World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

COST-AWARE FAULT DETECTION AND DIAGNOSIS FOR HVAC CHILLERS: A REVIEW OF METHODS, FAULT IMPACT QUANTIFICATION, AND REPAIR PRIORITIZATION

Adeola Ibrahim Taiwo *

Department of Industrial Engineering, Morgan State University, Baltimore, Maryland, United States.
* Corresponding Author

ORCID Details
Adeola Ibrahim Taiwo: https://orcid.org/0009-0003-4674-6203

World Journal of Advanced Research and Reviews, 2026, 32(02), 1440–1451

Publication history: Received on 13 October 2026; revised on 19 November 2026; accepted on 21 November 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.32.2.2985

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Chillers are among the largest electricity consumers in commercial buildings, and faults such as refrigerant leakage, condenser fouling and sensor drift can raise their energy use for months before they are noticed. Over two decades, fault detection and diagnosis (FDD) research has produced physics-based, statistical and machine-learning methods that classify chiller faults with high accuracy, largely using the ASHRAE RP-1043 laboratory dataset. This review examines that body of work from the viewpoint of the facility manager, who must decide which detected fault to repair first under limited maintenance budgets. We first summarize chiller fault taxonomies and the reported effects of each fault on energy use, capacity and equipment life. We then compare FDD methods by data needs, accuracy, interpretability and transferability across chillers. Next, we review how fault impacts have been quantified in energy and monetary terms, and how maintenance prioritization methods, such as failure mode and effects analysis and cost-sensitive learning, have been applied to HVAC systems. The review finds that most FDD studies report detection and diagnosis accuracy but rarely translate a diagnosed fault into its energy penalty, repair cost and urgency. Fault severity is seldom estimated, misclassification costs are usually treated as equal, and methods trained on one chiller often fail on another. To close these gaps, we propose a cost-aware FDD framework that links fault diagnosis, severity estimation, energy and cost impact modeling, and repair prioritization. We conclude with a research agenda covering labeled field datasets, severity-aware models, transfer learning and decision-support tools for building operators.

Keywords: Chiller; Fault detection and diagnosis; Machine learning; Fault impact; Maintenance prioritization; Cost-sensitive learning; HVAC

1 INTRODUCTION

Faults in heating, ventilation and air-conditioning (HVAC) equipment waste a large share of commercial building energy. An early U.S. assessment estimated that building faults waste about 205 TWh (0.7 quads) of energy each year, worth nearly US$14 billion [1]. More recent estimates place the added primary energy use between 90 and 530 TWh per year [2], and control-related problems are thought to waste 5–30% of commercial building energy [3]. Chillers deserve

particular attention because they account for a significant part of building energy use in large commercial facilities [4], and a degraded chiller can keep cooling the building while quietly consuming more power.

Automated fault detection and diagnosis (FDD) aims to find such problems early. Research on FDD for building systems has grown steadily for over two decades and has been summarized in several broad reviews [5–13]. Field evidence now shows that the technology pays off: a U.S. Department of Energy campaign covering thousands of buildings documented median annual energy savings of 9% for organizations using FDD software [14], and further work has examined how such savings are achieved and how FDD algorithms should be evaluated [3].

Despite this progress, a gap remains between what FDD algorithms produce and what building operators need. Most published methods output a fault label, such as "condenser fouling" or "refrigerant leak", and are judged by classification accuracy. An operator with a limited maintenance budget, however, must decide which detected fault to fix first. That decision depends on how severe the fault is, how much energy and money it wastes per day, what the repair costs, and whether the fault threatens equipment or comfort. Kim and Katipamula noted a lack of studies that estimate fault impact [7], and Li and O'Neill argued that ranking faults by their impact could help facility managers make decisions [15]. Notably, one of the earliest chiller fault surveys reported faults by both frequency and relative cost to repair [16], but this cost perspective largely disappeared as research shifted toward machine-learning classifiers.

This review examines chiller FDD from the decision-maker's point of view. Unlike earlier reviews that cover all HVAC systems and focus mainly on algorithms [5–13], it concentrates on chillers and follows the full chain from detection to repair decision. Its contributions are fourfold:

- A synthesis of chiller FDD research published between 2000 and 2026, organized by method family, data needs and practical readiness.
- A review of how fault impacts on energy, capacity and cost have been quantified, and how often FDD studies report them.
- An analysis of maintenance prioritization and cost-sensitive methods that could turn a diagnosis into a ranked repair decision.
- A proposed cost-aware FDD framework and a research agenda for closing the identified gaps.

The rest of the paper is organized as follows. Section 2 describes the review method. Section 3 summarizes common chiller faults and their effects. Section 4 compares FDD methods, and Section 5 reviews the datasets used to develop them. Section 6 covers fault impact quantification, and Section 7 reviews prioritization and decision-support approaches. Section 8 presents research gaps and the proposed framework, and Section 9 concludes.

Highlights

- Reviews chiller fault detection and diagnosis (FDD) research from 2000 to 2026.
- Maps fault types to their reported energy, comfort and cost impacts.
- Finds most FDD studies stop at classification accuracy, not repair decisions.
- Few studies rank detected faults by energy penalty and repair cost.
- Proposes a cost-aware framework linking detection, impact and prioritization.

2 REVIEW METHODOLOGY

This review follows a structured search and screening process reported in line with the PRISMA 2020 guideline [17]. The search covered publications from January 2000 to [month] 2026 in Scopus, Web of Science, ScienceDirect, IEEE Xplore and the ASHRAE Technical Library, supplemented by Purdue e-Pubs for the International Refrigeration and Air Conditioning Conference. National laboratory reports from LBNL, NREL, PNNL and ORNL were included because much of the field evidence on fault prevalence and savings appears there.

2.1 Search strings

Three groups of terms were combined with AND:

- *Equipment:* chiller OR "chilled water" OR "vapor compression" OR "central plant"
- *Task:* "fault detection" OR "fault diagnosis" OR FDD OR AFDD OR "anomaly detection" OR prognostics
- *Decision (second pass):* "fault impact" OR "energy penalty" OR cost OR prioritization OR "maintenance scheduling" OR FMEA OR "cost-sensitive"

2.2 Inclusion criteria

Studies were included if they (i) were peer-reviewed journal or conference papers, or technical reports from recognized research institutions; (ii) addressed faults in chillers or chiller plants, or offered fault-impact or prioritization methods transferable to chillers; and (iii) were written in English. Studies on air-handling units or rooftop units were kept only when they contributed a method for impact estimation, cost-aware decisions or datasets relevant to chillers.

2.3 Screening and coding

Titles and abstracts were screened first, followed by full-text review. Each included study was coded for: fault types covered; method family; data source (laboratory, simulation or field); evaluation metrics; whether fault severity was estimated; whether energy or cost impact was quantified; and whether a repair priority or decision was produced.

2.4 Search outcome

[To be completed by the authors after running the search: number of records identified (n = …), duplicates removed (n = …), records screened (n = …), full texts assessed (n = …) and studies included (n = …). A PRISMA flow diagram will be provided as Fig. 1.]

3 CHILLER FAULTS AND THEIR EFFECTS

Chiller faults fall into four broad groups: refrigerant-side faults, water-side and heat-exchanger faults, sensor faults, and control faults. The first two groups are "degradation" faults: they develop slowly, rarely trip an alarm, and mainly raise energy use and shorten equipment life [18]. That is why they are both costly and hard to notice without automated FDD.

3.1 Fault survey and benchmark fault set

Comstock, Braun and Groll surveyed major U.S. chiller manufacturers to identify the most common and most expensive chiller faults, reporting them by frequency of occurrence and relative repair cost [16]. Building on that survey, ASHRAE Research Project 1043 (RP-1043) introduced seven faults into a centrifugal chiller in the laboratory at four severity levels each [19]. The same team showed that all seven faults, at the levels tested, measurably changed chiller performance and should be detectable with relatively low-cost sensors [18]. This fault set has since become the de facto benchmark for chiller FDD research (Section 5).

Table 1: Common chiller faults, typical symptoms and main consequences (fault set based on [16,18,19]).

Group	Fault	Typical symptoms	Main consequence
Refrigerant side	Refrigerant leak (undercharge)	Lower evaporating pressure, higher suction superheat	Reduced capacity, higher energy per ton
Refrigerant side	Refrigerant overcharge	Higher condensing pressure, higher subcooling	Higher compressor power
Refrigerant side	Non-condensable gas in refrigerant	Elevated condenser pressure relative to water temperature	Higher lift and compressor power
Refrigerant side	Excess oil	Degraded heat transfer, abnormal oil temperatures	Lower efficiency, compressor wear
Water side	Condenser fouling	Larger condenser approach temperature	Higher lift and compressor power
Water side	Reduced condenser water flow	Larger condenser water temperature rise	Higher lift, possible high-pressure trips
Water side	Reduced evaporator water flow	Larger chilled-water temperature drop, low evaporating temperature	Freeze risk, reduced capacity
Sensor	Bias or drift in temperature, pressure or flow sensors	Inconsistent readings between related sensors	Wrong control actions, false FDD alarms
Control	Poor setpoints or staging	Chillers running at inefficient part load	Avoidable energy use

3.2 Sensor and control faults

Sensor faults deserve separate treatment because they corrupt both control and FDD. Wang and Cui showed that principal component analysis (PCA) can detect, isolate and estimate biased sensors in centrifugal chiller systems before system-level diagnosis is run [23]. Simulation studies indicate that the effect of incipient sensor errors can be large: in one multi-zone commercial building model, cooling energy ranged from 11.5% lower to 65% higher than the fault-free baseline depending on the sensor error [24].

3.3 Fault prevalence

How often each fault occurs in real chillers is still poorly documented. Recent large-scale U.S. studies of fault prevalence have analyzed data from commercial FDD software and service records [21,22], and a unified HVAC fault taxonomy has been developed to make such data comparable across tools [20]. However, this work has focused mainly on air-handling units, terminal units and rooftop units [20,21]. Field data on the frequency, severity and duration of chiller faults remain scarce, which limits any attempt to estimate expected annual fault costs.

4 CHILLER FDD METHODS

Chiller FDD methods have moved from physical models toward data-driven learning, and accuracy on laboratory data is now very high. Across building systems, process-history (data-driven) methods are the most widely used, while quantitative model-based methods are the least used because of their modeling effort [7]. For large HVAC systems, data-driven approaches are generally judged more promising than model-based or knowledge-based ones [9]. Table 2 compares the main method families; the subsections below summarize each.

Table 2: Main families of chiller FDD methods.

Method family	Representative chiller studies	Training data needed	Main strengths	Main limitations
Physics-based and gray-box models	[25,26]	Fault-free data plus design or manufacturer data	Physically interpretable; little or no fault data needed	Modeling effort; calibration for each chiller
Statistical process monitoring (PCA, control charts)	[23,27,28]	Fault-free data	Simple; good for sensor faults and detection	Weak at isolating root causes
One-class and novelty detection (SVDD)	[29]	Fault-free data	Suited to scarce fault data	Detects but does not diagnose on its own
Probabilistic and knowledge-based (Bayesian networks)	[30]	Expert knowledge plus data	Handles uncertain, incomplete evidence	Network design requires expertise
Supervised machine learning (SVM, LDA, ensembles)	[4,31,32]	Labeled fault and fault-free data	High accuracy; can estimate severity level	Needs labeled faults; may not transfer across chillers
Deep learning (CNN, RNN, hybrids)	[33–35]	Large labeled datasets	Learns features automatically; very high accuracy	Black-box; data-hungry
Transfer learning and domain adaptation	[37–40]	Source data plus limited target data	Reuses knowledge across chillers or conditions	Young field; few field validations

4.1 Physics-based and statistical methods

Early chiller FDD used mechanistic models. McIntosh et al. calibrated a detailed model of a 5,500-ton centrifugal chiller with site data and used it to identify operating faults [25]. Reddy later applied a generic assessment methodology to compare four chiller FDD methods under ASHRAE RP-1275 [26]. Statistical approaches followed: Wang and Cui used PCA for sensor faults [23], Zhao et al. combined support vector regression reference models with exponentially-weighted moving average (EWMA) control charts [27], and Beghi et al. used a semi-supervised PCA approach with a reconstruction-based contribution method to isolate faulty variables, tested on two water chillers [28].

4.2 Machine-learning classifiers

Because fault-free data are plentiful but fault data are scarce, Zhao et al. treated detection as a one-class problem and showed that support vector data description (SVDD) outperformed PCA-based detection [29]. The same group developed a diagnostic Bayesian network that mimics how chiller experts reason under uncertain evidence [30]. Han et al. used feature selection with multi-class SVMs and found that eight sensors centered on the core refrigeration cycle outperformed larger feature sets drawn from 64 candidates [31]. Li et al. used linear discriminant analysis to detect faults and assign them to a severity level [4], one of the few studies to estimate severity explicitly. Kim and Kim compared eight classifiers on RP-1043 data and found that the amount of training data and the fault severity strongly affected accuracy [32].

4.3 Deep learning and explainability

Deep networks now dominate recent work. Examples include a hybrid one-dimensional convolutional neural network and gated recurrent unit for time-series chiller data [33] and a deep CNN trained with self-produced data [34]. Li et al. proposed an explainable 1D-CNN that reached over 98.5% diagnosis accuracy for the seven RP-1043 faults while showing which inputs drove each decision [35]. Explainability matters for adoption: a study of a Singapore building with six chillers argued that operators need explanations before they will trust data-driven FDD alarms [36].

4.4 Transfer learning and field readiness

A model trained on one chiller often performs poorly on another. Transfer learning addresses this. Liu et al. tested four transfer strategies on two water-cooled screw chillers and reported maximum accuracy gains of 12.63% and 8.18% in two learning tasks when target data were limited [37]. Later studies transferred knowledge between a 703-kW screw chiller and the 316-kW RP-1043 chiller [38], across operating conditions [39], and under severe data scarcity using synthetic samples [40].

Field readiness remains a concern. Research models often rely on sensors that commercial chillers do not have. Studies using only field or factory-installed sensors found that diagnosis is feasible but more difficult, especially with imbalanced data [41,42], and a recent study built its feature libraries from sensors actually installed on field chillers [43].

4.5 What the methods report

Nearly all studies above report detection rates, diagnosis accuracy or F1 scores. Few report fault severity, and almost none translate a diagnosis into energy waste, cost or repair urgency. The output of a typical chiller FDD model is therefore a label, not a decision.

5 BENCHMARK DATASETS

Most chiller FDD studies still train and test on a single laboratory dataset, ASHRAE RP-1043, so reported accuracies say more about that chiller than about chillers in general. Table 3 summarizes the main data sources.

Table 3: Data sources used in chiller FDD research.

Data source	Equipment	Data type	Fault content	Typical use
ASHRAE RP-1043 [19]	One 316-kW (90-ton) centrifugal chiller	Laboratory, steady-state and transient	Seven faults at four severity levels plus fault-free tests	Training and benchmarking most published methods
Multi-laboratory labeled HVAC dataset [44]	Common commercial HVAC systems	Simulation, laboratory and field	Many fault types and severities, mostly a year of data per case	Benchmarking across systems and climates
Research-group laboratory chillers [37,38]	Water-cooled screw chillers	Laboratory	Selected faults	Testing transfer learning between chillers
Field building automation data [36,41,42]	Installed chillers	Field operation	Few labeled faults; imbalanced	Testing practical feasibility

5.1 ASHRAE RP-1043

RP-1043 produced fault-free and faulty data for a centrifugal chiller with seven faults introduced at four severity levels [19]. Its open availability made it the common yardstick, and it is used by most studies cited in Section 4 [4,27,29,31,32,35]. Its strengths are careful instrumentation and controlled fault levels. Its weaknesses are just as important: it covers one chiller of one type, faults are introduced one at a time, operating conditions follow a test matrix rather than real building loads, and it records no repair costs or energy prices.

5.2 Newer open datasets

Granderson et al. published what they describe as the largest known public dataset of building system operation in faulted and fault-free states, drawing on simulation, experimental facilities and field operation across climates, fault types and severities [44]. [Authors: confirm which chiller-plant cases are included in the current release before submission.] Such datasets make it possible to test methods beyond one laboratory chiller.

5.3 Evaluation practice

Frank et al. observed that far more effort has gone into developing FDD algorithms than into evaluating them, and proposed a general framework that defines what counts as a fault and as an evaluation sample [45]. Lin et al. likewise reviewed methods for evaluating algorithm performance alongside achieved field savings [3]. Neither the RP-1043 literature nor these frameworks yet include cost-weighted metrics, for example the expected cost of missing a severe refrigerant leak compared with a false alarm. Very high accuracies reported on RP-1043, often above 98% [35], should therefore be read as an upper bound rather than as expected field performance.

6 QUANTIFYING FAULT IMPACT

A fault's impact can be estimated, but the methods exist mostly outside the chiller FDD literature and are rarely connected to it. To rank faults, an operator needs at least the extra energy a fault causes and the cost of that energy over time. For a fault *f* present over *T* time steps of length Δt, the energy penalty and its cost can be written as:

$$\Delta E_f = \sum_{t=1}^{T} \left[P_f(t) - P_0(t) \right] \Delta t$$

$$C_f^{\mathrm{waste}} = \sum_{t=1}^{T} \left[P_f(t) - P_0(t) \right] p_e(t) \, \Delta t$$

where $P_f(t)$ is the measured chiller power with the fault, $P_0(t)$ is the expected fault-free power under the same load and weather, and $p_e(t)$ is the electricity price, which may include time-of-use rates. Demand charges, comfort penalties and accelerated wear would add further terms. The challenge lies in estimating $P_0(t)$ and in predicting how the penalty will grow if repair is delayed.

6.1 Laboratory and simulation approaches

The RP-1043 tests measured how chiller performance changes with each fault and severity level [18,19], which gives direct but equipment-specific impact data. Simulation offers broader coverage. Li and O'Neill reviewed fault models for HVAC systems and found that the number of validated fault models is still limited and that simulation platforms need improvement for impact analysis [15]. Kim et al. built a library of fault models in EnergyPlus for small commercial buildings [2], and Cheung and Braun developed gray-box inverse models that simulate fault impacts in vapor compression equipment, such as low refrigerant charge and heat-exchanger fouling, from limited laboratory data [46]. Comparable impact studies for rooftop units date back to Breuker and Braun [47].

6.2 Data-driven baselines

Many chiller FDD methods already build a fault-free reference model, for example a support vector regression model of performance indices [27]. The same reference model can estimate $P_0(t)$, so the gap between measured and expected power gives a running estimate of the energy penalty. This connection between FDD reference models and impact estimation is natural, yet it is rarely made explicit in the reviewed studies.

6.3 Economic impact

Few studies express fault impact in money. Verhelst et al. evaluated the economic impact of persistent sensor and actuator faults in office buildings with concrete core activation, combining energy and comfort costs [48]. At the national level, Roth et al. estimated the annual cost of building faults in the United States [1]. Li and O'Neill later combined failure mode and effects analysis with whole-building simulation to rank the criticality of faults [49]. For chillers specifically,

however, none of the studies reviewed here links a diagnosed fault and its estimated severity to a dollar-per-day cost of leaving it unrepaired.

7 FROM DETECTION TO DECISION

Tools for turning faults into ranked maintenance actions exist in reliability engineering and machine learning, but they have seldom been applied to chiller FDD output. Table 4 summarizes the approaches found and what each still lacks for chillers.

Table 4: Approaches for prioritizing detected faults.

Approach	Output	HVAC examples	Gap for chiller FDD
Manufacturer fault surveys	Fault frequency and relative repair cost	[16]	Dated; not linked to live diagnosis
Failure mode and effects analysis (FMEA)	Ordinal criticality ranking	[49,50]	Ordinal scores, not money; little chiller-specific work
Prognostics and time-to-failure models	Estimated time before failure	[51]	Few chiller datasets with recorded failures
Cost-sensitive learning	Classifier decisions that minimize expected misclassification cost	[52,53]	Applied mainly to feature selection, not repair ranking
Simulation-based impact ranking	Faults ranked by energy or comfort impact	[15,49]	Not driven by real-time diagnosis

7.1 Maintenance strategies and prognostics

HVAC maintenance is still mostly corrective (fix after failure) or preventive (fix on a schedule). Predictive maintenance uses FDD and prognostics to act on equipment condition instead. Yang et al. developed regression-tree models that predict time to failure for central heating and cooling plant equipment from control-system data, and found that about half of the modeled failure events could be predicted accurately [51]. Such prognostics tell an operator *when* a failure may occur but not *what it will cost* to wait.

7.2 Failure mode and effects analysis

FMEA ranks failure modes by their severity, occurrence and detectability. Yang et al. built an HVAC FMEA by mining work orders from building energy management systems, as a practical guide for data gathering and prognostics [50]. Li and O'Neill combined FMEA with whole-building energy simulation and used sensitivity analysis to rank fault criticality [49]. Both studies show that ranking is feasible, but FMEA scores are ordinal and do not express impact in energy or money.

7.3 Cost-sensitive learning

In cost-sensitive classification, the decision is the class with the lowest expected cost rather than the highest probability [52]. With $P(j \mid x)$ the probability that the true fault is j given data x, and $C(i, j)$ the cost of acting as if fault i were present when fault j is true, the decision rule is:

$$\hat{y}(x) = \arg\min_{i} \sum_{j} P(j \mid x) \, C(i, j)$$

Yan et al. brought this idea to chiller FDD, noting that false positives and false negatives carry different costs, and used it to select features [53]. That study is an important step, but it applied cost only to the choice of sensors. The cost matrix $C(i, j)$ itself, which should reflect energy waste, repair cost and risk for each chiller fault, has not been estimated from data in the studies reviewed.

7.4 Summary

The building blocks of a cost-aware decision exist separately: diagnosis (Section 4), severity estimation [4], impact models (Section 6), criticality ranking [49,50] and cost-sensitive decision rules [52,53]. What is missing is a framework that connects them, so that each diagnosed chiller fault arrives with an estimated cost of delay and a recommended priority.

8 RESEARCH GAPS AND A COST-AWARE FDD FRAMEWORK

The review points to one central gap: chiller FDD tells operators *what* is wrong but not *what it costs* or *what to fix first*. Seven specific gaps follow from Sections 3–7.

8.1 Research gaps

- **Accuracy without decisions.** Studies report detection and diagnosis accuracy but almost never a repair recommendation (Section 4.5).
- **Severity rarely estimated.** Few methods estimate how severe a diagnosed fault is [4], although impact depends strongly on severity [18].
- **Impact not linked to diagnosis.** Energy and cost impact methods exist [2,15,46,48] but are not connected to live FDD output, and none of the chiller studies reviewed reports a cost per day of delay.
- **Equal error costs.** Most classifiers treat all misclassifications alike. Cost-sensitive methods exist [52,53], but cost matrices for chiller faults have not been estimated from energy, repair and risk data.
- **Dependence on one benchmark.** Most results come from RP-1043 [19], and models often degrade on other chillers, operating conditions or field sensor sets [37–43].
- **Missing field data on prevalence and repair cost.** Chiller fault frequency, duration and repair cost are poorly documented compared with air-side equipment [16,20–22].
- **Trust and explainability.** Operators need to understand why a model raises an alarm before acting on it [35,36].

8.2 Proposed framework

To close these gaps, we propose a cost-aware FDD framework with six linked stages (Fig. 2).

- **Detect.** A fault-free reference model of chiller performance flags deviations. The same model supplies the expected fault-free power $P_0(t)$ used in Section 6.
- **Diagnose.** A classifier identifies the most likely fault and reports its probability $P(j \mid x)$, with an explanation of the inputs that drove the decision. Transfer learning adapts the classifier to a new chiller using limited local data.
- **Estimate severity.** A severity model places the fault on a continuous or ordinal scale, using fault-indicative features.
- **Quantify impact.** The energy-penalty and cost equations in Section 6 convert the fault into an energy penalty and a cost per day, using local tariffs, and project how that cost grows if repair is delayed.
- **Add repair cost and risk.** Service records, manufacturer data or expert surveys supply repair cost, and FMEA-style factors capture risks such as freeze-up or compressor damage.
- **Prioritize.** An optimization ranks repairs under budget and crew limits, and repair outcomes feed back to update the models.

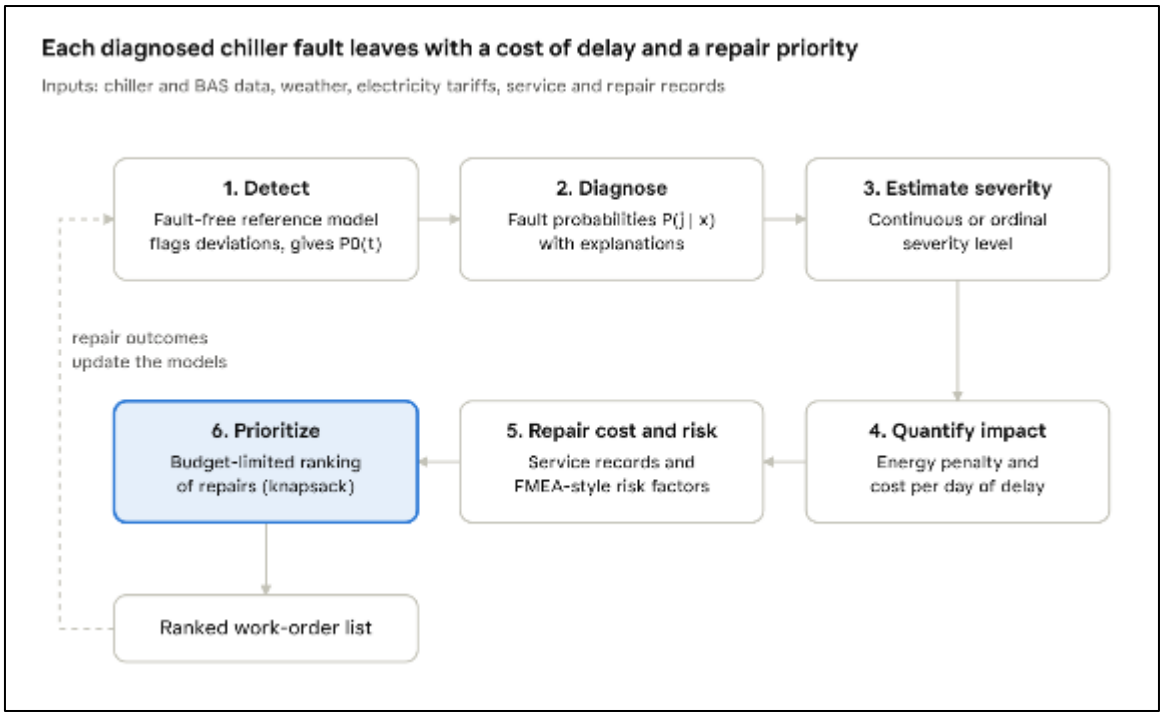


Figure 2: Proposed cost-aware chiller FDD framework · 6 stages and a feedback loop

Fig. 2. Proposed cost-aware chiller FDD framework. Stages 1–3 extend existing FDD; stages 4–6 add the impact, cost and prioritization steps that current studies lack. Repair outcomes feed back to update the models.

Stage 6 can be framed as a knapsack problem. Let $x_f = 1$ if fault *f* is repaired in the current maintenance window, c_f be its expected daily cost, τ_f the days of waste avoided by repairing now rather than later, R_f the repair cost and B the available budget:

$$\max_{x} \sum_{f} x_f \left(c_f \, \tau_f - R_f \right) \quad \text{s.t.} \quad \sum_{f} x_f R_f \le B, \quad x_f \in \{0, 1\}$$

Because diagnosis is uncertain, the expected daily cost combines diagnosis probabilities with the daily cost c_j of each candidate fault j:

$$c_f = \sum_{j} P(j \mid x) \, c_j$$

This formulation links every stage: better diagnosis sharpens $P(j \mid x)$, better severity and impact models sharpen c_j, and the optimization turns both into a ranked work-order list.

8.3 Research agenda

- **Field datasets with costs.** Labeled field data from several chiller types, including recorded repair dates, actions and costs.
- **Severity-aware models.** Methods that estimate severity on a continuous scale and are validated against measured energy penalties.
- **Data-based cost matrices.** Estimation of misclassification costs for chiller faults from energy, repair and risk data, and their use in training and evaluation.
- **Cost-weighted evaluation metrics.** Benchmarks that score FDD tools by avoided cost as well as accuracy, extending current evaluation frameworks [45].
- **Field-validated transfer learning.** Transfer methods tested on installed chillers with only factory sensors [41–43].
- **Low-GWP refrigerant transition.** Fault signatures for chillers using newer low-GWP refrigerants, since existing benchmark data come mostly from chillers with legacy refrigerants.
- **Decision-support tools.** Operator-facing tools that present cost of delay, priority and explanation together, tested with facility staff.

9 CONCLUSIONS

Chiller FDD research has matured into accurate, increasingly data-driven methods, but it has not yet answered the question operators face every day: which fault should be fixed first. This review of chiller FDD studies from 2000 to 2026 leads to four conclusions.

- **Diagnosis accuracy is no longer the main bottleneck.** On the RP-1043 benchmark, modern classifiers diagnose the seven standard faults with accuracies often above 98%. Field performance, transfer to other chillers and operator trust are now the harder problems.
- **Fault impact is rarely quantified alongside diagnosis.** Methods to estimate energy penalties exist, from laboratory sensitivity tests to gray-box and building simulation models, but they are seldom attached to FDD output, and chiller faults are almost never expressed as a cost per day of delay.
- **Decision tools exist but sit apart from FDD.** FMEA, prognostics and cost-sensitive learning offer ways to rank faults and weigh errors, yet in chiller research they have been applied separately and mostly to sensor selection rather than repair decisions.
- **A cost-aware framework can connect the pieces.** Linking detection, diagnosis, severity estimation, impact quantification, repair cost and budget-constrained optimization would turn chiller FDD from an alarm system into a decision-support tool.

Progress will depend on field datasets that record repairs and their costs, on severity-aware and transferable models, and on evaluation metrics that reward avoided cost as well as accuracy. Work in these directions would help building operators act on FDD results with confidence and capture more of the energy savings that FDD makes possible.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

I would like to express my appreciation to my colleagues in the Department of Industrial Engineering at Morgan State University, John Onofowokan and Wisdom Ukeije, for their time, critical peer review, and helpful suggestions throughout the writing of this article. I also acknowledge the foundational studies and reference materials that guided and enriched this work.

REFERENCES

[1] K.W. Roth, D. Westphalen, M.Y. Feng, P. Llana, L. Quartararo, Energy Impact of Commercial Building Controls and Performance Diagnostics: Market Characterization, Energy Impact of Building Faults and Energy Savings Potential, TIAX LLC report for the U.S. Department of Energy, 2005, NTIS PB2006-100567.

[2] J. Kim, S. Frank, J.E. Braun, D. Goldwasser, Representing small commercial building faults in EnergyPlus, Part I: Model development, Buildings 9 (11) (2019) 233. https://doi.org/10.3390/buildings9110233

[3] G. Lin, H. Kramer, J. Granderson, Building fault detection and diagnostics: Achieved savings, and methods to evaluate algorithm performance, Build. Environ. 168 (2020) 106505. https://doi.org/10.1016/j.buildenv.2019.106505

[4] D. Li, G. Hu, C.J. Spanos, A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis, Energy Build. 128 (2016) 519–529. https://doi.org/10.1016/j.enbuild.2016.07.014

[5] S. Katipamula, M.R. Brambley, Methods for fault detection, diagnostics, and prognostics for building systems—A review, Part I, HVAC&R Res. 11 (1) (2005) 3–25.

[6] S. Katipamula, M.R. Brambley, Methods for fault detection, diagnostics, and prognostics for building systems—A review, Part II, HVAC&R Res. 11 (2) (2005) 169–187. https://doi.org/10.1080/10789669.2005.10391133

[7] W. Kim, S. Katipamula, A review of fault detection and diagnostics methods for building systems, Sci. Technol. Built Environ. 24 (1) (2018) 3–21. https://doi.org/10.1080/23744731.2017.1318008

[8] Y. Zhao, T. Li, X. Zhang, C. Zhang, Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future, Renew. Sustain. Energy Rev. 109 (2019) 85–101.

[9] M.S. Mirnaghi, F. Haghighat, Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review, Energy Build. 229 (2020) 110492. https://doi.org/10.1016/j.enbuild.2020.110492

[10] J. Chen, L. Zhang, Y. Li, Y. Shi, et al., A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems, Renew. Sustain. Energy Rev. 161 (2022) 112395. † (full author list)

[11] Z. Chen, Z. O'Neill, J. Wen, O. Pradhan, T. Yang, X. Lu, et al., A review of data-driven fault detection and diagnostics for building HVAC systems, Appl. Energy 339 (2023) 121030. https://doi.org/10.1016/j.apenergy.2023.121030

[12] S.P. Melgaard, K.H. Andersen, A. Marszal-Pomianowska, R.L. Jensen, P.K. Heiselberg, Fault detection and diagnosis encyclopedia for building systems: A systematic review, Energies 15 (12) (2022) 4366. https://doi.org/10.3390/en15124366

[13] J. Bi, H. Wang, E. Yan, C. Wang, K. Yan, L. Jiang, B. Yang, AI in HVAC fault detection and diagnosis: A systematic review, Energy Rev. 3 (2) (2024) 100071.

[14] H. Kramer, G. Lin, C. Curtin, E. Crowe, J. Granderson, Proving the Business Case for Building Analytics, Lawrence Berkeley National Laboratory, 2020. https://doi.org/10.20357/B7G022

[15] Y. Li, Z. O'Neill, A critical review of fault modeling of HVAC systems in buildings, Build. Simul. 11 (5) (2018) 953–975. https://doi.org/10.1007/s12273-018-0458-4

[16] M.C. Comstock, J.E. Braun, E.A. Groll, A survey of common faults for chillers, ASHRAE Trans. 108 (1) (2002). † (pages)

[17] M.J. Page, J.E. McKenzie, P.M. Bossuyt, et al., The PRISMA 2020 statement: An updated guideline for reporting systematic reviews, BMJ 372 (2021) n71. https://doi.org/10.1136/bmj.n71

[18] M.C. Comstock, J.E. Braun, E.A. Groll, The sensitivity of chiller performance to common faults, HVAC&R Res. 7 (3) (2001). https://doi.org/10.1080/10789669.2001.10391274 † (pages)

[19] M.C. Comstock, J.E. Braun, Development of Analysis Tools for the Evaluation of Fault Detection and Diagnostics in Chillers, ASHRAE Research Project 1043-RP, Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN, 1999.

[20] Y. Chen, G. Lin, E. Crowe, J. Granderson, Development of a unified taxonomy for HVAC system faults, Energies 14 (17) (2021) 5581. https://doi.org/10.3390/en14175581

[21] E. Crowe, Y. Chen, H. Reeve, D. Yuill, A. Ebrahimifakhar, Y. Chen, L. Troup, A. Smith, J. Granderson, Empirical analysis of the prevalence of HVAC faults in commercial buildings, Sci. Technol. Built Environ. (2023). https://doi.org/10.1080/23744731.2023.2263324 † (volume and pages)

[22] S. Katipamula, R.M. Underhill, N. Fernandez, W. Kim, R.G. Lutes, D. Taasevigen, Prevalence of typical operational problems and energy savings opportunities in U.S. commercial buildings, Energy Build. 253 (2021) 111544. https://doi.org/10.1016/j.enbuild.2021.111544

[23] S. Wang, J. Cui, Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method, Appl. Energy 82 (3) (2005) 197–213.

[24] Y. Li, P. Im, S. Lee, Y. Bae, Y. Yoon, S. Lee, Sensor incipient fault impacts on building energy performance: A case study on a multi-zone commercial building, Buildings 13 (2) (2023) 520. https://doi.org/10.3390/buildings13020520

[25] I.B.D. McIntosh, J.W. Mitchell, W.A. Beckman, Fault detection and diagnosis in chillers: Model development and application, ASHRAE Trans. 106 (2) (2000). † (part and pages)

[26] T.A. Reddy, Application of a generic methodology to assess four different chiller FDD methods (RP-1275), HVAC&R Res. 13 (5) (2007) 711–729.

[27] Y. Zhao, S. Wang, F. Xiao, A statistical fault detection and diagnosis method for centrifugal chillers based on exponentially-weighted moving average control charts and support vector regression, Appl. Therm. Eng. 51 (2013) 560–572. † (journal and issue)

[28] A. Beghi, R. Brignoli, L. Cecchinato, G. Menegazzo, M. Rampazzo, F. Simmini, Data-driven fault detection and diagnosis for HVAC water chillers, Control Eng. Pract. 53 (2016). † (pages)

[29] Y. Zhao, S. Wang, F. Xiao, Pattern recognition-based chillers fault detection method using support vector data description (SVDD), Appl. Energy 112 (2013) 1041–1048.

[30] Y. Zhao, F. Xiao, S. Wang, An intelligent chiller fault detection and diagnosis methodology using Bayesian belief network, Energy Build. 57 (2013) 278–288. https://doi.org/10.1016/j.enbuild.2012.11.007

[31] H. Han, B. Gu, T. Wang, Z.R. Li, Important sensors for chiller fault detection and diagnosis (FDD) from the perspective of feature selection and machine learning, Int. J. Refrig. 34 (2) (2011) 586–599.

[32] I. Kim, W. Kim, Development and validation of a data-driven fault detection and diagnosis system for chillers using machine learning algorithms, Energies 14 (7) (2021) 1945. https://doi.org/10.3390/en14071945

[33] Z. Wang, Y. Dong, W. Liu, Z. Ma, A novel fault diagnosis approach for chillers based on 1-D convolutional neural network and gated recurrent unit, Sensors 20 (9) (2020) 2458. https://doi.org/10.3390/s20092458

[34] J. Gao, H. Han, Z. Ren, et al., Fault diagnosis for building chillers based on data self-production and deep convolutional neural network, J. Build. Eng. 34 (2021) 102043. † (full author list)

[35] G. Li, Q. Yao, C. Fan, C. Zhou, G. Wu, Z. Zhou, X. Fang, An explainable one-dimensional convolutional neural networks based fault diagnosis method for building heating, ventilation and air conditioning systems, Build. Environ. 203 (2021). † (article number)

[36] S. Srinivasan, P. Arjunan, B. Jin, A.L. Sangiovanni-Vincentelli, Z. Sultan, K. Poolla, Explainable AI for chiller fault-detection systems: Gaining human trust, Computer 54 (10) (2021) 60–68. https://doi.org/10.1109/MC.2021.3071551

[37] J. Liu, Q. Zhang, X. Li, et al., Transfer learning-based strategies for fault diagnosis in building energy systems, Energy Build. 250 (2021) 111256. https://doi.org/10.1016/j.enbuild.2021.111256 † (full author list)

[38] G. Li, L. Chen, J. Liu, X. Fang, Comparative study on deep transfer learning strategies for cross-system and cross-operation-condition building energy systems fault diagnosis, Energy (2023) 125943. https://doi.org/10.1016/j.energy.2022.125943 † (volume)

[39] Z. Chen, W. Zhang, W. Zhao, X. Yang, X. Zhang, Y. Li, Cross-condition fault diagnosis of chillers based on an ensemble approach with adaptive weight allocation, Energy Build. (2024) 115007. https://doi.org/10.1016/j.enbuild.2024.115007 † (volume)

[40] K. Yao, T. Zhao, T. Jin, M. Xiao, X. Zhang, Y. Rezgui, Y. Li, A GAN-enhanced transfer learning framework for cross-condition chiller fault diagnosis under severe data scarcity, Build. Environ. 303 (2026) 114926. https://doi.org/10.1016/j.buildenv.2026.114926

[41] Y.Q. Fan, X.Y. Cui, H. Han, H.L. Lu, Chiller fault diagnosis with field sensors using the technology of imbalanced data, Appl. Therm. Eng. 159 (2019) 113933. https://doi.org/10.1016/j.applthermaleng.2019.113933

[42] Y.Q. Fan, X.Y. Cui, H. Han, H.L. Lu, Feasibility and improvement of fault detection and diagnosis based on factory-installed sensors for chillers, Appl. Therm. Eng. 164 (2020) 114506. https://doi.org/10.1016/j.applthermaleng.2019.114506

[43] Z. Wang, J. Guo, P. Xia, L. Wang, C. Zhang, Q. Leng, K. Zheng, Feature selection for chillers fault diagnosis from the perspectives of machine learning and field application, Energy Build. (2024) 113937. https://doi.org/10.1016/j.enbuild.2024.113937 † (volume)

[44] J. Granderson, G. Lin, Y. Chen, A. Casillas, J. Wen, Z. Chen, P. Im, S. Huang, J. Ling, A labeled dataset for building HVAC systems operating in faulted and fault-free states, Sci. Data 10 (2023) 342. https://doi.org/10.1038/s41597-023-02197-w

[45] S. Frank, G. Lin, X. Jin, R. Singla, A. Farthing, J. Granderson, A performance evaluation framework for building fault detection and diagnosis algorithms, Energy Build. 192 (2019) 84–92. https://doi.org/10.1016/j.enbuild.2019.03.024

[46] H. Cheung, J.E. Braun, Inverse modeling to simulate fault impacts for vapor compression equipment, Part 2: System modeling and validation, in: Proc. 14th Int. Refrigeration and Air Conditioning Conf. at Purdue, West Lafayette, IN, 2012, Paper 2110.

[47] M.S. Breuker, J.E. Braun, Common faults and their impacts for rooftop air conditioners, HVAC&R Res. 4 (3) (1998) 303–318. https://doi.org/10.1080/10789669.1998.10391406

[48] J. Verhelst, G. Van Ham, D. Saelens, L. Helsen, Economic impact of persistent sensor and actuator faults in concrete core activated office buildings, Energy Build. 142 (2017) 111–127. https://doi.org/10.1016/j.enbuild.2017.02.052

[49] Y. Li, Z. O'Neill, An innovative fault impact analysis framework for enhancing building operations, Energy Build. 199 (2019). † (pages)

[50] C. Yang, W. Shen, Q. Chen, B. Gunay, A practical solution for HVAC prognostics: Failure mode and effects analysis in building maintenance, J. Build. Eng. 15 (2018) 26–32. † (pages)

[51] C. Yang, B. Gunay, Z. Shi, W. Shen, Machine learning-based prognostics for central heating and cooling plant equipment health monitoring, IEEE Trans. Autom. Sci. Eng. 18 (1) (2021) 346–355.

[52] C. Elkan, The foundations of cost-sensitive learning, in: Proc. 17th Int. Joint Conf. on Artificial Intelligence (IJCAI), Seattle, WA, 2001, pp. 973–978.

[53] K. Yan, L. Ma, Y. Dai, W. Shen, Z. Ji, D. Xie, Cost-sensitive and sequential feature selection for chiller fault detection and diagnosis, Int. J. Refrig. 86 (2018). https://doi.org/10.1016/j.ijrefrig.2017.11.003 † (authors and pages)