Engineering Resilient MEP Infrastructure for Saudi Arabia's Mega Projects: Integrating Construction Quality Assurance, Testing and Commissioning, and Lifecycle Operational Reliability under Vision 2030.

Abstract

Saudi Arabia's mega projects depend upon mechanical, electrical and plumbing (MEP) services that sustain environmental control, electrical continuity, water provision, fire protection and safe building operation. Yet conventional completion procedures often separate construction inspection from functional acceptance and subsequent maintenance, leaving latent interface defects undiscovered until occupation. This critical review examines how construction quality assurance, structured testing and commissioning, and risk-informed lifecycle management can operate as a single resilience system. Thirty peer-reviewed studies published between 2020 and 2025 were examined through a transparent, theme-based synthesis of evidence on quality governance, model-enabled coordination, commissioning, operational diagnostics and infrastructure resilience. The analysis distinguishes demonstrated technological capabilities from proposed management arrangements and does not present simulated outcomes as field measurements. It develops an assurance-to-reliability framework linking asset requirements, verified installation, performance testing, operational baselines and feedback-led intervention. Particular attention is given to high ambient temperatures, airborne dust, intensive occupancy, complex subcontracting and interfaces among critical systems in Saudi developments. Findings indicate that information continuity and risk-based acceptance are more defensible indicators of future serviceability than completion percentages alone. The paper specifies practical hold points, evidence packages, failure-consequence measures and responsibility arrangements suitable for owner-led adoption. The framework provides a testable research agenda rather than an unverified claim of superior project performance.

Keywords: *mechanical, electrical and plumbing; construction quality assurance; testing and commissioning; asset integrity; operational reliability; Saudi Vision 2030*

UNDER PEER REVIEW IN IJAR

1. Introduction

Saudi Arabia's Vision 2030 development programme is creating complex transport hubs, destination resorts, mixed-use districts, healthcare buildings and public facilities whose performance depends on coordinated building services. Mechanical plant, electrical distribution, plumbing networks, controls

and fire-safety interfaces must operate as interdependent systems rather than discrete construction packages. The importance of governance in Saudi construction quality has been documented, including the influence of leadership, organisational routines and project-specific processes [1]. Research on commissioning likewise identifies shortcomings in documentation and continuity between delivery and operation [2]. The engineering concern is therefore not simply whether equipment has been installed, but whether its required functions have been verified under credible operating and abnormal conditions.

An MEP installation may achieve a satisfactory visual inspection while remaining operationally fragile. A cooling plant can pass individual component tests without demonstrating coordinated demand response; an emergency generator can start successfully yet fail to maintain acceptable downstream loads after transfer; and a pressure test can confirm pipe integrity without establishing long-term flow, control and maintainability. Continuous commissioning research shows the value of comparing operational measurements with expected building performance, rather than treating acceptance as a single event [3]. More recent reviews of digital twins also caution that connecting representations to operational reality involves unresolved modelling and data-quality challenges [4]. These distinctions matter when construction-stage claims of digital completion are used as proxies for resilience.

The design and handover environment compounds those risks. Multiple specialist contractors often install highly interdependent services under compressed programmes, with late architectural revisions, access conflicts, fragmented inspection records and changing operational requirements. A resilient outcome depends upon preserving the engineering intent behind each inspection or test, making deviations visible, and transferring usable evidence to the facility operator. This article interprets resilience as the ability to provide essential services, absorb foreseeable disturbances, recover within agreed limits and adapt maintenance decisions as conditions change. It does not equate resilience with the mere availability of standby equipment or sophisticated digital platforms.

The Saudi operating context also demands a careful separation of environmental exposure from design assumptions. Hot summers influence condenser performance, electrical loading and the time available to restore cooling; dust can affect filters, heat-transfer surfaces and exposed equipment; and expansive sites can complicate response times and replacement logistics. These considerations are not universal failure predictions for every project. They are scenario variables that ought to appear in design-basis records, commissioning conditions and asset-criticality assessments. Similarly, facility types differ: a passenger interchange, hotel, entertainment venue and healthcare building impose

UNDER PEER REVIEW IN IJAR

distinct tolerances for service interruption. The assurance framework therefore begins with an owner-agreed definition of essential functions and acceptable downtime rather than presuming that all mechanical or electrical equipment merits identical testing.

2. Research aim and questions

The aim is to construct an evidence-informed lifecycle assurance framework that explains how construction-phase MEP quality controls can become defensible operational reliability. Three objectives guide the review: first, to identify which installation and information-quality controls best protect downstream functional performance; second, to examine how commissioning should verify system interactions, degraded modes and recoverability; and third, to translate acceptance evidence into risk-based maintenance priorities after handover. The corresponding questions concern traceability between requirements and tests, the operational significance of unresolved defects, and the governance arrangements necessary to maintain verified performance. These objectives differ from assessments focused only on design productivity, digital adoption or construction schedule gains. The intended contribution is an auditable decision logic rather than a purportedly validated Saudi project case study.

3. Review design and interpretive method

A structured critical narrative review was selected because the topic combines construction management, building-services engineering, commissioning, information management and reliability analysis; their outcomes cannot be meaningfully pooled into one effect estimate. Evidence discovery focused on scholarly publisher records, DOI-indexed article descriptions and the technical themes of MEP coordination, construction inspection, functional commissioning, building operations, failure diagnostics and thermal or electrical resilience. Eligibility was restricted to research published during 2020–2025 with identifiable article metadata and a DOI. Thirty relevant publications were retained to represent distinct parts of the proposed lifecycle chain. Standards and statutory obligations are discussed as engineering requirements but are not misrepresented as peer-reviewed experimental evidence.

Each article was interpreted according to its research setting, method, operational domain, type of evidence and limits of transferability. Site-tested inspection methods were distinguished from conceptual digital-twin architectures; operational fault-detection demonstrations were separated from studies of organisational barriers; and resilience optimisation studies were not assumed to establish construction quality outcomes. The synthesis proceeded through five analytical categories:

UNDER PEER REVIEW IN IJAR

installation conformity, interface coordination, functional acceptance, asset-information handover and risk-informed operation. Contradictions were retained where a technological intervention improved information visibility without proving improvements in reliability or safety. The selection is purposive rather than exhaustive; therefore no fabricated database-search totals, exclusion flow or statistical meta-analysis is reported.

The attached Saudi construction paper informed the broad article architecture, notably the separation of methods, contextual discussion, evaluation and limitations. Its case-specific design-phase improvement figures were deliberately excluded from the evidence base because they do not demonstrate MEP commissioning effectiveness or lifecycle fault reduction. The present review makes no claim to have interviewed asset owners, accessed confidential commissioning logs or observed reliability over time. Instead, its propositions are evaluated for engineering coherence against the included literature and converted into performance measures that could subsequently be tested on comparable Saudi facilities. This distinction between conceptual synthesis and empirical validation is essential to avoid overstating the strength of the conclusions.

Critical appraisal considered whether a publication's findings could be translated into an observable engineering control. For example, geometric inspection research informs how installed location might be verified, whereas fault-detection research informs how operational deviations might be recognised; neither automatically proves that an entire facility is more reliable. A claim was treated as transferable when the physical mechanism, required data and operating boundary could be clearly specified. Evidence of feasibility was not escalated into evidence of economic benefit. The resulting framework is an interpretive synthesis that can be falsified: if projects adopting its linked assurance gates show no improvement in independently observed service continuity or defect recovery, the value of the proposed integration must be reconsidered. This explicit testability reduces the risk of presenting the literature review as a promotional account of technology.

4. From design coordination to construction conformity

The installed quality of MEP systems begins with decisions that determine access, clearances, loads, maintainability and the possibility of testing. Generative optimisation of MEP plant spaces demonstrates the value of examining multiple layouts before construction, although geometric feasibility alone cannot establish serviceability [5]. Research using graph-based analysis of MEP clashes similarly shows that conflict resolution depends on understanding component relationships, not merely counting detected collisions [6]. Design reviews should therefore test removal routes,

UNDER PEER REVIEW IN IJAR

isolation boundaries, lifting provisions, sensor access and safe separation between services. An installation can be geometrically coordinated but deficient if equipment cannot later be calibrated, removed or isolated without disabling an adjacent critical function.

Site evidence must establish that built conditions remain consistent with that intent. Mixed-reality comparison has been investigated for MEP progress and geometric verification, but detection performance depends upon the representation and comparison method [7]. Inspection knowledge linked to building models can help field teams access the correct acceptance criteria at the point of work [8]. Geometric reconstruction from point clouds can also support quality inspection and later maintenance records [9]. These tools strengthen observation, yet none replaces instrument calibration, witnessed functional checks or engineering judgement about pressure, electrical protection, flow balance and fire-system interactions. The proposed control principle is to pair every digital observation with a named requirement, acceptance threshold, responsible inspector and retained evidence artefact.

Information structure is a second determinant of conformity. Automated duct modelling can support the creation of coherent service representations from fragmented drawings [10], while object-classification research exposes persistent problems with incomplete semantics and building-performance exchange [11]. Large MEP models may contain redundant objects that hamper useful facility management [12]. Consequently, the owner should specify asset identifiers, system boundaries, tagging conventions, control-point names and maintainable equipment classes before installation begins. Information that cannot identify the installed asset, its governing test or its physical location has limited reliability value. The practical measure of model quality is not file size or visual sophistication, but the proportion of safety- and operation-critical assets with accurate, retrievable and verified records.

Design responsibility must also include defining a measurable performance envelope before procurement. Chillers require more than nominal cooling-capacity schedules: review should address expected water temperatures, part-load behaviour, redundancy philosophy and allowable recovery time. Emergency electrical equipment requires a clear list of essential loads, protection selectivity and acceptable interruption duration. Water systems require an operating range for pressure and quality, while drainage resilience depends on capacity, backflow safeguards and accessible maintenance routes. The same information should appear in procurement submittals, inspection criteria and commissioning scripts, so a substitution is evaluated against system performance rather than only dimensional fit. Any approved deviation needs to update the governing requirement and its

UNDER PEER REVIEW IN IJAR

downstream tests. Without this link, apparently compliant purchases can unintentionally change the service outcome the owner expects.

5. Construction quality assurance at system boundaries

Quality assurance should follow physical and functional boundaries rather than subcontract packages alone. An inspection-and-test plan ought to identify hold points before concealment, energisation, pressurisation and integrated operation. For chilled-water networks this includes material certification, support and expansion provisions, cleanliness, pressure integrity, flushing, balancing and recorded control-valve authority. For power systems the evidence chain includes cable identity, termination torque, insulation measurements, protection settings, earth continuity and the controlled sequence of energisation. For life-safety interfaces, witnessed cause-and-effect tests should verify that an alarm produces the intended response across ventilation, access control, lifts and emergency supply. Each hold point needs a release decision, not merely a document upload.

The hardest defects frequently sit between trades. A valve installed correctly by the mechanical contractor may be commanded by the wrong controller point; a maintained emergency distribution board may supply equipment whose control circuit lacks ride-through capability. These interface failures are not resolved by more intensive inspections within individual disciplines. An interface register should therefore map initiating event, participating assets, expected state changes, test prerequisites and evidence owners. A cross-trade witness is required whenever one party cannot establish the outcome alone. Defect classification must consider service consequence and detectability, so an undocumented critical interlock receives higher attention than a minor non-functional finish issue even when both count as one outstanding item.

Documentation quality is itself an engineering barrier. Asset information requirement studies show the importance of defining needed records, their intended uses and the stakeholder responsible for supplying them [13]. Work on construction, warranty and maintenance interoperability likewise illustrates why information should persist after contractor demobilisation [14]. For a Saudi mega project, the accepted dossier should link equipment tags, certificates, redlined drawings, inspection releases, calibrated test instruments, commissioning scripts, deviations, training and warranty boundaries. Records should be searchable by asset and system, not only by contractor or document number. Table 1 translates these principles into illustrative acceptance controls. Its entries are proposed governance criteria, not measurements obtained from a project.

UNDER PEER REVIEW IN IJAR

In construction quality control, selective inspection is sometimes unavoidable because the volume of repeated installations is large. Sampling, however, should not be uniform across consequence classes. Concealed joints, critical electrical terminations, fire-stopping penetrations and difficult-to-access valves warrant traceability that permits a failed sample to trigger examination of related work installed by the same method or crew. Photographs should include location, date, asset identification and relevant measurement rather than being accepted as uncontextualised visual proof. Non-conformance closure should record cause, correction, independent reinspection and whether a comparable defect could exist elsewhere. A repeated defect is a process signal that may require revising work instructions or supervision, not simply a sequence of isolated items to be cleared from a dashboard.

Table 1. Proposed discipline-specific assurance gates and lifecycle evidence requirements.

MEP domain	Assurance hold point	Witnessed acceptance evidence	Lifecycle implication
Cooling and ventilation	Pressure, flushing, balancing and access before concealment	Calibrated readings; part-load and changeover sequences	Transfer an operating baseline and maintainable isolation
Electrical and emergency power	Cable identification, protection settings, earthing, energisation release	Insulation and protection records; representative transfer-load test	Demonstrate service continuity and restoration authority
Water and drainage	Hydraulic integrity, cleanliness, isolation and backflow prevention	Pressure, flow, quality and abnormal-condition records	Traceable valves, alarms and response procedures
Fire and life safety	Cause-and-effect matrix approved across affected trades	Witnessed integrated alarm, ventilation and power interactions	Retain evidence for later changes to occupancy or controls
Controls and communications	Point-to-point mapping, time synchronisation and version control	Alarm, trend, fail-safe and manual-override verification	Maintain reliable diagnosis and accountable change history

UNDER PEER REVIEW IN IJAP

Note: Controls are an interpretive framework, not reported observations.

6. Testing and commissioning as performance assurance

Commissioning must verify the owner-required function under realistic loads, sequences and interaction states. Component pre-commissioning establishes that installation is safe for functional testing; system commissioning demonstrates controlled operation; integrated systems testing evaluates whether several systems respond correctly to shared events. This hierarchy avoids the common error of treating energisation or a satisfactory manufacturer start-up sheet as final acceptance. Model-supported commissioning research demonstrates how control-system configuration and simulated expectations may be carried into verification [15]. Nevertheless, automated scripts require independent checking of point mapping, logic version, instrument accuracy and failure-state behaviour. A passed software test is not evidence of a safe physical response unless the final device and actual operating pathway were exercised.

A commissioning script needs an explicit initial condition, stimulus, predicted response, permitted tolerance, measurement location and restoration step. For example, a smoke-control scenario should establish occupancy assumptions and damper states, initiate the relevant detection event, observe fan responses and pressure relationships, check alarms at the operator workstation, then demonstrate an orderly return to normal. Emergency-power testing should impose representative load steps and verify selective protection, transfer timing and the continuity of essential services. Cooling systems should be assessed at minimum and part loads, during equipment changeover and after sensor or pump faults. Tests that exclude credible abnormal states risk verifying convenience rather than resilience.

Seasonal and staged acceptance are especially important for Saudi facilities. A cooling plant handed over in mild weather may not have demonstrated its high-temperature capacity, while a district-scale development may receive new loads months after an initial area opens. The programme should therefore distinguish provisional readiness from final performance acceptance and document deferred tests with owners, dates, operational restrictions and contingency measures. Continuous commissioning research supports comparison of measured operation with a defined performance baseline [3]. A baseline should preserve operating modes, ambient conditions, set points, alarm thresholds and energy or flow measurements so that later deviations can be meaningfully interpreted, rather than judged against a single unrepresentative test-day value.

UNDER PEER REVIEW IN IJAR

Commissioning closure should also be risk weighted. A critical defect may demand a no-occupancy decision, temporary operating constraint or demonstrable independent mitigation; a non-critical issue can be accepted subject to a time-bound rectification plan. The essential record is a signed decision that describes residual risk, affected service, compensatory action and accountable party. The framework prohibits converting open issues into completed work solely because a programme milestone is imminent. Figure 1 depicts the required flow from design intent to evidence-backed operation. Repeated verification following software changes, equipment replacement or altered occupancy is part of assurance continuity, not an admission that initial testing failed.

Quantitative test acceptance should avoid both false precision and subjective pass marks. The design team should determine allowable deviations from approved performance criteria, while the commissioning authority identifies instrument accuracy, measurement uncertainty and the number of repetitions needed to support a decision. Results recorded as merely satisfactory obscure whether equipment was near a threshold. Trend records can reveal unstable control loops or delayed responses that a single spot reading would miss. Where a test cannot safely impose a full fault, the record must disclose the substituted method, assumptions and remaining evidence gap. Combining witnessed physical tests, control-event logs and calibrated measurements provides stronger assurance than relying on any one source. It also gives future maintainers a defensible reference when comparing degraded operating conditions with the original accepted state.

UNDER PEER REVIEW IN IJAP

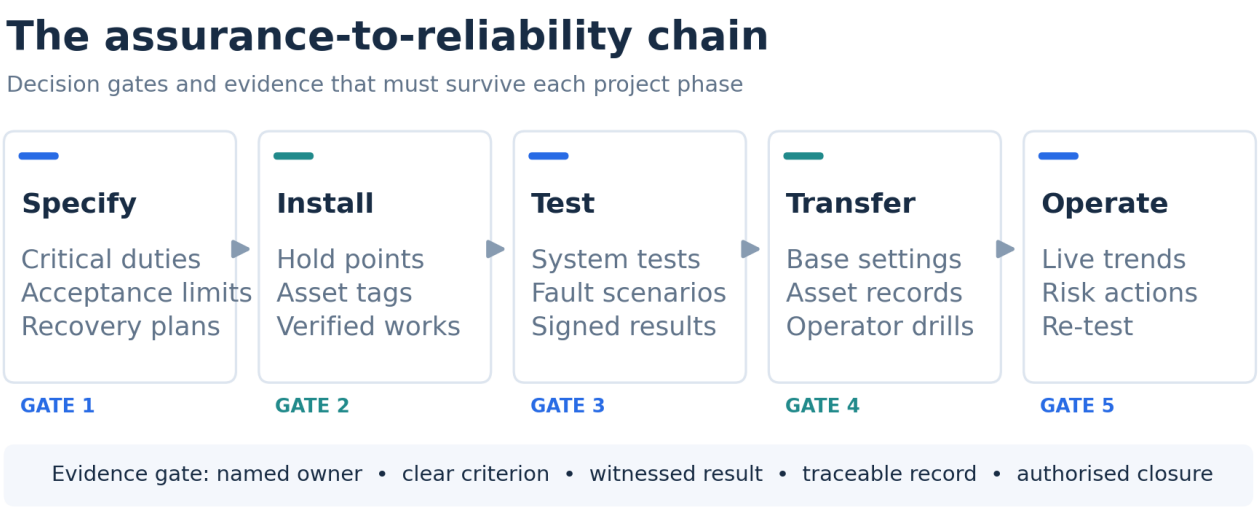


Figure 1. Proposed assurance-to-reliability chain across project delivery and operation.

7. Operational reliability, diagnostics and maintainability

An operational asset register becomes useful only when acceptance records are linked to deterioration mechanisms and decisions. Virtual modelling between building information models and operating data demonstrates a route towards more dependable HVAC fault diagnosis [16]. Work using real building-automation information provides examples of fault recognition without relying exclusively on ideal laboratory labels [17]. However, detection algorithms often identify abnormal patterns rather than root causes or safety consequences. Operators must evaluate the plausibility of each alert against sensor integrity, maintenance history, weather, occupancy and known control changes before intervening. False alerts consume labour; missed critical faults can weaken resilience. Diagnostic performance should be measured alongside the quality of corrective action.

Risk prioritisation must join probability, consequence and detection capability. Research on air-handling-unit fault prioritisation links observed fault occurrence with performance impacts [18], while building-level diagnostic approaches show the importance of tracing equipment symptoms across connected systems [19]. Industry evidence also identifies organisational and commercial barriers to the adoption of automated diagnostics [20]. The proposed decision rule assigns each failure mode an affected essential service, expected functional loss, redundancy status, time to detect, feasible recovery pathway and verification test. Criticality is then reviewed whenever operating conditions change. A reliable maintenance programme cannot simply maximise the number of predictive models; it must demonstrate that high-consequence degradation leads to timely, accountable intervention.

A successful predictive alert should create a closed work order, identify the asset and probable failure mechanism, specify safe access, and require a post-repair functional test. Unsupervised analysis of fan-coil units illustrates how methods using normal building automation data can identify unusual operating patterns, although dimensionality reduction may hide important behaviours [21]. Maintenance teams should therefore retain raw context for contested events and monitor diagnosis accuracy, repeat faults, mean time to restoration and the fraction of critical assets with a current tested operating baseline. Table 2 proposes a concise reliability dashboard. A low count of alarms, by itself, is not evidence of health when sensors, alarm routing or operator response are unreliable.

Resilience also requires evaluating dependencies beyond the building boundary. Studies of critical-building microgrids illustrate how power continuity may be supported through alternative dispatch and supply arrangements [22]. Building-climate reviews emphasise the importance of defining service resilience under rising thermal stress [23], and lifecycle optimisation research shows why

UNDER PEER REVIEW IN IJAR

uncertain future climate conditions should be considered in design decisions [24]. For Saudi mega projects, a practical resilience scenario might combine extreme ambient heat, partial electrical loss, restricted water supply and delayed spare-part delivery. Appropriate responses include load shedding priorities, emergency cooling or ventilation strategies, accessible isolation, spare components and recovery procedures that are tested, funded and understood by operators.

For decision support, a simple risk model can be more useful than a complex opaque score. Each failure scenario should be described through service consequence, observed or plausible occurrence, detectability, recovery resources and dependency on other assets. Owners may rank these factors ordinally if reliable frequencies are absent, provided they document the rationale and avoid presenting rankings as probabilistic forecasts. A standby pump does not necessarily reduce consequence if both pumps share a vulnerable controller, power feeder or suction arrangement. Conversely, an accessible bypass and rehearsed isolation procedure may shorten downtime even where equipment redundancy is limited. This distinction separates resilience engineering from equipment-count accounting and encourages investment in independent failure paths, response readiness and procedures that can actually be executed under pressure.

Table 2. Illustrative risk-based operational dashboard for critical MEP systems.

Failure or degradation	**Service consequence to assess**	**Operational evidence**	**Owner response**
Chiller or pump capacity loss	Thermal comfort and essential cooling exposure	Load-normalised temperatures, flows, trend and recovery time	Prioritise redundancy, repair and re-test
Protection or transfer failure	Loss of emergency electrical continuity	Switchgear events, settings, witnessed transfer records	Restrict operation pending functional clearance
Incorrect controls mapping	Undetected command or interlock failure	Point mapping, alarms and as-tested logic revision	Reconcile version, then repeat the scenario test

UNDER PEER REVIEW IN IJAR

Failure or degradation	Service consequence to assess	Operational evidence	Owner response
Drainage, supply or isolation defect	Flooding, pressure loss or water hygiene risk	Trend, leakage, sample and isolation records	Contain, isolate and verify reinstatement
Sensor, network or analytics outage	Degraded visibility and delayed fault identification	Sensor-health, data completeness and local-control evidence	Use independent local controls and restore telemetry

Note: Each indicator requires project-specific thresholds, data quality checks and assigned authority.

8. Information continuity and accountable governance

A digital representation is valuable only when it preserves the relationship between what was specified, what was installed, what was tested and what is currently happening. Building-operation case studies have identified the promise of digital twins alongside difficulties of data interpretation and governance [25]. A recent review of operation-and-maintenance applications also identifies gaps in information integration and regional evidence [26]. The owner should mandate a controlled system-of-record connecting asset identifiers, final configuration, commissioned settings, software revisions, manuals, spares, maintenance tasks and live telemetry. Data stewardship must include approval rights and change histories. Otherwise, a visually impressive model may become obsolete after the first control alteration or equipment replacement.



PEER REVIEW IN JAR

Figure 2. Closed-loop operational decision model linking observation, risk and verified restoration.

Interactive operational twins demonstrate the feasibility of linking building automation, three-dimensional context and user-facing information [27]; broader reviews of construction data ecosystems likewise emphasise integration challenges [28]. Such capabilities do not eliminate cybersecurity and safety obligations. Remote monitoring interfaces need role-based access, segmentation between business and operational networks, secure change control and procedures for operating during digital outages. A digital-twin service can fail without the underlying mechanical system failing, so resilience plans must retain independent local control, clear manual overrides and offline emergency instructions. Figure 2 represents the proposed operational feedback loop, including a decision gate that separates a data alert from authorised maintenance action.

The governance model assigns requirements to the client and designer, installation proof to contractors, functional witnesses to commissioning authorities, acceptance decisions to the owner, and continuing baseline management to the operator. Responsibility overlaps must be resolved through named decision rights rather than a general instruction to collaborate. A conceptual operational framework can clarify these interfaces [29], while automated multi-drawing MEP modelling illustrates the potential to improve the consistency of transferred asset representations [30]. Contract deliverables should therefore include both technical evidence and information usability tests. The acceptance question becomes whether a facility team can locate a critical asset, interpret its test history and execute its recovery procedure without relying on the original subcontractor.

Operational information acceptance should be tested through tasks performed by the eventual maintainers. A representative exercise would ask an engineer to find an installed asset, identify its isolation points, retrieve the latest commissioning result, recognise the relevant alarm, obtain a replacement specification and close a corrective work order with a post-repair test. Missing identifiers or incompatible control-point naming should be classified as handover defects when they impede those tasks. This exercise is more demanding than confirming that a common data environment contains the required folders. It also exposes training and language needs, access-permission failures and gaps between contractor asset schedules and the maintenance-management platform. The operator's participation before practical completion is consequently a form of technical verification rather than a ceremonial transfer of documents.

UNDER PEER REVIEW IN IJAR

9. Discussion: application to Saudi mega projects

The integrated synthesis reveals a recurring mismatch between delivery efficiency and service resilience. Digital coordination can reduce visibility gaps, but it does not automatically establish correct protection settings, pressure relationships, operating sequences or recovery arrangements. Installation conformity becomes reliability only when functional testing creates credible baselines and when maintenance teams can retrieve and challenge those baselines. This argument is particularly consequential in Saudi programmes combining fast-track construction with phased occupation, complex logistics and high cooling demand. A single universal checklist would be inadequate; acceptance depth should reflect system consequence, operational dependencies and the maturity of the eventual operator.

Implementation should proceed through a pilot involving a bounded set of critical services, such as a central cooling plant, emergency electrical distribution and fire-alarm interfaces. Before the pilot, the owner would define service-loss tolerances, critical equipment identifiers, minimum evidence completeness, deferred-test rules and intervention authority. During commissioning, the team would record results against credible normal and fault conditions. During the first operating period, maintainers would monitor detection latency, failed-test recurrence, restoration duration and the percentage of risk-critical deviations resolved within an agreed window. Comparison with conventional handover packages would then be possible without inventing savings or attributing unrelated operational improvements to the framework.

The literature does not support a general claim that integrated digital assurance necessarily lowers faults by a particular percentage. Many cited studies concern single technologies, individual buildings or simulated scenarios; direct longitudinal evidence from Saudi mega-project MEP systems remains limited. Organisational incentives, climatic exposure, contractor capability and operating regimes could materially alter outcomes. The review's purposive source selection, reliance on accessible article metadata and absence of original field data also constrain causal inference. Future studies should use multi-project prospective designs, independent test witnesses, comparable system criticality and predeclared performance indicators. Particular attention should be paid to cases in which higher inspection effort yields more detected defects initially but fewer service interruptions later.

A phased adoption strategy could first establish minimum asset information and high-consequence interface tests, then introduce automatic performance analytics only after instrumentation and control-point reliability have been demonstrated. This sequencing reduces the likelihood of

UNDER PEER REVIEW IN IJAR

automating poor-quality evidence. Evaluation should include a counterfactual or comparable baseline wherever feasible, with comparable building use, climate exposure and service criticality. Researchers should report the cost of additional inspection and data stewardship as well as any operational benefit, because apparent savings may depend on how those costs are allocated. An independent assurance team should review failed and deferred tests rather than accepting only completed records. Finally, measured improvements should be linked to actual service outcomes, such as avoided downtime or shorter verified recovery, rather than the number of digital models or work orders produced.

10. Conclusions

Resilient MEP infrastructure requires a chain of accountable engineering decisions, not merely successful installation or digital handover. This review has integrated evidence on Saudi quality management, MEP coordination, inspection technologies, commissioning procedures, operational diagnostics and climate-sensitive resilience into a lifecycle assurance framework. The distinctive contribution is a set of linked decision gates: verified requirements, construction conformity, functional and abnormal-mode acceptance, usable asset information, and risk-based corrective feedback. Proposed figures and tables express testable engineering controls rather than invented project outcomes. For Saudi Arabia's major developments, the practical priorities are rigorous interface testing, auditable deferred-test management, owner-defined information requirements and explicit recovery responsibilities. The research agenda is to evaluate these arrangements across actual facilities and establish whether they improve essential-service continuity, defect detection and restoration performance over time.

The proposed framework should be judged against service-based outcomes rather than its documentation volume. A longitudinal investigation would compare facilities with respect to high-consequence defects discovered after occupation, mean time to restore essential functions, repeat commissioning failures and the accuracy of transferred asset information. Data collection should record exposure conditions and occupancy patterns to distinguish infrastructure deterioration from changing demand. Owners could also investigate whether independent witnessing improves the early identification of common-cause failures and whether the financial burden of assurance falls disproportionately on contractors. Such evaluation would clarify the appropriate level of oversight for different system criticalities. Until comparative evidence is available, the framework remains a reasoned engineering proposition. Its value lies in making assumptions, acceptance evidence,

UNDER PEER REVIEW IN IJAP

unresolved risks and recovery responsibilities visible to the people who must operate the infrastructure safely.

References

[1] Trigunarsyah B, Moohialdin A, Islam MS (2023) Quality management system framework for construction companies in the Kingdom of Saudi Arabia. International Journal of Productivity and Quality Management 39(3):302–329. https://doi.org/10.1504/IJPQM.2023.132262

[2] Singh J, Anumba CJ (2024) Building commissioning process and documentation: a literature review and directions for future research. International Journal of Construction Management 24(1):75–85. https://doi.org/10.1080/15623599.2023.2211409

[3] Jradi M, Liu N, Arendt K, Mattera CG (2020) An automated framework for buildings continuous commissioning and performance testing – a university building case study. Journal of Building Engineering 31:101464. https://doi.org/10.1016/j.jobe.2020.101464

[4] AlBalkhy W, Karmaoui D, Ducoulombier L, Lafhaj Z, Linner T (2024) Digital twins in the built environment: definition, applications, and challenges. Automation in Construction 162:105368. https://doi.org/10.1016/j.autcon.2024.105368

[5] Pestana E, Paice A, West S (2024) Optimizing MEP design in early AEC projects through generative design. Automation in Construction 165:105566. https://doi.org/10.1016/j.autcon.2024.105566

[6] Hu Y, Xia C, Chen J, Gao X (2023) Clash context representation and change component prediction based on graph convolutional network in MEP disciplines. Advanced Engineering Informatics 55:101896. https://doi.org/10.1016/j.aei.2023.101896

[7] Tao B, Bosché F, Li J (2024) Mixed Reality-based MEP construction progress monitoring: evaluation of methods for mesh-to-mesh comparison. Automation in Construction 168:105852. https://doi.org/10.1016/j.autcon.2024.105852

[8] Liu H, Liu D, Chen J (2025) Natural language-extracted and BIM-referenced knowledge base for construction quality inspection via augmented reality. Automation in Construction 179:106500. https://doi.org/10.1016/j.autcon.2025.106500

[9] Kim M, Lee D (2023) Automated two-dimensional geometric model reconstruction from point cloud data for construction quality inspection and maintenance. Automation in Construction 154:105024. https://doi.org/10.1016/j.autcon.2023.105024

UNDER PEER REVIEW IN IJAR

[10] Park S, Shin M, Jang JY, Koo B, Kim TW (2024) Automated process for generating an air conditioning duct model using the CAD-to-BIM approach. Journal of Building Engineering 109529. https://doi.org/10.1016/j.jobe.2024.109529

[11] Utkucu D, Ying H, Wang Z, Sacks R (2024) Classification of architectural and MEP BIM objects for building performance evaluation. Advanced Engineering Informatics 61:102503. https://doi.org/10.1016/j.aei.2024.102503

[12] Guo Z, Li Q, Wang Q, Qiao S, Mei T, Zuo W (2024) Reducing redundancy in large MEP building information models for facility management using oriented bounding boxes and hash sets. Automation in Construction 165:105514. https://doi.org/10.1016/j.autcon.2024.105514

[13] Benn M, Stoy C (2023) Implementing asset information requirement templates for corporate real estate management: a study in the chemical industry. Developments in the Built Environment 16:100224. https://doi.org/10.1016/j.dibe.2023.100224

[14] Palha RP, Hüttl RMC, Silva AJC (2024) BIM interoperability for small residential construction integrating warranty and maintenance management. Automation in Construction 166:105639. https://doi.org/10.1016/j.autcon.2024.105639

[15] Zech P, Hammes S, Goldin E, Geisler-Moroder D, Breu R, Pfluger R (2025) From BIM to Digital Twin: a transformation process through advanced control modeling and automated commissioning using daylight and artificial lighting as examples. Energy and Buildings 115184. https://doi.org/10.1016/j.enbuild.2024.115184

[16] Yoon S, Lee J, Li J, Wang P (2024) Virtual in-situ modeling between digital twin and BIM for advanced building operations and maintenance. Automation in Construction 168:105823. https://doi.org/10.1016/j.autcon.2024.105823

[17] Soultanzadeh MB, Ouf MM, Nik-Bakht M, Paquette P, Lupien S (2024) Fault detection and diagnosis in light commercial buildings' HVAC systems: a comprehensive framework, application, and performance evaluation. Energy and Buildings 316:114341. https://doi.org/10.1016/j.enbuild.2024.114341

[18] Gopalan S, Rijs A, Chitkara S, Thamban A, Kramer R (2024) Fault prioritisation for Air Handling units using fault modelling and actual fault occurrence data. Energy and Buildings 319:114476. https://doi.org/10.1016/j.enbuild.2024.114476

[19] Gao T, Marié S, Béguery P, Thebault S, Lecoeuche S (2024) Integrated building fault detection and diagnosis using data modeling and Bayesian networks. Energy and Buildings 113889. https://doi.org/10.1016/j.enbuild.2024.113889

UNDER PEER REVIEW IN IJAR

[20] Andersen KH, Melgaard SP, Johra H, Marszal-Pomianowska A, Jensen RL, Heiselberg PK (2024) Barriers and drivers for implementation of automatic fault detection and diagnosis in buildings and HVAC systems: an outlook from industry experts. Energy and Buildings 303:113801. https://doi.org/10.1016/j.enbuild.2023.113801

[21] Rajabi F, McArthur JJ (2024) Applying OPTICS with and without PCA for fault detection of fan coil units using building automation system data. Energy and Buildings 317:114368. https://doi.org/10.1016/j.enbuild.2024.114368

[22] Rosales-Asensio E, Icaza D, González-Cobos N, Borge-Diez D (2023) Peak load reduction and resilience benefits through optimized dispatch, heating and cooling strategies in buildings with critical microgrids. Journal of Building Engineering 68:106096. https://doi.org/10.1016/j.jobe.2023.106096

[23] Peri G, Cirrincione L, Mazzeo D, Matera N, Scaccianoce G (2024) Building resilience to a warming world: a contribution toward a definition of "Integrated Climate Resilience" specific for buildings – literature review and proposals. Energy and Buildings 315:114319. https://doi.org/10.1016/j.enbuild.2024.114319

[24] Chen R, Samuelson H, Zou Y, Zheng X, Cao Y (2024) Improving building resilience in the face of future climate uncertainty: a comprehensive framework for enhancing building life cycle performance. Energy and Buildings 302:113761. https://doi.org/10.1016/j.enbuild.2023.113761

[25] Zhao J, Feng H, Chen Q, Garcia de Soto B (2022) Developing a conceptual framework for the application of digital twin technologies to revamp building operation and maintenance processes. Journal of Building Engineering 49:104028. https://doi.org/10.1016/j.jobe.2022.104028

[26] Elshabshiri A, Ghanim A, Hussien A, Maksoud A, Mushtaha E (2025) Integration of Building Information Modeling and Digital Twins in the Operation and Maintenance of a building lifecycle: a bibliometric analysis review. Journal of Building Engineering 99:111541. https://doi.org/10.1016/j.jobe.2024.111541

[27] Ma T, Xiao F, Zhang C, Zhang J, Zhang H, Xu K, Luo X (2025) Digital twin for 3D interactive building operations: integrating BIM, IoT-enabled building automation systems, AI, and mixed reality. Automation in Construction 176:106277. https://doi.org/10.1016/j.autcon.2025.106277

[28] Li F, Laili Y, Chen X, Lou Y, Wang C, Yang H, Gao X, Han H (2023) Towards big data driven construction industry. Journal of Industrial Information Integration 35:100483. https://doi.org/10.1016/j.jii.2023.100483

UNDER PEER REVIEW IN IJAR

494 [29] Borkowski AS (2024) Digital twin conceptual framework for the operation and maintenance
495 phase in the building's lifecycle. Archives of Civil Engineering 70(3):139–152.
496 https://doi.org/10.24425/ace.2024.150975
497 [30] Zhao Q, Shi H, Zhou L, Lv G (2025) Automated BIM generation for MEP systems from CAD
498 data using multi-drawing graph integration. Automation in Construction 180:106542.
499 https://doi.org/10.1016/j.autcon.2025.106542

UNDER PEER REVIEW IN IJAR