Sustainable HVAC Operations and EnergyEfficient Maintenance Frameworks for Smart Buildings in Saudi Arabia under Vision 2030

Abstract

Coolingintensive buildings in Saudi Arabia face a linked operational challenge: reducing electricity consumption without compromising thermal comfort, ventilation, equipment reliability, or service continuity. This critical review examines sustainable heating, ventilation and air conditioning (HVAC) operations and energyefficient maintenance frameworks for existing and emerging smart buildings within the objectives of Saudi Vision 2030. Literature published during 2020–2025 is organized across climateresponsive operation, supervisory control, condition monitoring, fault detection, maintenance decisionmaking, and implementation governance. The analysis distinguishes demonstrated technical capability from deployable facilitymanagement practice and avoids treating simulated savings as guaranteed field outcomes. It identifies weaknesses in crossclimate transferability, inadequate fault labels, uncalibrated energy baselines, limited lifecycle costing, and incomplete representation of occupant health. A staged operationstomaintenance framework is proposed that combines reliable sensing, fault triage, human authorization, verified intervention, and continuous measurement. The central finding is that sustainable HVAC performance depends on institutionalizing verifiable operational routines alongside digital optimization. For Saudi portfolios, the priority is robust, climatespecific and maintainable integration rather than technology adoption alone. The review concludes with research priorities and measurable implementation indicators.

Keywords

Smart buildings; sustainable HVAC; energyefficient maintenance; predictive maintenance; fault detection and diagnosis; Saudi Arabia; Vision 2030; building energy management.

UNDER PEER REVIEW IN IJAR

1. Introduction

The strategic case for improving the performance of cooling systems in Saudi Arabia is unusually strong. Long cooling seasons, high outdoor temperatures, variable humidity and rapid development of buildings create sustained demand for mechanical conditioning. An efficient chiller or packaged air conditioner at installation can subsequently waste energy because of sensor drift, fouled coils, malfunctioning valves, inappropriate schedules, simultaneous heating and cooling, or uncontrolled setpoints. Equipment specifications alone consequently provide an incomplete picture of operational sustainability. The building must remain efficient through changing occupancy, weather, maintenance capacity and asset age. Krarti and Howarth [1] evaluated the economic relevance of higherefficiency airconditioning equipment in the Saudi residential stock, while Saudi retrofit research underlines the influence of the envelope on cooling demand [2,3].

Smart buildings add connected controls, submetering, analytics and increasingly predictive functions. These technologies enable closer observation of energy use, but they do not automatically produce improved outcomes. The quality of metadata, sensors, control sequences and commissioning determines whether a promising anomaly can be translated into a serviceable maintenance task. Abuimara et al. [4] emphasize the interdependence of metadata inference, fault diagnostics and occupantresponsive control. Chen et al. [5] similarly show why datadriven HVAC diagnostics must be examined in terms of process, equipment and evaluation metrics rather than model accuracy in isolation. In practice, digital intelligence can fail when technicians distrust warnings, faults are poorly prioritized, or energy improvements are not verified against a fair baseline.

Vision 2030 creates an appropriate policy context for improving resource efficiency, service quality and technological capabilities. Nevertheless, an alignment claim should be operationally testable: kilowatthours avoided, peak demand reduced, comfort hours maintained, failures prevented and maintenance expenditure understood. This review therefore focuses on the interface between everyday HVAC operations and lifecycle maintenance. It evaluates how predictive control, conditionbased intervention and facility management can work together in Saudi smart buildings. The contribution is a decision framework that separates measured evidence, transferable technical principles and implementation proposals instead of presenting overseas or simulated results as achieved Saudi savings.

2. Aim, objectives and research questions

The aim is to critically synthesize recent research on sustainable HVAC operations and energyefficient maintenance and to derive an implementable framework for smart buildings in Saudi Arabia. The objectives are to characterize climaterelated operating constraints; compare

UNDER PEER REVIEW IN IJAR

control and diagnostic approaches; evaluate maintenance prioritization and measurement practices; examine tradeoffs among energy, reliability, comfort and cost; and identify research requirements for locally validated implementation. Three questions guide the review. First, which operational and maintenance interventions have the clearest physical or empirical rationale for coolingdominated buildings? Second, how should buildingmanagement data be converted into trustworthy and actionable decisions? Third, what governance and evaluation arrangements are necessary to scale these interventions without compromising indoor environmental quality?

These questions deliberately emphasize sustained performance rather than the technical novelty of individual algorithms. A model that detects a faulty sensor but does not trigger reliable inspection has limited facilitylevel benefit. Likewise, a control optimization that reduces cooling consumption at the expense of uncomfortable or inadequately ventilated occupied spaces should not be characterized as a successful sustainability intervention. Evaluating the whole causal chain connects energy engineering to operational accountability.

3. Review methodology

A structured critical narrative review design was selected because the relevant evidence crosses experimental controls, simulation studies, diagnostic reviews, policyoriented analyses and operational case studies whose endpoints are not sufficiently uniform for a meaningful pooled effect estimate. The review window was restricted to publications dated 2020 through 2025. Search concepts were organized into four combinations: (HVAC OR cooling OR chiller) AND (energy efficiency OR sustainable operation); (fault detection OR predictive maintenance OR condition monitoring) AND building; (model predictive control OR artificial intelligence OR building management system) AND HVAC; and (Saudi Arabia OR hot arid OR hot humid) AND building energy. Academic discovery locations appropriate for replication include Scopus, Web of Science, ScienceDirect and Google Scholar, with backward reference checks from key reviews.

Eligibility priorities were peerreviewed studies, transparent research methods, explicit relevance to building cooling or management, and operationally interpretable results. Studies limited to unrelated industrial plants, purely theoretical algorithms without plausible HVAC transferability, and publications outside the specified years were excluded from the substantive citation set. Saudi studies were privileged for climatic context; international studies were retained where they offered a transferable method rather than direct evidence of Saudi performance. The two supplied sample papers informed sectioning and discussion scope, but the 2026 sample was not included among the 2020–2025 references. The 2025 façade simulation was treated as contextual evidence regarding envelope–cooling interaction, not as proof of HVAC maintenance savings [30].

UNDER PEER REVIEW IN IJAR

Extraction fields included climate, building type, system boundary, intervention, outcome definition, validation design, energy baseline, operational constraints and stated limitations. Evidence was thematically synthesized under demand reduction, control, diagnostics, maintenance, measurement and governance. Interpretation privileges field validation over simulation when making performance claims and distinguishes normalized energy from absolute consumption. No original experimental measurements, formal exhaustive database screening count or statistical metaanalysis are claimed. This is a reproducible review protocol and critically curated synthesis, not a completed registered systematic review. Methodological rigor is strengthened through explicit scope, source traceability and careful assessment of transferability rather than invented screening statistics.

4. Saudi climatic conditions and demandside context

Saudi Arabia cannot be represented by a single generic desert weather file. Inland cities often face intense dry heat and large daily swings, while coastal locations experience high latent cooling loads because of humidity. Building activity varies across offices, hospitality, healthcare, retail, education and residential developments. A maintenance plan suitable for an intermittently occupied office may be unsafe or ineffective in a hospital that operates continuously. Regionalized weather data, occupancy schedules and asset inventories should therefore precede portfoliowide algorithm deployment. Climate adaptation research highlights the risks of applying standard designs and forecasts without testing local conditions [6].

Reducing the cooling load upstream also improves the performance of maintenance programs. Roof and wall insulation, solarcontrol glazing, external shading, reduced infiltration and appropriate ventilation design can lower equipment operating hours or peak loads. Saudi studies on façade retrofits and insulation materials demonstrate the need to consider both thermal performance and lifecycle impacts [2,3,7]. These envelope actions complement operational changes but must not be conflated with HVAConly savings. An energy audit should distinguish reductions caused by fabric upgrades, replacement of plant, control adjustments, changed occupancy and weather. A combined project may deliver valuable overall savings without establishing which subsystem produced them.

Dust and airborne contamination introduce additional operational concerns. Filters, heattransfer surfaces and outdoor air intakes require sitespecific inspection, particularly where sand and construction activities are common. Blanket calendar replacement can be wasteful, yet delayed maintenance can increase pressure drop, fan power, complaints and risk to equipment. Differential pressure and actual filter condition should inform intervention, subject to health and manufacturer requirements. Humid coastal conditions add condensate management and microbialcontrol obligations. Sustainability should therefore include careful maintenance of air quality and moisture control rather than simply relaxing ventilation to reduce energy.

UNDER PEER REVIEW IN IJAR

5. Operational optimization beyond static setpoints

Conventional HVAC control frequently uses fixed schedules, manually tuned setpoints and reactive overrides. These approaches are understandable and maintainable, but their rigidity wastes energy when occupancy and weather differ from design assumptions. Demandbased schedules, optimized start–stop logic, supplyair temperature reset, chilledwater temperature reset and coordinated plant staging can reduce unnecessary run time when implemented with adequate safeguards. The engineering principle is to provide the needed environmental service with the least appropriate resource input rather than to seek the lowest energy figure independently of service quality. Buildingautomation strategies should be commissioned as sequences of operation and checked routinely against sensor trends [4,8].

Model predictive control uses a representation of building dynamics and forecasts of disturbances to optimize future control actions over a moving time horizon. Reviews by Taheri et al. [9] and Kim et al. [10] describe its ability to handle interacting variables, while also noting the burden of model development, tuning and practical deployment. Its most credible uses in coolingdominated facilities include anticipating load peaks, coordinating thermal storage, limiting simultaneous equipment demands and maintaining comfort constraints. Yet forecasts can be inaccurate and models can drift after equipment replacement or occupancy changes. A robust controller requires fallback modes, constraint monitoring, operator visibility and periodic recalibration. A sophisticated optimization model without these provisions may be less sustainable than a wellcommissioned conventional sequence.

Datadriven predictive control can learn patterns without complete firstprinciples models, but extrapolation beyond training conditions remains a weakness. Xin et al. [11] compare modelbased and modelfree strategies and clarify why neither universally dominates. In a Saudi heatwave, a model trained mainly on milder weather can make unreliable recommendations precisely when failures are most consequential. Control proposals should be tested against seasonal holdout data, highoccupancy episodes and abnormal plant conditions. Moreover, automated changes should be reversible and accompanied by a clear explanation of their intended effect, particularly in critical buildings.

6. Fault detection, diagnosis and condition monitoring

Fault detection asks whether system behavior is abnormal; diagnosis asks which underlying failure plausibly caused it; prognosis estimates how the condition may evolve. These are distinct activities with different evidential requirements. A temperature residual may indicate a stuck valve, wrong sensor placement, air in the circuit or a changed schedule. Treating every anomaly as a specific failed component risks needless parts replacement. Chen et al. [5] classify datadriven diagnostic methods, while the systematic reviews of deep learning and broader

UNDER PEER REVIEW IN IJAP

artificial intelligence show considerable methodological development alongside persistent validation challenges [12,13]. The usefulness of a diagnostic alert depends on precision, recall, falsealarm workload, fault severity and physical interpretability.

Sensor quality is a firstorder constraint. Miscalibrated temperatures, missing flow measurements, time synchronization errors and unreliable control point names can propagate through even technically accurate predictive models. Zhang et al. [14] describe how sensor selection influences diagnostic feasibility and cost. Good practice begins with a point inventory, common naming rules, calibration intervals and automated dataquality checks. Historical baselines should record equipment states and maintenance events so that a postrepair change can be interpreted. Monitoring should focus on observable failure signatures: increasing compressor cycling, abnormal approach temperatures, changing fan pressure, persistent valve saturation, unexpected simultaneous loads and equipment operation outside occupancy periods.

Deep networks can discover nonlinear and temporal signatures, yet benchmark accuracy does not establish field robustness. Taheri et al. [15] examine recurrent architectures for HVAC diagnosis, while Wang [16] evaluates transformerbased diagnosis using labeled operational airhandlingunit data from a continuously operating hospital. Such work advances realistic evaluation, but transfer from one hospital or equipment family to a Saudi hotel remains uncertain. Different sensors, control conventions and fault prevalence can reduce diagnostic reliability. An operational system should therefore expose uncertainty, allow engineers to contest alerts and update fault labels after actual inspection. Human feedback is a design requirement, not an afterthought.

7. Maintenance as an energymanagement intervention

Maintenance is often evaluated through labor expenditure, workorder closure or asset availability, whereas sustainability studies emphasize energy use and carbon. The two perspectives should be connected. A deteriorating condenser, misconfigured economizer or drifting actuator can increase electricity demand before a complete functional failure occurs. Proactive intervention can restore efficiency while reducing the chance of service disruption. However, predictive maintenance is only preferable when its monitoring and intervention costs are proportionate to the consequences of failure. Criticality, remaining useful life, repairability, spares and safety should determine the depth of monitoring for each asset [5,17].

A workable hierarchy begins with preventive tasks required for safety, hygiene and warranty compliance. Conditionbased maintenance then uses measurements, alarms and inspections to decide when servicing is technically justified. Predictive methods add a forecast of future degradation, but only where a valid historical record supports inference. The tiers should coexist. Filters may be changed according to measured pressure drop within prescribed limits; bearings may be inspected after vibration anomalies; controls faults may be corrected following trend

UNDER PEER REVIEW IN IJAR

analysis; and essential emergency equipment may retain mandatory timebased tests. Replacing all planned maintenance with an algorithm would remove useful engineering safeguards.

A critical implementation challenge is the separation between building automation and computerized maintenance management systems. Analytics produce little benefit when an alert is not assigned an owner, linked to asset identification, ranked by consequence and followed through to verification. Each credible fault should create a traceable chain: evidence, diagnosis, authorization, intervention, functional test and energy or reliability outcome. Abuimara et al. [4] support this operationally integrated perspective, and recent reviews note that scalable diagnostic systems require more than strong prediction scores [13,17]. A maintenance framework should also record alerts that were false, deferred or resolved without repair, because these cases reveal the practical workload and future training needs.

Table 1. Operational faults and maintenance responses in coolingdominated buildings.

System / signal	Possible explanation	Verification and action	Performance safeguard
AHU, filter pressure increasing	Filter loading or sensor drift	Inspect filter and validate differentialpressure sensing	Maintain air quality and minimum airflow
Chiller, increasing kW per cooling output	Fouled heat exchanger, controls, degraded efficiency	Check operating points, flow, temperatures and service logs	Compare similar loads and weather
Variableairvolume box, persistent reheat	Faulty damper, setpoint conflict, programming error	Trend airflow, damper command and reheat demand	Preserve comfort and ventilation
Pump or fan cycling frequently	Incorrect staging or unstable feedback	Review sequence, sensor and actuator response	Check equipment wear and safety limits
Occupied cooling complaints	Sensor error, insufficient capacity, poor distribution	Check occupiedzone conditions before reset	No energy savings at comfort expense

8. Integrated energy and maintenance performance

UNDER PEER REVIEW IN IJAR

Performance assessment should use multiple mutually constraining indicators. Site electricity consumption is necessary but susceptible to shifts in weather, occupancy and operating hours. Weathernormalized kilowatthours per square meter, coolingsystem coefficient of performance, chiller plant kilowatts per ton of refrigeration and fan specific power provide more informative comparisons when measurement boundaries are stable. Asset availability, mean time between failures, mean repair time and backlog of highrisk faults capture reliability. Thermal comfort exceedance hours, relative humidity and ventilation compliance are guardrails against energyonly optimization. Environmental assessment can add refrigerant leakage and emissions intensity, while financial review should include lifecycle costs rather than simple equipment payback [18,19].

The baseline must be explicitly defined before an intervention. A short comparison between an exceptionally hot week and a mild week is not credible evidence of savings. Where possible, use adequate preintervention metering, comparable operating regimes, weather normalization and a documented treatment of coincident changes. Controls reprogramming, equipment replacement and occupant behavior shifts should be logged separately. Savings estimates also need uncertainty bounds and a postintervention persistence check. A model may initially perform well and then lose its advantage following a sensor replacement, a schedule override or a staff change. Continuous commissioning is therefore an evidence process, not a onetime technical activity.

Carbon claims require equal care. A decrease in electrical consumption does not imply a fixed percentage reduction in emissions unless the applicable electricity emissions factor and calculation period are specified. Similarly, onsite photovoltaic generation offsets purchased energy but does not necessarily improve an HVAC unit's intrinsic efficiency. Evaluations should report gross HVAC consumption, onsite generation and purchased electricity separately. This boundary discipline improves the interpretability of project claims and supports responsible Vision 2030 reporting. It also allows decision makers to distinguish investments that reduce peak electrical demand from those that primarily shift energy usage between hours.

Table 2. Suggested performance scorecard for Saudi smartbuilding HVAC operations.

Dimension	Indicator	Required comparison	Interpretation warning
Energy	Weathernormalized HVAC kWh/m²; plant kW/RT	Documented baseline, occupancy and weather	Do not mix gross and net electricity
Reliability	Critical failures; downtime; repeat	Same asset and	Improved detection may raise initial alarm

UNDER PEER REVIEW IN IJAR

	fault rate	reporting interval	counts
Comfort / health	Temperature and humidity excursions; ventilation compliance	Occupiedperiod thresholds	Energy reduction is not beneficial if service degrades
Maintenance	Confirmed faults closed; false alarm burden	Inspection logs and technician hours	Alert count alone is not success
Financial / carbon	Total lifecycle cost; emissions estimate	Prices, system boundary and emissions factor	Assumptions and uncertainty must be declared

9. Crossstudy interpretation and evidence limitations

Across the reviewed studies, agreement is strongest at the level of engineering direction: buildings benefit from lower avoidable cooling loads, accurate monitoring, fault correction, appropriate controls and measurementbased verification. Agreement is weaker concerning universal percentage savings. Different reference systems, climates, maintenance histories and occupancy constraints can generate widely divergent results. The simulated smartfaçade study set in Dammam, for example, reports a reduction in cooling demand under its defined model assumptions [30]. That result is informative about the cooling load and solargain interaction, but it cannot be extrapolated as the expected savings from predictive HVAC maintenance. Its simulation boundary excludes several realworld degradation and organizational effects relevant to ongoing facilities management.

Buildingdesign literature adds another important caution. AIassisted optimization and surrogate models can accelerate exploration of envelope and control configurations, but predictions inherit assumptions embedded in the simulation or training data [20,21,22]. Research addressing hotclimate retrofit decisions [2,23] suggests that operational savings should be interpreted within a broader building system. A poorly insulated envelope can force efficient HVAC equipment to run longer; conversely, a retrofit that reduces demand can create an opportunity to resize or better stage plant. This crosssystem view does not justify adding separate estimated savings from different studies because interventions interact and their reported baselines are rarely comparable.

There is also a publication and deployment bias. Technical articles commonly report welldefined testbeds, data availability and successful implementations, whereas the facility management population includes legacy buildings, incomplete documentation, poorly configured networks

UNDER PEER REVIEW IN IJAR

and constrained staff budgets. Negative or inconclusive deployment outcomes are less visible. Consequently, the review treats claimed accuracy or savings as conditional on validation design and avoids numerical synthesis where study boundaries differ. Future research should publish failed deployment lessons, maintenance labor requirements and the durability of savings across several cooling seasons.

10. Proposed Saudi HVAC operations and maintenance framework

The proposed framework has six linked stages, summarized in Figure 1. Stage one establishes the asset and service baseline: identify systems, document control sequences, map sensors and characterize occupied comfort and ventilation requirements. Stage two cleans and validates data, including point units, timestamps and calibration status. Stage three establishes a performance baseline by climate zone, season and building use, with explicit measurement uncertainty. Stage four runs a layered analytics stack: transparent rulebased checks first, calibrated anomaly detection second and predictive control only where justified. Stage five converts prioritized alerts into inspected and authorized maintenance actions. Stage six verifies the effect on energy, comfort and reliability and returns confirmed labels to the diagnostic system. This sequence deliberately makes feedback and governance integral rather than optional.

Implementation should proceed through a constrained pilot rather than simultaneous rollout. Select a building with dependable metering, engaged operators and equipment whose likely faults have clear physical signatures. Review seasonal trends, interview technicians about recurring issues, and target a few interventions with measurable consequences. Comparison should include the labor required to handle false alarms as well as energy reductions. After verification, expand to similar assets only when point mappings and climate differences are addressed. Portfolio governance can define common naming, rolebased permissions, data retention, cyber controls, supplier responsibilities and change logs. These arrangements determine whether a smart building remains operable when a vendor contract expires.

Maintenance workorder priority can be defined through a transparent score combining safety consequences, service criticality, estimated energy waste, confidence in diagnosis and repair effort. The score should not replace professional judgment: a suspected ventilation failure in an occupied hospital demands attention even if its estimated energy penalty is small. Conversely, a minor nighttime scheduling fault may offer a lowrisk early efficiency improvement. Figure 2 illustrates the conceptual relationship between operating discipline, diagnostic confidence and verified savings. It is a framework schematic rather than an empirical performance chart. A governance committee should review recurrent faults, operator overrides, comfort complaints and unverified savings at planned intervals.

UNDER PEER REVIEW IN IJAR

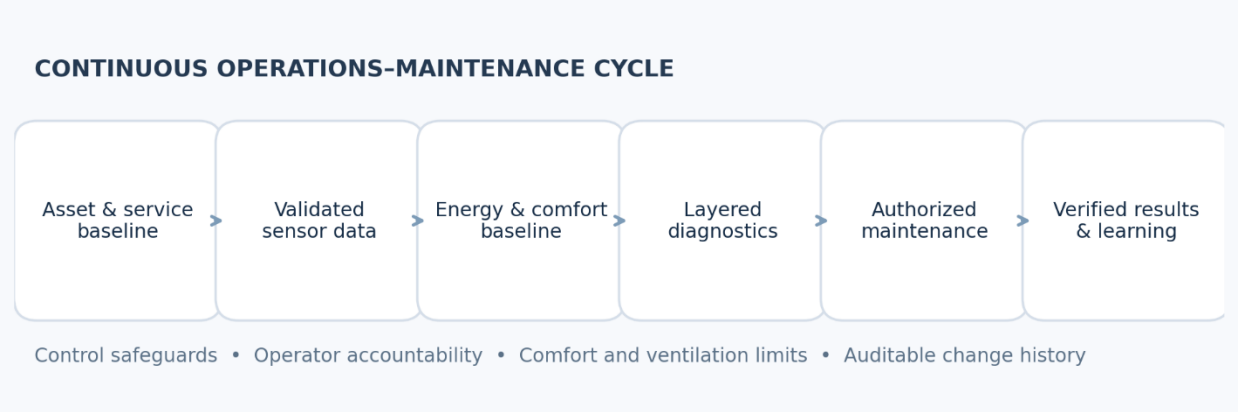


Figure 1. Proposed sixstage HVAC operations and maintenance cycle (authordeveloped conceptual figure).

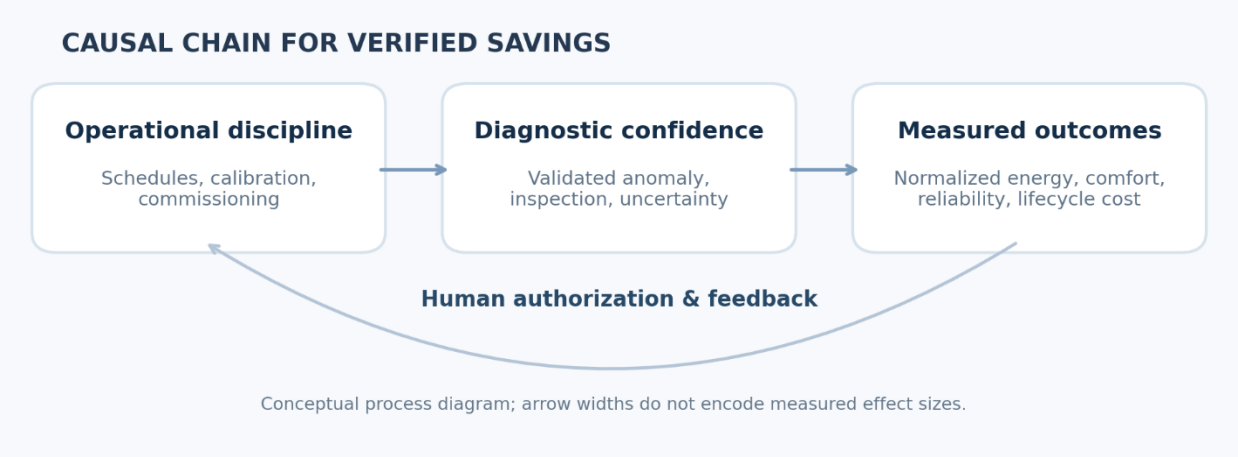


Figure 2. Conceptual relationship between operations, diagnostic reliability and validated outcomes (not empirical effectsize data).

11. Managerial and policy implications

Facility owners should specify operational performance at procurement, not merely equipment ratings or the presence of a smart control platform. Contracts can require reliable metering, open data access, documented control sequences, fault resolution workflows and seasonal acceptance tests. Performancebased service clauses should reward verified outcomes while protecting

comfort and health. Owners should also require access to historical trend data and sensible offboarding arrangements, limiting dependence on proprietary dashboards. For operators, training must join control theory with routine fault diagnosis, refrigerant handling, air quality and data literacy. Reliable human decisions remain decisive when algorithms encounter unfamiliar conditions.

For Saudi policymakers and portfolio administrators, a practical priority is standardizing measurable reporting across public and private building types. Rather than assuming that one algorithm fits every region, guidance could encourage local calibration for Riyadh, Dammam, Jeddah and other climate contexts. Building and energy information models can support lifecycle decisions when they are connected to real asset registers rather than used as static handover files [24,25]. Publicly reported demonstration projects should disclose measurement periods, baseline definitions, operating constraints and uncertainty. These details would enable replication, comparison and professional learning, thereby translating national sustainability ambitions into an auditable engineering practice.

11.1. Phased implementation and field verification

A portfolio evaluation should separate quick operational corrections from capitalintensive changes. In the first month, facility teams can inspect scheduling exceptions, nighttime consumption, sensor calibration and frequent overrides. The second phase can test supply temperature resets, coordinated staging, filterpressure monitoring and fault dashboards, provided comfort and ventilation conditions are held within approved limits. A later phase may evaluate variablespeed retrofits, equipment replacement and thermal storage against normalized baseline demand. This ordering keeps lowcost, reversible adjustments visible before major procurement commitments. It also reduces the risk of attributing savings from a replacement chiller to an analytics product installed simultaneously. The underlying lesson of Saudi school energy research is that building type and initial consumption must be documented before predictive methods are generalized [29]. A portfolio should include facilities with poor documentation rather than selecting only digitally mature demonstration buildings. Selection transparency is important because a narrowly successful pilot can overstate likely performance across an organization with heterogeneous assets.

Practical fault investigation should reconcile control trends with physical checks. A persistent high supplyair temperature may reflect a faulty sensor, poor chilledwater flow, incorrect valve command, heattransfer deterioration or an unusual occupancy load. An analyst can identify the time interval and correlated variables, but a technician must verify hydraulic conditions, air distribution and instrument readings before authorizing repair. After intervention, the same operating state should be sampled again to determine whether the problem actually resolved. If energy intensity remains unchanged, the team should reconsider the diagnosis rather than recording an assumed saving. This iterative process reduces wasted repairs and creates labeled

UNDER PEER REVIEW IN IJAR

examples grounded in engineering inspection. It also supports staff development: junior technicians learn how patterns connect to physical mechanisms instead of relying solely on dashboard fault labels. Where safety is implicated, the relevant operating procedure and professional authority must override algorithmic priorities.

For procurement and reporting, each project dossier should record the intervention boundary, dataset coverage, missingdata rate, comfort exclusions, seasonal adjustments, responsible personnel and independent verification method. A maintenance contractor should not be rewarded merely for generating alarms, nor penalized for correctly identifying faults that require ownerfunded replacement. Shared incentives can instead emphasize closure of verified highimpact faults, reduction of repeated service calls and sustained energy performance. Transparency is essential when a proposed control change affects the occupant experience. Operators should explain intended setpoint bands, exception procedures and the basis for responding to complaints. In large Saudi developments, these practices can be established during commissioning and carried into routine operation, avoiding the common separation between project delivery and longterm facilities management. The objective is a managed, trustworthy improvement cycle rather than a collection of isolated digital products.

12. Research gaps and future agenda

The highestpriority evidence gap is prolonged field validation of integrated operations and maintenance programs in Saudi buildings. Many methods are evaluated on simulated or narrowly labeled datasets; fewer studies demonstrate persistence over successive seasons, vendor changes and maintenance team turnover. Multisite research should compare office, residential, educational, hospital and hospitality assets across dry inland and humid coastal climates. Experimental designs should document preexisting efficiency, weather normalization and all major interventions. An evidence base assembled in this manner could begin to establish realistic distributions of savings rather than promotional point estimates.

A second gap concerns the joint optimization of health, comfort, water and energy. Ventilation reductions and warmer setpoints can decrease electricity use while creating adverse conditions for some occupants or uses. Future models should report humidity excursions, indoor air quality, comfort heterogeneity and operational exceptions alongside kilowatthours. A third gap is economic: diagnostic models rarely report the full cost of sensor maintenance, data engineering, falsealarm investigations and technician time. Comparative research should measure total lifecycle value and assess whether simple rules outperform complex methods when deployment cost is included. These are especially important questions for small or legacy buildings with limited budgets.

Other priorities include faultlabel interoperability, cyberresilient supervisory control, uncertaintyaware alarms, explainable recommendations and transfer learning tested under

UNDER PEER REVIEW IN IJAR

climate shift [26,27,28]. Emerging AI methods should be assessed against transparent conventional baselines and humanintheloop procedures rather than only against other highcomplexity models. Researchers should also study refrigerant leakage monitoring and embodied impacts of premature plant replacement, since operational electricity is not the sole environmental burden. Data sharing frameworks must preserve cybersecurity and occupant privacy while allowing independent evaluation. These directions would make algorithmic performance more relevant to the assetowner decisions that determine whether efficiencies endure.

13. Limitations

This review is a structured critical narrative synthesis rather than an exhaustively screened systematic review. The cited literature spans multiple building types, methods and performance definitions; it does not support a pooled national energysavings estimate. Several papers concern design, modeling or international field conditions and are used to derive transferable principles, not direct Saudi operating outcomes. The provided samples are used as structural references, but only one falls within the requested publication period. Future research may revise the relative importance of specific controls or algorithms as larger local field datasets become available. The proposed framework is conceptual and requires implementation testing before quantitative claims of effectiveness can be made. In evaluation reports, every percentage must have a named denominator, an identified period and an explicit comparator. Without those three elements, even a correct numerical value can mislead a decision maker. A credible sustainability claim should be reproducible by the operations team that maintains the system after the research group departs. This supports technical accountability across procurement, commissioning, operation and independent performance review. Local studies must test robust models under summer peaks, coastal humidity, missing data, irregular occupancy, equipment aging, maintenance delays, and service restrictions before drawing credible conclusions about persistent portfolio wide.

14. Conclusion

Sustainable HVAC management in Saudi smart buildings is best understood as an integrated operational capability. The literature supports the value of avoiding unnecessary cooling demand, commissioning controls, monitoring faults and acting on verified equipment conditions. It does not support a universal claim that any single AI tool guarantees energy savings. A credible strategy joins climatic calibration, accessible sensors, transparent diagnostics, disciplined work orders and continuous measurement while safeguarding comfort, indoor air quality and reliability. The proposed framework provides an adaptable path from basic monitoring to conditionbased maintenance and selective predictive optimization. For Vision 2030, the strongest

UNDER PEER REVIEW IN IJAR

contribution of smartbuilding technology will be demonstrated through repeatable, auditable performance improvement across actual Saudi building portfolios.

References

1. Krarti, M.; Howarth, N. (2020). Transitioning to high efficiency air conditioning in Saudi Arabia: A benefit cost analysis for residential buildings. Journal of Building Engineering, 31, 101457. https://doi.org/10.1016/j.jobe.2020.101457
2. Aloshan, M.; Aldali, K. (2024). Empirical study of facade retrofits for optimizing energy efficiency and cooling in school buildings in Saudi Arabia. Energy Reports, 12, 4105–4128.
3. Alsaqabi, Y. et al. (2023). Technoenvironmental assessment of insulation materials in Saudi Arabia: Integrating thermal performance and LCA. Buildings, 13, 331. https://doi.org/10.3390/buildings13020331
4. Abuimara, T.; Hobson, B. W.; Gunay, H. B.; O'Brien, W. (2022). A datadriven workflow to improve energy efficient operation of commercial buildings: A review with realworld examples. Building Services Engineering Research and Technology, 43(4). https://doi.org/10.1177/01436244211069655
5. Chen, Z. et al. (2023). A review of datadriven fault detection and diagnostics for building HVAC systems. Applied Energy, 339, 121030. https://doi.org/10.1016/j.apenergy.2023.121030
6. Kutty, N. A. et al. (2024). A systematic review of climate change implications on building energy consumption: Impacts and adaptation measures in hot urban desert climates. Buildings, 14, 13. https://doi.org/10.3390/buildings14010013
7. Alhashmi, M. et al. (2021). Framework for developing a lowcarbon energy demand in residential buildings using communitygovernment partnership: An application in Saudi Arabia. Energies, 14, 4954. https://doi.org/10.3390/en14164954
8. Ali, D. M. T. E.; Motuzienė, V.; DžiugaitėTumėnienė, R. (2024). AIdriven innovations in building energy management systems: A review of potential applications and energy savings. Energies, 17, 4277. https://doi.org/10.3390/en17174277
9. Taheri, S.; Hosseini, P.; Razban, A. (2022). Model predictive control of heating, ventilation, and air conditioning (HVAC) systems: A stateoftheart review. Journal of Building Engineering, 60, 105067. https://doi.org/10.1016/j.jobe.2022.105067
10. Kim, D. et al. (2022). Energy modeling and model predictive control for HVAC in buildings: A review of current research trends. Energies, 15, 7231. https://doi.org/10.3390/en15197231
11. Xin, X. et al. (2024). A comprehensive review of predictive control strategies in heating, ventilation, and airconditioning (HVAC): Modelfree vs model. Journal of Building Engineering, 94, 110013. https://doi.org/10.1016/j.jobe.2024.110013

UNDER PEER REVIEW IN IJAR

12. Deep learning in fault detection and diagnosis of building HVAC systems: A systematic review with meta analysis. (2023). Energy and AI, 12, 100235. https://doi.org/10.1016/j.egyai.2023.100235
13. AI in HVAC fault detection and diagnosis: A systematic review. (2024). Energy Reviews, 3(2), 100071. https://doi.org/10.1016/j.enrev.2024.100071
14. Zhang, L. et al. (2021). Sensor impact evaluation and verification for fault detection and diagnostics in building energy systems: A review. Advances in Applied Energy, 3, 100055. https://doi.org/10.1016/j.adapen.2021.100055
15. Taheri, S.; Ahmadi, A.; MohammadiIvatloo, B.; Asadi, S. (2021). Fault detection diagnostic for HVAC systems via deep learning algorithms. Energy and Buildings, 250, 111275. https://doi.org/10.1016/j.enbuild.2021.111275
16. Wang, S. (2025). Automated fault diagnosis detection of air handling units using real operational labelled data and transformerbased methods at 24hour operation hospital. Building and Environment, 113257. https://doi.org/10.1016/j.buildenv.2025.113257
17. Advanced fault detection, diagnosis and prognosis in HVAC systems: Lifecycle insight, key challenges, and promising approaches. (2025). Renewable and Sustainable Energy Reviews, 115867. https://doi.org/10.1016/j.rser.2025.115867
18. Onweh, C. C.; AlHabaibeh, A.; Manu, E. (2025). A review of energy efficiency strategies in smart buildings: Integrating occupant comfort, HVAC optimisation, and building automation. Research and Reviews in Sustainability, 1(1), 48–60. https://doi.org/10.5334/rss.9
19. Akbulut, L. et al. (2025). A systematic review of building energy management systems (BEMSs): Sensors, IoT, and AI integration. Energies, 18, 6522. https://doi.org/10.3390/en18246522
20. Mehraban, M. H. et al. (2025). AIenhanced automation of building energy optimization using a hybrid stacked model and genetic algorithms. Results in Engineering, 26, 104994.
21. Elwy, I.; Hagishima, A. (2024). The artificial intelligence reformation of sustainable building design approach: A systematic review on building design optimization methods using surrogate models. Energy and Buildings, 323, 114769.
22. Di Stefano, A. G.; Ruta, M.; Masera, G. (2023). Advanced digital tools for datainformed and performancedriven design: A review of building energy consumption forecasting models based on machine learning. Applied Sciences, 13, 12981.
23. Aloshan, M. A. (2025). Achieving nearzero energy in hot climates: Retrofitting building envelopes for existing homes. Heliyon, 11, e42170. https://doi.org/10.1016/j.heliyon.2025.e42170
24. Forouzandeh, N.; Tahsildoost, M.; Zomorodian, Z. S. (2021). A review of webbased building energy analysis applications. Journal of Cleaner Production, 306, 127251.
25. Mihalakakou, G. et al. (2025). Coupling simulationbased and machine learning methodologies for energy optimization and environmental impact mitigation in buildings. Journal of Building Engineering, 112, 113809.

UNDER PEER REVIEW IN IJAR

26. Vamvakas, D. et al. (2025). Generative AI for sustainable smart environments: A review of energy systems, buildings, and usercentric decisionmaking. Energies, 18, 6163.
27. Liu, H.; Du, Z.; Xue, T.; Jiang, T. (2025). An integrated artificial intelligencedriven approach to multicriteria optimization of building energy efficiency and occupants' comfort: A case study. Journal of Building Engineering, 103, 111944.
28. Farzaneh, H. et al. (2021). Artificial intelligence evolution in smart buildings for energy efficiency. Applied Sciences, 11, 763.
29. Alshibani, A. (2020). Prediction of the energy consumption of school buildings. Applied Sciences, 10, 5885. https://doi.org/10.3390/app10175885
30. Youssef, H. A. H. et al. (2025). A digital twin approach for AIcontrolled smart façades: Empirical validation and strategic implementation in achieving Saudi Vision 2030. International Journal of Recent Advances in Multidisciplinary Research, 12(10), 11864–11873. https://doi.org/10.64485/ijramr.6274.10.2025

UNDER PEER REVIEW IN IJAP