Global Academic Journal of Economics and Business

Available online at https://www.gajrc.com
DOI: https://doi.org/10.36348/gajeb.2026.v08i06.031

OPEN ACCESS



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

Original Research Article

AI Driven Predictive Maintenance and Reliability Optimisation of HVAC and MEP Systems in Saudi Smart Buildings under Vision 2030

Muhammad Junaid Raza[1]*
[1]Independent Researcher

***Corresponding Author**
Muhammad Junaid Raza
Independent Researcher

Article History
Received: 22.09.2026
Accepted: 14.11.2026
Published: 18.11.2026

Abstract: Saudi Arabia's expanding stock of digitally connected buildings creates a need to maintain cooling and mechanical, electrical and plumbing (MEP) services without compromising safety, occupant comfort or energy efficiency. This critical review examines how artificial intelligence can support predictive maintenance of chillers, air handling units, pumps, fans and related building infrastructure. Drawing on 30 peer reviewed publications issued between 2020 and 2025, it distinguishes fault detection from diagnosis, prognosis and maintenance decision making; compares statistical learning, deep neural networks, hybrid physical data models and digital twins; and evaluates the constraints of real building management system data. The synthesis finds that published methods offer useful early warning and diagnostic capabilities, but that laboratory accuracy does not establish reduced downtime, improved maintainability or positive lifecycle economics. Missing labels, shifting operating conditions, sensor faults, cybersecurity exposure and weak links to work order systems complicate practical adoption. A staged implementation framework is proposed for Saudi facilities that combines asset criticality analysis, trustworthy telemetry, interpretable fault assessment, human approval and measured post intervention verification. The review argues that predictive maintenance should be judged through reliability, cost, energy and service continuity outcomes rather than algorithmic performance alone. It identifies priorities for longitudinal, multi building evaluation under hot climate operation and for integrating predictions with accountable engineering decisions in support of Vision 2030.

Keywords: Predictive Maintenance, HVAC, MEP, Smart Buildings, Fault Detection and Diagnosis, Building Management Systems, Reliability, Digital Twins, Saudi Vision 2030.

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.

1. INTRODUCTION

Reliability in smart buildings is often discussed as a technology problem, although its consequences are operational. An air handling unit that drifts out of calibration may remain functional while consuming extra electricity, disturbing thermal comfort and generating repeated complaints. A chiller fault can threaten service continuity even when local controls continue to report normal operation. Contemporary reviews describe considerable promise for data driven fault detection, yet also demonstrate how strongly performance depends on data quality and deployment conditions [1–4]. It is to decide when intervention is warranted, what evidence supports that decision and whether the intervention actually preserves building performance.

The Saudi setting intensifies this problem. Cooling intensive commercial, institutional and mixed use buildings combine extended operating seasons with high reliability expectations. Large portfolios may include different equipment vintages, proprietary building management systems (BMS), subcontracted maintenance and varying degrees of sensor coverage. A method tuned to one air handling unit may not transfer to another; an apparent anomaly may reflect a change in occupancy, weather or control sequence rather than degradation. Reviews of fault diagnosis and digital twins consistently identify interoperability, transferability and contextual interpretation as barriers to adoption [5–8]. These constraints make maintenance engineering judgment indispensable even when sophisticated models are available.

This paper treats HVAC and associated MEP services as an integrated operational system. HVAC encompasses chillers, pumps, fans, coils, valves, air handling units and controls; related MEP assets include electrical distribution, generators, water systems and fire protection interfaces. The research literature is most mature for HVAC and building energy systems, so the evidence is not assumed to transfer automatically to every mechanical, electrical or life safety subsystem. The contribution is a critical, asset oriented synthesis connecting AI methods to maintenance governance, reliability evidence and implementation decisions. This framing complements broader examinations of digital building operation and decarbonisation while concentrating on the maintenance actions required to make predictive insights useful [9–12].

2. Aim, Objectives and Research Questions

The aim is to evaluate how AI supported predictive maintenance can improve the reliability and operational performance of HVAC and interconnected MEP assets in Saudi smart buildings, while identifying the conditions under which technically promising models become defensible engineering decisions.

The objectives are to: (i) classify AI and hybrid approaches used for fault detection, diagnosis and remaining useful life assessment; (ii) establish the data and contextual information needed for dependable building scale deployment; (iii) compare how the literature evaluates reliability, comfort, energy and cost effects; (iv) analyse operational, organisational and cybersecurity barriers; and (v) develop a context sensitive implementation and validation framework for Saudi facility portfolios.

Three questions guide the synthesis. RQ1: Which approaches are best suited to common HVAC and building service failure mechanisms? RQ2: What separates promising diagnostic accuracy from demonstrated maintenance value? RQ3: Which technical and managerial safeguards are necessary for accountable deployment across heterogeneous Saudi buildings?

3. REVIEW METHODOLOGY

3.1 Review Design and Evidence Boundary

A structured critical narrative review was selected because the subject combines machine learning classification, engineering reliability, BMS integration and organisational maintenance practice. The goal was to compare mechanisms and practical applicability rather than pool incompatible performance statistics. The evidence base contains 30 peer reviewed journal publications published from 2020 through 2025. The two reference articles supplied for this manuscript informed its scholarly organisation and Saudi context framing; they were not treated as evidence that a separate systematic database search had occurred. Publisher records, article metadata and references cited within the supplied papers were used to locate and cross check the selected literature. This process supports a transparent, purposive synthesis but does not constitute an exhaustive systematic review or a meta analysis.

3.2 Search Concepts and Eligibility

Search concepts combined the terms predictive maintenance, automated fault detection and diagnosis, HVAC, chillers, air handling units, smart buildings, building management systems, digital twin, machine learning, operational energy and facility management. Titles and abstracts were screened for an explicit connection between building operation and at least one of these domains: detection of abnormal operation, identification of faults, estimation of degradation, maintenance scheduling, asset information integration or evaluation of operational performance. Studies of purely industrial machinery were excluded unless their method was directly evaluated in the built environment. General construction digitalisation studies were retained only where they illuminated implementation constraints relevant to operation and maintenance.

Publications outside the 2020–2025 window were not included in the numbered reference list. Manufacturer brochures and unverified performance claims were not used as evidence. Because underlying datasets, search logs and independent full text eligibility decisions were not available for every source, no fabricated identification counts, reviewer agreement statistics or screening flow were reported. This limitation is stated explicitly to prevent the appearance of a reproducible systematic review where the supporting audit trail does not exist.

3.3 Appraisal and Synthesis Procedure

Each source was appraised against five practical questions: What physical asset or building context was studied? Which telemetry, engineering labels and contextual features were required? Did evaluation distinguish normal operation, faults and changing control modes? Were the reported outcomes algorithmic, operational or economic? Finally, were limitations of real world transfer and maintenance integration acknowledged? Findings were compared narratively across sensor based detection, supervised classification, deep sequence modelling, digital twin integration and implementation studies. This evidence matrix informed Table 1 and the proposed lifecycle architecture in Figure 1. The synthesis avoids numerical aggregation because faults, training data, baselines and measurement boundaries are not sufficiently comparable across studies [1-13].

4. Conceptual Foundations of Predictive Building Maintenance

4.1 Distinguishing the Four Decision Levels

Condition monitoring records equipment state; anomaly detection flags departures from expected behaviour; diagnosis seeks a probable cause; prognosis estimates future degradation or remaining useful life; and maintenance optimisation chooses an appropriate intervention. These tasks are frequently grouped under predictive maintenance, but each requires different evidence. A classifier trained on simulated refrigerant leaks cannot automatically predict when an installed compressor will fail. Likewise, a residual from a temperature model may reveal a malfunction without identifying whether the cause is a faulty sensor, blocked coil, leaking valve or changed supervisory control. Recent assessments emphasise the need to separate fault recognition from failure consequences and repair decisions [2-14].

For building operators, the relevant unit of analysis is an asset service relationship rather than an isolated component. A chilled water pump affects cooling distribution; an unreliable sensor can produce misleading alarms; a faulty valve may encourage compensatory control that obscures a developing problem. Systems based diagnosis models and fault ontologies can preserve these relationships, improving consistency between analytics and engineers' causal reasoning [15, 16]. Risk assessment must also distinguish comfort and energy defects from hazards affecting electrical safety, ventilation integrity or critical services. Automated recommendations should not bypass mandated inspection, protection systems or competent engineering approval.

4.2 Reliability and Maintenance Outcomes

Reliability improvement is not captured by classification accuracy alone. Facilities need evidence about unexpected stoppages, mean time between failures, mean time to restoration, maintenance backlog, work order recurrence, asset availability and cost per operating hour. Energy consumption, indoor comfort, service level compliance and occupant disruption are complementary outcomes. The baseline matters: comparing a model with a poorly maintained schedule may overstate value, while comparison with experienced condition based maintenance is more demanding. Studies of real world prototypes demonstrate that operator usability, data exchange and maintenance workflow integration are as important as model selection [17, 18].

5. AI Methods and Their Engineering Applicability

5.1 Statistical Learning and Supervised Classification

Supervised approaches learn relationships between measured operating variables and labelled fault conditions. Typical inputs include supply and return temperatures, chilled water flow, valve position, fan speed, electrical current, pressure, setpoints and ambient conditions. Random forests, support vector machines and gradient boosted learners can capture nonlinear interactions and rank features with less data than some deep architectures. Their usefulness nevertheless depends on reliable fault labels and representative operating regimes. Experimental datasets with artificially introduced faults may contain distinctive signatures that are absent in field operation. Recent chiller and air handling unit studies therefore move beyond fault identification toward severity estimation and reduced false positive exposure [19, 20].

Model validity requires more than a randomly shuffled accuracy score. Time series samples from adjacent minutes are highly correlated, and equipment signatures recur across weather and occupancy cycles. If observations from the same fault episode appear in training and testing sets, measured performance can be misleading. Stronger designs separate buildings, assets or time periods and explicitly test model behaviour during seasonal peaks and control mode changes. Calibration and alert prevalence also deserve attention: a low false positive rate on balanced laboratory data can still produce excessive tickets when real faults are rare. Comparative reviews call for operationally meaningful metrics and common evaluation protocols [1-5].

5.2 Unsupervised, Sequence and Hybrid Approaches

Unsupervised methods help where confirmed failure labels are scarce. Clustering, autoencoders and residual based change detection learn normal operating envelopes and identify deviations; however, a deviation is not proof of a component fault. Sequence methods such as long short term memory networks can exploit temporal structure, which matters for progressive degradation, transient load changes and cyclic equipment behaviour. Published work uses digital twins and sequence models to detect emerging indoor climate problems, and field oriented work examines how abnormal air handling unit behaviour relates to energy use [21, 22]. These approaches remain vulnerable to concept drift when equipment is reconfigured or control sequences change.

Hybrid approaches combine physical insight with learned relationships. For example, an engineering rule can establish permissible operating modes, while a data driven model evaluates residuals only within those modes. Bayesian networks can connect abnormal measurements to plausible upstream causes; digital twin based workflows can attach semantic labels and equipment relationships to observations [15, 16]. The hybrid architecture is attractive in safety sensitive settings because it can expose assumptions and allow engineers to challenge a diagnosis. Yet model complexity does not guarantee performance. A transparent residual monitored over time may be preferable to an opaque deep model when asset coverage, training labels and maintenance capacity are limited.

5.3 Transfer Learning, Interpretability and Model Governance

A multi building portfolio rarely has identical chillers, sensors or maintenance histories. Transfer learning attempts to reuse knowledge from an information rich source asset in a target asset with fewer examples, but shifts in operating range, vendor control logic and instrumentation can invalidate the transfer. Research on cross building fault diagnosis and transformer based adaptation demonstrates the potential of this approach while also showing why rigorous domain testing matters [23, 24]. Explanations based on input importance may help prioritise inspections, although they should not be mistaken for physical proof of causation.

Operational model governance consequently needs clear ownership, version control and retraining triggers. A fault model should document its intended asset classes, data requirements, training boundary, diagnostic labels, confidence reporting and conditions under which it must abstain. Engineering teams should record whether alerts were accepted, rejected or escalated, thereby creating an auditable feedback loop. Reviews of machine learning FDD identify inconsistent metadata, missing labels and limited transferability as persistent obstacles; a maintenance programme cannot treat them as minor software details [3-6].

Table 1: Critical evidence map of AI approaches in building maintenance, based on the reviewed literature

Evidence domain	Typical asset/data	What the literature supports	Key validation need
Supervised fault classification	Chillers; AHUs; labelled BMS points	Recognises known fault patterns	Unseen buildings, balanced and rare faults
Anomaly and sequence models	Temperature, flow, vibration, current	Flags temporal deviation and degradation	False alerts, drift and time to action
Hybrid or Bayesian diagnosis	Topology, rules and sensor residuals	Links symptoms to plausible causes	Engineering confirmation and safety constraints
Transfer learning	Mixed equipment and control regimes	Potential adaptation across assets	Cross site testing and uncertainty
Digital twin integration	BIM, asset registers and work orders	Adds identity, context and traceability	Interoperability, governance and lifecycle cost

6. Asset Level Applications across HVAC and MEP

6.1 Chillers, Pumps and Hydronic Networks

Chillers generate substantial maintenance interest because degraded heat exchange, refrigerant problems, fouling and incorrect controls affect both service and electricity consumption. A diagnostic workflow combines refrigerant side and water side measurements, status points and operating context. Sensors should be checked before models infer compressor, condenser or valve faults. Severity aware studies show the value of distinguishing conditions that warrant immediate attention from low impact anomalies [19]. Maintenance decisions should account for redundancy: an early warning on a duty chiller in a single plant building may justify different action from the same warning on a standby machine.

Hydronic pumps and distribution systems add further complexity. Pump current, differential pressure, flow, speed and valve demand may expose poor control, cavitation risk or capacity degradation,

but their interpretation requires equipment curves and knowledge of operating modes. A spike in energy use does not uniquely identify mechanical wear. When developing predictive maintenance rules, engineers should test against known maintenance records, commissioning values and observed post repair change, rather than inferring avoided failures from algorithmic alerts alone. Real world facility digital twin prototypes support the feasibility of linking these information streams but also reveal the burden of information integration [17].

6.2 Air Handling Units, Terminal Devices and Ventilation

Air handling units illustrate the importance of multiple fault mechanisms. Temperature sensor bias, stuck dampers, filter loading, leaking heating or cooling coils, fan degradation and inappropriate setpoints can produce overlapping symptoms. Published automated diagnosis frameworks using BMS data report the potential to identify these states, yet operational conditions and sparse labels constrain generalisation [20]. A well designed alert should include asset identification, diagnostic evidence, affected operating conditions and a proposed verification check. Work order priority must reflect actual air quality or comfort consequences, not the novelty of the algorithm.

Ventilation failure also creates a reliability question beyond energy consumption. Excessive optimisation that reduces airflow without adequate verification may compromise environmental quality or statutory obligations. Models should respect minimum outdoor air requirements, smoke control interfaces, fire damper logic and equipment interlocks. Any AI supported adjustment to ventilation or life safety functions must remain subordinate to engineered controls and applicable codes. The literature on digital building operation supports integrated monitoring, but does not justify automatic intervention in safety critical systems without case specific validation [9-25].

6.3 Electrical and Wider MEP Interfaces

MEP service continuity depends on relationships across systems. Electrical switching, standby generation, building controls, water supply and fire protection can determine whether a cooling fault becomes a local discomfort event or a widespread disruption. The directly reviewed machine learning evidence mainly concerns HVAC and building energy assets. For other MEP equipment, transferable principles include sensor validation, asset criticality, work order traceability and condition trending, while specific diagnostic models require separate substantiation. Treating a building as an interconnected asset network helps avoid isolated local optimisation that transfers risk to another subsystem.

7. Integrated Smart Building Implementation Framework

Figure 1 translates the literature into an engineering decision sequence. The first layer is an asset register linking equipment identifiers, location, design duty, maintainability, failure criticality and service dependencies. The second layer obtains telemetry from BMS, submetering and inspection records with time synchronisation and quality checks. Third, analytics compare current observations with an asset and mode specific baseline. Fourth, a diagnostic layer combines model evidence with engineering rules and maintenance history. Finally, the system creates a reviewable recommendation rather than an uncontrolled actuation command. Confirmed outcomes return to the data store to update labels and measure whether intervention created value.

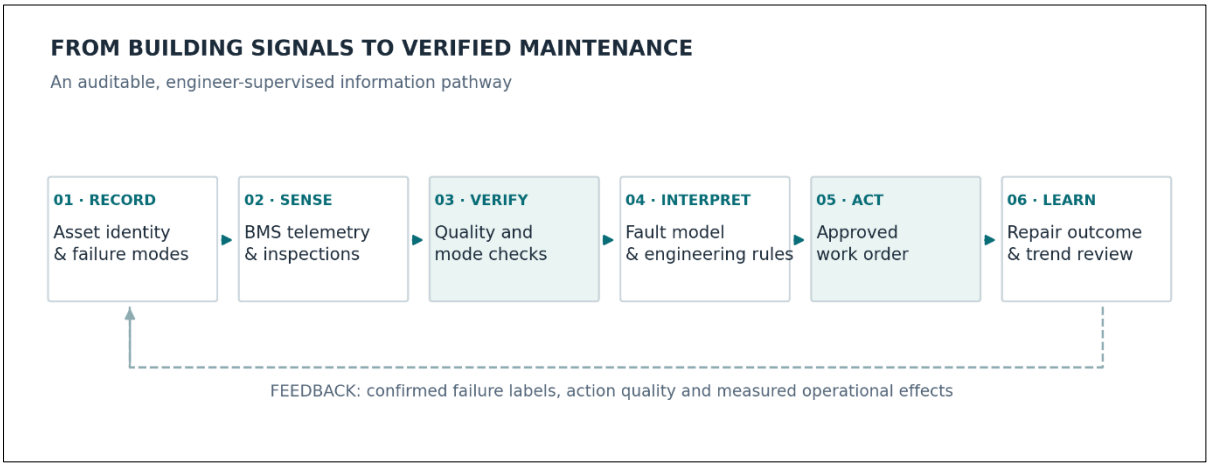


Figure 1: Conceptual workflow for AI assisted HVAC and MEP maintenance. Arrows represent decision and information flows rather than measured effects

The framework is intended for incremental rollout. Sites should first resolve asset naming, incomplete points lists and unclear responsibility for alarms. They can then establish normal ranges and a small number of high value use cases: persistent valve leakage, air handling unit fan degradation, chiller performance drift or recurrent sensor errors. Hybrid rules and data driven models may coexist; a site need not implement a comprehensive digital twin before beginning useful analytics. Research on BMS derived features, BIM knowledge models and digital twins indicates that contextual data can improve fault interpretation, but also increases integration effort [16-26].

Human oversight remains a designed component of the architecture. Maintenance engineers validate whether an alert agrees with recent operating history and physical inspection. Supervisors decide urgency using criticality, service level commitments, spares availability and access requirements. Vendors receive actionable work orders rather than unexplained risk scores. After intervention, the team checks the same sensor signals and service outcomes that triggered the alert. This closed loop turns predictive maintenance from a dashboard feature into a reliability management practice.

8. Synthesis of Evidence and Performance Measurement

The strongest convergence across reviews concerns the ability of machine learning to detect non obvious deviations in multi variable HVAC data [1–6]. Applied studies add evidence that diagnostic models can be integrated with BMS data, chiller and air handling unit monitoring, and building information models [19-26]. However, published performance figures are not interchangeable: they describe different buildings, faults, sensor combinations, training procedures and testing periods. Similarly, the availability of a digital twin is not evidence of improved reliability unless relevant intervention outcomes were measured.

Table 1 distinguishes technical capability from evaluation needs. Detection methods are comparatively mature where labels and sensors are available; prognosis and maintenance scheduling require longitudinal evidence of deterioration and completed repairs. This difference explains why facility managers should select outcome measures before choosing algorithms. Minimum evaluation should include the number and severity of confirmed incidents, detection lead time, false alerts per asset month, engineer investigation time, repeat faults, downtime hours, energy normalised for weather and occupancy, and maintenance expenditure. Where sufficient data exist, comparisons should account for changes in asset condition and portfolio composition rather than attributing every improvement to AI.

A credible pilot uses a pre specified baseline and intervention protocol. Comparisons may be made across matched equipment groups or sequential operating periods, but seasonality and coincident refurbishment must be disclosed. Maintenance costs should include deployment, licensing, data cleaning, engineering review and false ticket handling. Reliability outcomes should consider both avoided unplanned interruptions and new risks caused by unnecessary interventions. Published work on operator oriented digital twins and data driven energy workflows supports the need for end to end evaluation and feedback, rather than analytics in isolation [10-27].

9. Saudi Operational Context and Deployment Priorities

Saudi smart building portfolios need methods resilient to prolonged cooling demand, dust exposure, variation in occupancy, high daytime temperatures and differences between newly commissioned and ageing facilities. These conditions are contextual factors for model design, not universal failure rates established by this review. Design priorities include testing under peak cooling conditions, comparing duty and standby equipment, recognising changes in operating modes and protecting service continuity at critical sites. A recent Saudi oriented study of AI enabled digital twins identifies implementation issues involving infrastructure, standardisation and professional capability, reinforcing the need to coordinate technical models with organisational readiness [28].

Vision 2030 provides a strategic motivation to improve building operations through reliable services, efficient resource use and digital capability. The direct contribution of a proposed predictive maintenance programme must nevertheless be demonstrated at the level of measurable facility outcomes. The review does not claim that a specific algorithm is required by national policy or that digitalisation itself proves compliance. Facilities departments should define a small number of decision relevant indicators, obtain permission to use operational data and preserve appropriate segregation between analytics networks and safety or control systems.

An adoption sequence is given in Table 2. In the initial phase, portfolio leaders establish asset criticality and identify consistent records from BMS, computer aided facility management systems and work orders. The second phase checks data completeness and label credibility. The third phase runs the model in shadow mode, without automatic

dispatch, to observe alert frequency and technical plausibility. The fourth phase authorises targeted maintenance actions with engineer sign off. The final phase measures reliability, cost and energy outcomes and determines whether the model warrants expansion. Saudi research on digital twins and sustainable buildings is supportive of integrated information systems, but operational governance remains decisive [28–30].

Table 2: Staged implementation and acceptance gates proposed for Saudi facility portfolios

Stage	Required practice	Release criterion	Responsible function
1. Asset selection	Criticality, failure modes, service impact	Approved priority asset list	Facility engineering
2. Data readiness	Point mapping, sensor QA, repair labels	Auditable completeness threshold	BMS/CAFM teams
3. Shadow evaluation	Compare alerts with engineer assessment	Acceptable false ticket burden	Reliability lead
4.Controlled intervention	Approve work orders and safety checks	Confirmed diagnosis and sign off	Maintenance manager
5.Outcome verification	Track downtime, energy, cost and comfort	Measured benefit against baseline	Owner and operations

10. DISCUSSION AND RESEARCH GAPS

10.1 The Missing Link between Diagnostic Models and Avoided Failures

A recurring weakness is the assumption that identifying an abnormal signature establishes an economic benefit. In practice, early detection only creates value when the diagnosis is actionable, an intervention is feasible, and the repair prevents a more costly outcome. A model that repeatedly warns about minor sensor bias may increase maintenance workload rather than improve service. Conversely, a seemingly modest early warning signal can matter greatly for a single point of failure asset. Future studies should connect alerts to work orders, confirmed causes, intervention decisions, recurrence and observed service continuity. This evaluation should follow equipment across maintenance cycles rather than end at classifier testing.

10.2 Transferability and the Representativeness Problem

Models are often evaluated on highly structured research datasets, whereas operating buildings contain missing values, manual overrides, unusual sequences and incomplete maintenance records. The literature on transfer learning demonstrates how adaptation can help, but also exposes domain dependence [23, 24]. Multi building Saudi studies should report equipment type, age, sensing configuration, climate exposure, BMS vendor, operating modes and maintenance practice. Dataset partitioning should prevent leakage between related fault episodes. Reporting should distinguish an unseen time period at one site from a genuinely unseen site; these are materially different claims about generalisation.

10.3 Explainability, Governance and Human Factors

Explainability is relevant when engineers must justify a work order, review a vendor claim or reject a risky recommendation. Feature attribution graphics are helpful but insufficient if they ignore hydraulic and thermodynamic constraints. More informative explanations connect observed symptoms to equipment topology, operating mode and plausible causes. Building wide Bayesian reasoning, semantic fault tags and BIM linked digital twins offer routes toward such explanations [15-26]. Still, governance should specify who can override the model, how false alerts are classified and which outcomes trigger retraining. These decisions belong within maintenance quality assurance, not solely an information technology team.

10.4 Research Agenda for Hot Climate Building Portfolios

Priority research includes longitudinal field validation across diverse Saudi building typologies; calibrated prognosis using confirmed degradation and repair histories; uncertainty aware alerts that abstain on unfamiliar operating conditions; and joint appraisal of reliability, energy, comfort and workload. Comparative investigations should test interpretable baselines against advanced deep models rather than assuming complexity delivers higher operational value. Another need is practical integration between BMS event streams, asset registers and service contractor records. Interoperability and data ownership arrangements should be evaluated as factors affecting replicability and deployment cost. These questions are more relevant to engineering managers than isolated benchmark gains.

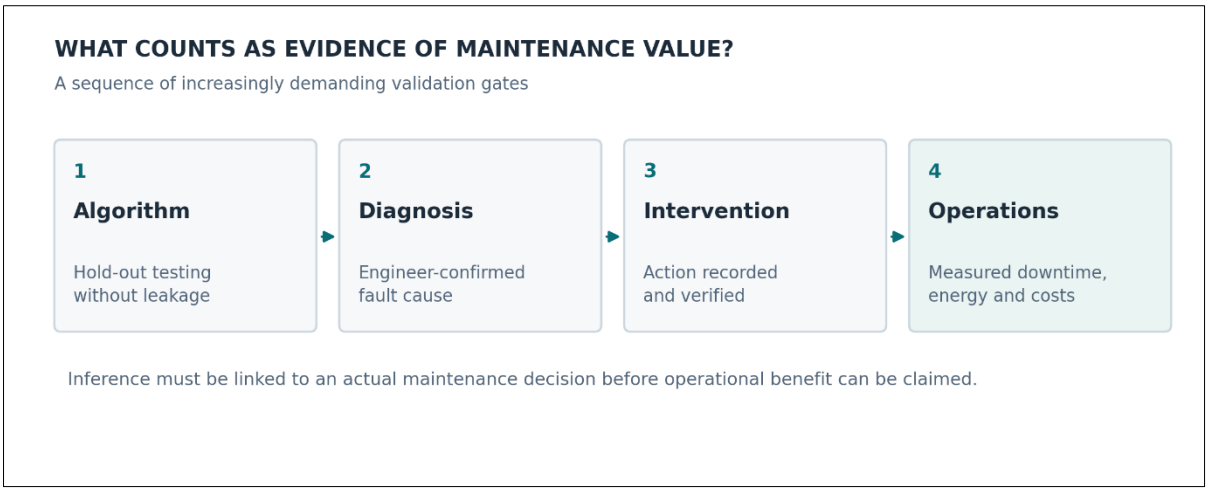


Figure 2: Proposed evidence maturity pathway from diagnostic feasibility to measurable maintenance value. The stages are conceptual, not quantitative results

11. Practical, Managerial and Policy Implications

For facility engineers, an implementable programme begins with failure mode analysis and a narrowly defined list of measurable faults. Sensors must be maintained as assets, with calibration checks and documented quality thresholds. For maintenance managers, the most important procurement questions concern verifiable baseline performance, integration with existing work order processes, alert triage effort and exit arrangements for proprietary analytics. Contracts should distinguish software availability from demonstrated equipment reliability. An operator should be able to retrieve the data and reasoning behind a maintenance recommendation after a service incident.

For building owners, investment evaluation should consider the whole lifecycle: instrumentation, data engineering, licences, validation, training and cyber controls against avoided downtime, energy use and emergency repairs. Portfolio wide expansion should follow a documented pilot rather than a blanket assumption of savings. For policymakers and professional bodies, common asset metadata, secure exchange formats, engineering competency development and transparent reporting templates may offer greater near term value than mandating a particular algorithm. These implications are consistent with the literature's emphasis on information quality and institutional readiness [9-30].

12. Limitations

This article is a critical review of 30 selected peer reviewed publications, not an exhaustive systematic review. Studies differ in fault definitions, operating conditions and outcome metrics, preventing defensible pooling of accuracy or economic gains. Several sources concern general built environment digitalisation rather than direct HVAC predictive maintenance trials. The synthesis therefore uses those papers to interpret data integration and governance, not to substantiate asset specific performance. Saudi specific field evidence remains limited within the selected set. No new operational dataset, maintenance experiment, field survey or validated cost model was produced. The framework and figures are conceptual; they require independent pilot evaluation before being described as empirically proven.

13. CONCLUSION

AI can strengthen HVAC and MEP maintenance when predictions are embedded within reliable engineering practice. The reviewed literature supports automated detection of complex anomalies, improved contextual diagnosis and emerging integration with building information models and digital twins. It does not establish that advanced algorithms consistently outperform disciplined baseline maintenance across every building or asset class. For Saudi smart buildings, successful implementation depends on quality telemetry, clear failure definitions, mode aware validation, accountable human decisions and a documented connection between alerts and completed repairs. A staged programme that begins with critical assets and evaluates downtime, cost, energy and service continuity offers a more defensible route than deploying technology without an operational baseline. Future research should prioritise longitudinal multi site evidence, transferability under high cooling loads and transparent quantification of realised maintenance value. Such work can contribute to the reliable, efficient and digitally capable building stock envisaged under Vision 2030.

REFERENCES

1. Chen, Z., O'Neill, Z., Wen, J., et al. (2023). A review of data driven fault detection and diagnostics for

building HVAC systems. Applied Energy, 339, 121030. https://doi.org/10.1016/j.apenergy.2023.121030

2. Bi, J., Wang, H., Yan, E., Wang, C., Yan, K., Jiang, L., & Yang, B. (2024). AI in HVAC fault detection and diagnosis: A systematic review. Energy Reviews, 3(2), 100071. https://doi.org/10.1016/j.enrev.2024.100071
3. Nelson, W., & Culp, C. (2022). Machine learning methods for automated fault detection and diagnostics in building systems—A review. Energies, 15(15), 5534. https://doi.org/10.3390/en15155534
4. Matetić, I., Štajduhar, I., Wolf, I., & Ljubic, S. (2023). A review of data driven approaches and techniques for fault detection and diagnosis in HVAC systems. Sensors, 23(1), 1. https://doi.org/10.3390/s23010001
5. Guo, Y., Liu, Y., Wang, Y., Wang, Z., Zhang, Z., & Xue, P. (2024). Advance and prospect of machine learning based fault detection and diagnosis in air conditioning systems. Renewable and Sustainable Energy Reviews, 205, 114853. https://doi.org/10.1016/j.rser.2024.114853
6. Melgaard, S. P., Andersen, K. H., Marszal Pomianowska, A., Jensen, R. L., & Heiselberg, P. K. (2022). Fault detection and diagnosis encyclopedia for building systems: A systematic review. Energies, 15(12), 4366. https://doi.org/10.3390/en15124366
7. Nelson, W., & Dieckert, C. (2024). Machine learning based automated fault detection and diagnostics in building systems. Energies, 17(2), 529. https://doi.org/10.3390/en17020529
8. 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
9. Zhang, A., Yang, J., & Wang, F. (2023). Application and enabling digital twin technologies in the operation and maintenance stage of the AEC industry: A literature review. Journal of Building Engineering, 80, 107859. https://doi.org/10.1016/j.jobe.2023.107859
10. Abuimara, T., Hobson, B. W., Gunay, B., & O'Brien, W. (2022). A data driven workflow to improve energy efficient operation of commercial buildings: A review with real world examples. Building Services Engineering Research and Technology, 43(4). https://doi.org/10.1177/01436244211069655
11. Arsecularatne, B., Rodrigo, N., & Chang, R. (2024). Digital twins for reducing energy consumption in buildings: A review. Sustainability, 16(21), 9275. https://doi.org/10.3390/su16219275
12. Jørgensen, B. N., & Ma, Z. G. (2025). Energy efficiency and decarbonization strategies in buildings: A review of technologies, policies, and future directions. Applied Sciences, 15(21), 11660. https://doi.org/10.3390/app152111660
13. Cespedes Cubides, A. S., & Jradi, M. (2024). A review of building digital twins to improve energy efficiency in the building operational stage. Energy Informatics, 7, 11. https://doi.org/10.1186/s42162 024 00313 7
14. Es sakali, N., Zoubir, Z., Kaitouni, S. I., Mghazli, M. O., Cherkaoui, M., & Pfafferott, J. (2024). Advanced predictive maintenance and fault diagnosis strategy for enhanced HVAC efficiency in buildings. Applied Thermal Engineering, 254, 123910. https://doi.org/10.1016/j.applthermaleng.2024.123910
15. 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, 304, 113889. https://doi.org/10.1016/j.enbuild.2024.113889
16. Xie, X., Merino, J., Moretti, N., Pauwels, P., Chang, J. Y., & Parlikad, A. (2023). Digital twin enabled fault detection and diagnosis process for building HVAC systems. Automation in Construction, 146, 104695. https://doi.org/10.1016/j.autcon.2022.104695
17. Asare, K. A. B., Liu, R., Anumba, C. J., & Issa, R. R. A. (2024). Real world prototyping and evaluation of digital twins for predictive facility maintenance. Journal of Building Engineering, 97, 110890. https://doi.org/10.1016/j.jobe.2024.110890
18. Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. IEEE Access, 8, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358
19. Prabhu Bam, R. V., Prabhu Gaonkar, R. S., & George, C. P. (2024). A machine learning framework for detection and severity estimation of faults for chillers and air handling units in HVAC systems. Energy and Buildings, 313, 114235. https://doi.org/10.1016/j.enbuild.2024.114235
20. Babadi Soultanzadeh, M., Ouf, M. M., 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
21. Hu, W., Wang, X., Tan, K., & Cai, Y. (2023). Digital twin enhanced predictive maintenance for indoor climate: A parallel LSTM autoencoder failure prediction approach. Energy and

Buildings, 301, 113738. https://doi.org/10.1016/j.enbuild.2023.113738
22. Yeom, S., Kim, J., & Hong, T. (2025). Hybrid AI model for fault detection and energy consumption analysis of air handling unit systems with supervised and unsupervised learning. Building and Environment, 113272. https://doi.org/10.1016/j.buildenv.2025.113272
23. Chen, L., Li, G., Liu, J., et al. (2024). Fault diagnosis for cross building energy systems based on transfer learning and model interpretation. Journal of Building Engineering, 91, 109424. https://doi.org/10.1016/j.jobe.2024.109424
24. Wang, Z. C., Li, D., Cao, Z. W., Gao, F., & Li, M. J. (2024). A modified transformer and adapter based transfer learning for fault detection and diagnosis in HVAC systems. Energy Storage and Saving, 3(2), 96–105. https://doi.org/10.1016/j.enss.2024.02.004
25. Omrany, H., Al Obaidi, K. M., Husain, A., & Ghaffarianhoseini, A. (2023). Digital twins in the construction industry: A comprehensive review of current implementations, enabling technologies, future directions. Sustainability, 15, 10908. https://doi.org/10.3390/su151410908
26. Hosseini Gourabpasi, A., & Nik Bakht, M. (2024). BIM based automated fault detection and diagnostics of HVAC systems in commercial buildings. Journal of Building Engineering, 87, 109022. https://doi.org/10.1016/j.jobe.2024.109022
27. Almatared, M., Sulaiman, M., Alghamdi, A., & Nasrallah, E. (2025). An artificial intelligence based predictive maintenance strategy using long short term memory networks for optimizing HVAC system performance in commercial buildings. Buildings, 15(22), 4129. https://doi.org/10.3390/buildings15224129
28. Alnaser, A. A., & Elmousalami, H. (2025). Benefits and challenges of AI based digital twin integration in the Saudi Arabian construction industry: A correspondence analysis (CA) approach. Applied Sciences, 15, 4675. https://doi.org/10.3390/app15094675
29. Alnaser, A. A., Maxi, M., & Elmousalami, H. (2024). AI powered digital twins and Internet of Things for smart cities and sustainable building environment. Applied Sciences, 14, 12056. https://doi.org/10.3390/app142412056
30. Piras, G., Muzi, F., & Tiburcio, V. A. (2024). Digital management methodology for building production optimization through digital twin and artificial intelligence integration. Buildings, 14, 2110. https://doi.org/10.3390/buildings14072110.