take to get the process back to equilibrium. Resilience investment should then be directed at the actual failure pathways that have been demonstrated rather than simply being driven by a general preference for duplication.

7 Proposed Risk-Based Safety and Resilience Framework

The proposed Safety Assurance and Operational Resilience Framework consists of six interlinked domains: system definition, risk characterisation, barrier assurance, response and continuity planning, independent governance and learning. They should be seen as forming a continuous assurance cycle rather than as a series of discrete documents completed at commissioning. Figure 1 shows the physical and organisational interfaces to be addressed and Figure 2 the feedback loop from risk evidence to operating decisions. The framework builds on water safety planning by making resilience thresholds and independently verified return to service criteria an essential part of the assurance process rather than simply being treated as separate elements of emergency management.



Figure 2: Six-domain cycle linking risk information, operating decisions and independent assurance. Original conceptual synthesis

Domain 1 is about establishing the service model. The operator will define the water source, the treatment path, the distribution dependencies, the permitted end uses and the health protection obligations. Each of the critical functions will have an accountable owner, a minimum performance requirement and a set of upstream dependencies. A plant level statement such as 'maintains safe drinking water production during a utility interruption' is not good enough without specifying what treatment trains are available, for how long the interruption can be tolerated, what the water quality evidence must be and under what circumstances isolation is required. The boundaries will also identify the contractors and the inter-operator handovers where the responsibility changes but the risk remains.

Domain 2 builds a scenario based risk picture. The traditional approaches of likelihood and severity are still relevant for prioritisation but should be augmented with barrier analysis, dependency mapping and explicit uncertainty. A low frequency

event with potentially very severe and irreversible health outcomes should not be subsumed by a single composite score. Each material scenario must have a credible initiating condition, plausible escalation pathways, the relevant users or workers impacted and a justification for the controls that are assumed to be independent. Lane and Hrudey's critique of risk matrices underlines the importance of having defensible category definitions and the natural hazard methods provide a framework for a systematic assessment of the impact of external disturbances [8-11].

Domain 3 is about verifying the safeguards. The barrier register should contain the purpose of the control, the operating envelope, the proof test method, the monitoring frequency, the evidence owner and the failure response rule. For a membrane integrity control the evidence could be differential pressure trends, calibrated instrumentation, lab confirmation and maintenance records. For confined space work the evidence could be a permit audit, gas test validity and rescue drill results. Barrier effectiveness should be reviewed following a modification to the equipment or a significant process change, rather than just as part of the annual certification. Table 1 takes a number of representative failure pathways and translates them into testable controls and decisions.

Table 1: Illustrative barrier-assurance register for representative water-sector scenarios

Exposure pathway	Critical barrier	Verification evidence	Decision if control is impaired
Off-specification treated water	Quality limits, sampling and automated isolation	Calibrations, laboratory checks, release records	Divert flow; withhold release pending verified quality
Chemical or H_2S exposure during maintenance	Isolation, atmosphere testing, ventilation and rescue readiness	Permit sampling, detector checks, practical drills	Suspend entry; restore controls before authorisation
Loss of electrical supply	Transfer protection, standby generation, safe shutdown logic	Load-bank test, fuel inventory, transfer records	Enter restricted mode or stop treatment safely
Manipulated process measurement	Secure control, diverse sensing and manual validation	Access logs, configuration audits, functional challenge	Isolate suspect signals; enforce validated safe state
Flooded access or electrical room	Asset elevation, drainage, equipment segregation	Site surveys, inspections and scenario exercise	Reconfigure access and preserve critical protective loads

Domain 4 is all about continuity around safe operating states. The defined modes could be normal production, restricted operation, isolated storage supply, safe shutdown and validated restart. The criteria for transitioning between modes should be based on measured conditions rather than on the subjective judgement of a manager. Where an alarm indicates a possible loss of disinfection control the appropriate resilience action may be rapid diversion or a controlled shutdown even at the cost of reduced throughput. Minimum service and maximum tolerable interruption should be specified by user category with particular attention to hospitals and other critical customers. The plans will require drills that cover communications, staffing, backup power and water quality release criteria.

Domain 5 is about governance and independent challenge. The process is managed by the frontline operating staff, with engineering and risk specialists ensuring standards are met. Assurance reviewers then assess the adequacy of the evidence and the need for escalation, and senior management ultimately accept or reject the residual exposure within their authority. There should be formal closure and time limits for any severe exceptions. Risk acceptance in no way should be used to replace the mandatory water quality or occupational protections. Capital reviews should require evidence that the proposed investment will lead to a reduction in a particular consequence pathway and procurement should take into account maintainability, spare availability, cyber support and commissioning tests [12-30].

Domain 6 is about turning events into organisational learning. Near misses, adverse process trends, drill failures, cyber alerts and non-conformities should all be investigated for both root cause and for any weaknesses in the barriers. The actions taken to rectify the issue must be linked to the appropriate design standard, training, maintenance interval or contingency assumption. Assurance confidence should be reduced if calibration is overdue, an exercise is missed or a change is untested, even in the absence of any recorded incident. That's a very dynamic view of the situation which is quite different to the notion that the fact that

we've had no incidents recently is proof that we're going to be safe in the future.

8 Implementation, Indicators and Decision Rights

Implementation should start with a pilot in one representative treatment or reuse corridor rather than a full sector wide scoring exercise from day one. The pilot should involve asset and dependency mapping, identification of credible high consequence scenarios and the creation of a barrier register from the existing procedures and maintenance information. Management can then pick a small number of decision relevant measures, establish a baseline quality and test the escalation rules via tabletop and controlled operating exercises. The lessons learned from the pilot will then tell you what can be standardised across sites and what has to be site specific in the light of source water chemistry, plant design, customer exposure and local hazard conditions.

A good set of indicators will be able to tell the difference between leading assurance signals and lagging outcomes. The leading indicators would be the % of critical barriers within proof test interval, completeness of calibration records, number of unresolved high consequence deviations, success of backup power transfer tests, validated cyber recovery capability and contractor competence coverage. The lagging indicators would be off spec releases, occupational exposure events, unplanned loss of safe service and recovery time. The indicators should be broken down by consequence severity; you shouldn't have a large number of minor alarms masking one unacceptable event which could have a significant impact on health. Table 2 sets out some practical measurement rules and escalation triggers, without making any claims about the actual performance of the system.

Table 2: Proposed evidence-led indicators for a pilot assurance programme; triggers are decision rules, not national benchmarks

Assurance dimension	Suggested metric	Evidence source	Illustrative trigger for review
Safety-critical barriers	Share with current proof-test evidence	Barrier register and test logs	Overdue or failed verification
Water-quality release	Confirmed off-specification events and closure time	Laboratory and supervisory records	Release without satisfactory confirmation
Workforce protection	Permit and gas-monitoring conformity	Permit audits and detector logs	Entry under impaired controls
Continuity readiness	Verified duration of safe minimum service	Transfer drills and storage tests	Exercise fails required service state
Cyber-physical recovery	Time to trustworthy manual or restored control	Incident exercise and configuration records	Unvalidated control restoration
Action accountability	Age and severity of unresolved exceptions	Corrective-action and governance records	High-consequence exception lacks accountable closure

A transparent decision rights architecture is just as important. The operating team must be able to stop unsafe activity at any time. The technical authorities must be able to validate the exceptions and the conditions for resumption. The incident commander must be able to co-ordinate the impact of multi-asset disruption. The oversight bodies must be able to challenge the persistent existence of unresolved risk. No performance target should be in any way incentivising the suppression of negative findings. A defensible return to service gate requires evidence that the initiating conditions have been controlled, the protective barriers are working, water quality is confirmed where it matters and the operators can maintain the restored state without any need for exceptional improvisation.

Commercial arrangements should be in place to underpin these decisions. The contracts should define the arrangements for access to test data, access to safety critical spares, update support for the control systems, incident notification and the obligations in the case of a prolonged disruption. The procurement evaluation should take into account the degree of proof testability, maintainability, the quality of the interoperable records and the recovery capability as much as it does capacity, cost and energy efficiency. Where renewable power or hybrid desalination is being introduced the contracts should be designed to test the continuity of supply under realistic transfer and dispatch conditions rather than simply assume that the diversity of the technology will in itself provide resilience [15–20]. Governance will only be effective if the risk evidence is capable of changing the specifications, the budgets or the operating permissions.

Assurance evidence also requires an information architecture. A key requirement is that there be an identifier which links the design requirement, the instrument record, the maintenance activity, the calibration history, the observed

deviation and the management approval. This is particularly important when you have different companies operating the desalination, transmission and distribution assets as a certificate issued for one component may not in itself prove that the protection pathway is intact. Periodic reviews should select samples of the evidence and re-trace the decisions from the initial alarm through to isolation, laboratory confirmation, the corrective work and the release. Missing timestamps, unexplained overrides and delayed action closure are all potential weaknesses. Where commercial confidentiality prevents information from being shared then the interfaces should still define the minimum safety data, access rights and notification intervals that are required to enable the responsible decisions to be made. This traceability can be of great assistance in terms of speeding up the assurance process during an incident and also enables an independent reviewer to see the difference between robust controls and those which are merely reassuring in terms of documentation. The evidence should be available to the decision makers as the operational situation changes.

9 DISCUSSION AND RESEARCH IMPLICATIONS

The evidence is that we are moving from an asset centric view of reliability to a safety conditioned view of service resilience. That resolves the apparent contradiction between the need for uninterrupted production and the need for protective shutdown. If safe delivery cannot be demonstrated then an interruption may in fact be the correct thing to do to avoid a much bigger public health failure. The right optimisation problem is therefore not to maximise uptime in isolation but to maintain or restore the essential service within the safety constraints that are enforceable. Operational targets should be understood in that hierarchy, particularly where the output targets will be used to determine contractor payments, plant performance reviews and investment case justification.

The approach outlined also provides an explanation for why digitalisation can both improve and compromise assurance. The increased amount of connected measurements provides greater situational awareness and enables faster analysis. The increased reliance on automation however creates new potential for shared failure. Experimental attack detection accuracy cannot simply be transplanted into a heterogeneous set of operating systems and the presence of an analytic anomaly in no way establishes contamination or equipment damage [25–29]. Trustworthy resilience will require diversity in the sensing, secure configuration, the judgement of the experienced operator and clear authority to revert to a validated safe state. The value of digitalisation is in enabling better defensible decisions rather than in the creation of more dashboards.

A further point is about institutional co-ordination. Water quality risk moves across the production, transmission, reuse and customer interfaces and formal accountability may be split between the asset owner, the contractor and the various oversight bodies. A single assurance case would make those interdependencies visible without in any way obliterating the legitimate differences between the requirements for drinking water and wastewater. It could be used to support Saudi's strategic objectives by tying in safe reuse, energy diversification and service continuity to specific evidence requirements. The review does not however show that any existing national programme has reached the proposed level of maturity or that one particular regulatory regime is in fact the best in all circumstances.

A third point is methodological. Future work should assess the framework using longitudinal operational data and controlled scenarios rather than confidence scores alone. Appropriate research would involve before and after assessment of barrier verification, cross site comparisons adjusted for site configuration and tests of whether the escalation rules actually lead to a reduction in the number of unsafe releases or recovery errors. The outputs of such work should include a statement of uncertainty, near misses and any unintended consequences. It would also be useful to look at how safety decisions are actually made when there are competing pressures from the perspective of water production, financial efficiency and customer continuity; purely technical models are not going to be able to capture the full range of organisational mechanisms at play.

Risk communication is part of continuity planning. The public need practical, timely and relevant information that reflects the actual exposure pathway. A vague announcement may do nothing to improve the situation and could actually erode public confidence. The plan must identify which organisation has the authority to impose precautionary restrictions, what evidence would be required to lift those restrictions and how conflicting messages from suppliers and downstream agencies would be resolved. Essential facilities will need pre-arranged alternatives which won't simply assume that the road will be open and that there will be an unlimited number of tankers available. Public communication exercises should include a test of the clarity of the message and the effectiveness of the communication channels in the event of a telecommunications failure. Worker communications are just as important and particularly so for contractors who are arriving in during the recovery

phase and who may not be familiar with the new isolation boundaries. These social and organisational dependencies are not captured by pump based reliability models but can be the decisive factor in whether a technically correct response actually protects people.

10 LIMITATIONS AND CONCLUSION

This is all constrained by the purposive nature of the literature we've selected, the diversity of study designs and the lack of access to Saudi plant level incident, audit and performance data. The 30 papers we've referenced do identify relevant mechanisms but don't give us a national risk ranking or any validated quantitative effect sizes. There are huge differences in treatment technology, regulatory practice and workforce capability which means that international examples cannot be readily transferred. The conceptual framework has not been field tested so the indicators and decision gates should be viewed as design propositions which would need to be verified in the local context rather than as established performance standards.

Safety assurance and operational resilience are interdependent but not synonymous. The most robust water infrastructure won't just survive a disruption, it will demonstrate that every permitted operating state (including recovery) maintains the requisite level of protection for workers, the public and the environment. The framework we propose links risk scenarios to verifiable barriers, safe service thresholds, independent oversight and a culture of continuous learning. For Saudi Arabia's Vision 2030 it is the practical application of this that is of value in ensuring that the modernisation of its infrastructure is demonstrably safe and reliable. Future work should be directed at testing these relationships in a range of treatment configurations and publishing the results in a transparent manner on the reliability of the barriers, the recovery decisions and the resultant public service outcomes.

REFERENCES

1. Alhajri MF, Jamil R, Abdalla EME, et al. (2025). Status of wastewater treatment and reuse in Saudi Arabia and its forecasting for the year 2035. Applied Water Science, 15:167. https://doi.org/10.1007/s13201-025-02484-4
2. Al-Hamzah AA, Fellows CM. (2024). Drinking water quality in the Kingdom of Saudi Arabia. Water, 16:1810. https://doi.org/10.3390/w16131810
3. Dawoud MA, Ewea HA, Alaswad SO. (2022). The future of wastewater treatment and reuse in Kingdom of Saudi Arabia. Desalination and Water Treatment, 263:127–138. https://doi.org/10.5004/dwt.2022.28217
4. Alao JO, Eze SU, Onyenweife GI, et al. (2025). Enhancing water security through integrated storage mechanisms and rainwater harvesting for sustainable development. Discover Sustainability, 6:941. https://doi.org/10.1007/s43621-025-01914-2
5. Mortula MM, Ahmed MA, Sadri AM, et al. (2020). Improving resiliency of water supply system in arid regions: integrating centrality and hydraulic vulnerability. Journal of Management in Engineering, 36(5). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000817
6. Lebu S, Lee A, Salzberg A, Bauza V. (2024). Adaptive strategies to enhance water security and resilience in low- and middle-income countries: a critical review. Science of the Total Environment, 925:171520. https://doi.org/10.1016/j.scitotenv.2024.171520
7. Leštáková M, Logan KT, Rehm I-S, et al. (2024). Do resilience metrics of water distribution systems really assess resilience? A critical review. Water Research, 248:120820. https://doi.org/10.1016/j.watres.2023.120820
8. Lane K, Hrudey SE. (2023). A critical review of risk matrices used in water safety planning: improving risk matrix construction. Journal of Water and Health, 21:1795–1811. https://doi.org/10.2166/wh.2023.129
9. Graham J, Russell K, Gilpin B. (2023). When the implementation of water safety plans fail: rethinking the approach to water safety planning following a serious waterborne outbreak and implications for subsequent water sector reforms. Journal of Water and Health, 21:1562–1571. https://doi.org/10.2166/wh.2023.188
10. Vorstius C, Rowan JS, Brown I, et al. (2024). Supporting the development of climate-resilient water safety plans through a risk-screening framework for drinking water source protection. Journal of Water and Climate Change, 15:5207–5218. https://doi.org/10.2166/wcc.2024.293
11. Amadio J, Kanakoudis V, Dimkić D, et al. (2024). Natural hazard risk analysis in the framework of water safety plans. Journal of Water and Health, 22:451–466. https://doi.org/10.2166/wh.2024.124
12. Pankow N, Krause S, Schaum C. (2024). Resilience adaptation through risk analysis for wastewater treatment plant operators in the context of the European Union Resilience Directive. Water, 16:3098. https://doi.org/10.3390/w16213098
13. Poirier K, Lotfi M, Garg K, et al. (2023). A comprehensive review of pre- and post-treatment approaches to achieve sustainable desalination for different water streams.

Desalination, 566:116944. https://doi.org/10.1016/j.desal.2023.116944
14. Politano A, Al-Juboori RA, Alnajdi S, et al. (2024). 2024 roadmap on membrane desalination technology at the water-energy nexus. Journal of Physics: Energy, 6:021502. https://doi.org/10.1088/2515-7655/ad2cf2
15. Schär S, Bischi A, Baccioli A, et al. (2023). Optimization of sustainable seawater desalination: modeling renewable energy integration and energy storage concepts. Energy Conversion and Management, 293:117447. https://doi.org/10.1016/j.enconman.2023.117447
16. Sajna MS, Elmakki T, Schipper K, et al. (2024). Integrated seawater hub: a nexus of sustainable water, energy, and resource generation. Desalination, 571:117065. https://doi.org/10.1016/j.desal.2023.117065
17. Cai Y, Wu J, Shi SQ, et al. (2023). Advances in desalination technology and its environmental and economic assessment. Journal of Cleaner Production, 397:136498. https://doi.org/10.1016/j.jclepro.2023.136498
18. Alawad SM, Mansour RB, Al-Sulaiman FA, Rehman S. (2023). Renewable energy systems for water desalination applications: a comprehensive review. Energy Conversion and Management, 286:117035. https://doi.org/10.1016/j.enconman.2023.117035
19. Tashtoush B, Alyahya W, Al Ghadi M, et al. (2023). Renewable energy integration in water desalination: state-of-the-art review and comparative analysis. Applied Energy, 352:121950. https://doi.org/10.1016/j.apenergy.2023.121950
20. Mahmoudi A, Bostani M, Rashidi S, Valipour MS. (2023). Challenges and opportunities of desalination with renewable energy resources in Middle East countries. Renewable and Sustainable Energy Reviews, 184:113543. https://doi.org/10.1016/j.rser.2023.113543
21. Toledo M, Muñoz R. (2025). Odour prevention strategies in wastewater treatment and composting plants: a review. Journal of Environmental Management, 375:124402. https://doi.org/10.1016/j.jenvman.2025.124402
22. Shi L, Yang Q, Xu Z, Yang C. (2024). Technologies for in-situ H2S control in wastewater treatment plants: a review. Journal of Water Process Engineering, 65:105716. https://doi.org/10.1016/j.jwpe.2024.105716
23. Austigard ÅD, Smedbold HT, Svendsen KvH. (2023). Risk characteristics of hydrogen sulphide exposure in wastewater collection and treatment related occupations. Annals of Work Exposures and Health, 67:216–227. https://doi.org/10.1093/annweh/wxac065
24. Goh KC, Kurniawan TA, Goh HH, et al. (2024). Strengthening infrastructure resilience for climate change mitigation: case studies from the Southeast Asia region with a focus on wastewater treatment plants in addressing flooding challenges. ACS ES&T Water, 4:3725–3740. https://doi.org/10.1021/acsestwater.4c00591
25. Li Z, Ma W, Zhong D, et al. (2024). Applications of machine learning in drinking water quality management: a critical review on water distribution system. Journal of Cleaner Production, 481:144171. https://doi.org/10.1016/j.jclepro.2024.144171
26. Essamlali I, Nhaila H, El Khaili M. (2024). Advances in machine learning and IoT for water quality monitoring: a comprehensive review. Heliyon, 10:e27920. https://doi.org/10.1016/j.heliyon.2024.e27920
27. Gulzar Q, Mustafa K. (2024). An analytical survey of cyber-physical systems in water treatment and distribution: security challenges, intrusion detection, and future directions. Security and Privacy, 7:e440. https://doi.org/10.1002/spy2.440
28. Neshenko N, Bou-Harb E, Furht B, Behara R. (2024). Machine learning and user interface for cyber risk management of water infrastructure. Risk Analysis, 44:833–849. https://doi.org/10.1111/risa.14209
29. Liu K, Wang M, Ma R, et al. (2022). Detection and localization of cyber attacks on water treatment systems: an entropy-based approach. Frontiers of Information Technology & Electronic Engineering, 23:587–603. https://doi.org/10.1631/FITEE.2000546
30. Al-Obaidi M, Alsarayreh AA, Rashid FL, et al. (2024). Hybrid membrane and thermal seawater desalination processes powered by fossil fuels: a comprehensive review, future challenges and prospects. Desalination, 583:117694. https://doi.org/10.1016/j.desal.2024.117694.