STM JOURNALS

Journal of
Refrigeration, Air Conditioning, Heating and Ventilation

ISSN: 2394-1952
Volume 13, Issue 1, 2026
January–April
DOI (Journal): 10.37591/JoRACHV

https://journals.stmjournals.com/jorachv

Review JoRACHV

A Review on AI and Machine Learning for Predictive Maintenance and FDD in RAC Systems

S. Ramabalan[1,*], S. Mahalakshmi[2], V. Sathiya[3], N. Godwin Raja Ebenezer[4]

Abstract

The paper reviews the existing AI/ML methods first in the general context of predictive maintenance and FDD of RAC systems, then specifically focusing on granular cooling appliances. Perspectives and insights are provided on the reasons why potentially valuable models do not make it into practice more often, and where future research and development should be headed. New emerging topics for decision support systems to include domain knowledge and physics-based modeling to enhance data-driven methods are also highlighted. Finally, the paper ends by making a reasoned comparison among the different generations of methods, introducing the current state of the art regarding other standard comparative metrics, and discussing the most promising solutions. This paper aims to inspire collaborative research in intelligent refrigeration and air conditioning (RAC) maintenance systems by reviewing current literature on state-of-the-art technologies, key conversations, and emerging trends. In addition, the review systematically categorizes existing approaches into data-driven, knowledge-driven, and hybrid frameworks, outlining their respective advantages, limitations, and applicability across various operational scenarios. Emphasis is placed on the role of high-quality datasets, sensor deployment strategies, and feature engineering in improving model robustness and generalization capability. The challenges associated with real-time implementation, scalability, cybersecurity, and integration with building management systems are also discussed in detail. Furthermore, the paper examines the impact of explainable AI, transfer learning, federated learning, and digital twin technologies in advancing predictive maintenance and fault detection and diagnosis (FDD) performance. Practical considerations such as computational cost, model interpretability, deployment feasibility, and lifecycle cost-benefit analysis are evaluated to bridge the gap between academic research and industrial adoption. By synthesizing theoretical advancements with practical constraints, the study provides a comprehensive roadmap for researchers, practitioners, and policymakers aiming to develop resilient, energy-efficient, and intelligent RAC systems for next-generation smart buildings and sustainable infrastructure.

Keywords: Predictive Maintenance, Fault Detection and Diagnosis, Artificial Intelligence, Refrigeration Systems, Energy Efficiency, Digital Twin

***Author for Correspondence**
S. Ramabalan
E-mail: cadsrb@gmail.com

[1]Professor, Department of Mechanical Engineering, Selvam College of Technology, Namakkal, Tamil Nadu, India.
[2]Professor, Department of Information Technology, Selvam College of Technology, Namakkal, Tamil Nadu, India.
[3]Professor, Department of Computer Science and Engineering, Selvam College of Technology, Namakkal, Tamil Nadu, India.
[4]Professor, Department of Mechanical Engineering, MAM College of Engineering, Trichy, Tamil Nadu, India.

Received Date: February 06, 2025
Accepted Date: February 07, 2025
Published Date: February 26, 2025

Citation: S. Ramabalan, S. Mahalakshmi, V. Sathiya, N. Godwin Raja Ebenezer. A Review on AI and Machine Learning for Predictive Maintenance and FDD in RAC Systems. Journal of Refrigeration, Air Conditioning, Heating and Ventilation. 2026; 13(1): 15–25p.

INTRODUCTION

Background and Significance

Intelligent systems integrate advanced sensors with data analytics that use machine learning and artificial intelligence to support predictive maintenance and optimize control strategies. This review organizes previous work along three facets: monitoring, diagnosis, and maintenance management; before presenting well-structured insights into the state of existing commercial technologies and unresolved research issues. This is expected to foster interdisciplinary thinking and create new research opportunities.

Maintenance based on the real-time health index provided by machine learning models normally informs a technician when to take an action. AI goes a step further by directly implementing an automatic action in the cyber-physical control system. AI-optimized controllers can set the best configuration for each specific operating condition. This can even lead to wear reduction and preventive action, like for instance, proactive predictive maintenance when slowly wearing but non-critically tolerated parts can be replaced during the next scheduled maintenance [3, 4]. The economic implications are quite substantial - predictive maintenance costs 25-30% less than its reactive counterpart, downtime is 35-45% shorter, and breakdowns decrease by 70-75% [5].

Evolution of FDD in RAC Systems

The journey to intelligent FDD has seen three primary phases: in the 1980s-2000s, early methods depended on expert knowledge, rule-based systems, and physical models [6]. While effective, these methods had limited adaptability, as they were based on a precisely codified and predetermined knowledge base. As a result, wider-scale implementations were prohibitively expensive[7,8].

The development of IoT/ML solutions for the proactive and predictive maintenance of RAC systems will be a crucial step towards their increased resiliency and efficient operation. This also represents another important path for future works, focusing on effective extensions needed to provide the full fault diagnostic and predictive capabilities of the currently available methodologies. Other future research initiatives include extending the work to span a systemic analysis of the results achieved through integration of AI/ML and Big Data analysis moving from predictive to prescriptive maintenance.

This development can be compared to Industry 4.0, where IoT connects cyber-physical systems which are connected to the cloud leading to big data applications and intelligent learning [9]. For RAC, this drift is vital because of the high energy intensity of the realm and modern system complexity [10]. Especially, VRF systems, transcritical cycles, and heat pumps.

Review Scope and Objectives

The main objectives of this review are as follows; (1) For the context of RAC FDD, categorize and assess AI/ML approaches, (2) Overview their historical progression and transition, (3) Identify performance results and compare bases of performance, (4) Analyze challenges, (5) Identify the latest directions of study, (6) Consolidate suggestions for the research community. We concentrate our discussion on vapor compression systems, but where appropriate, we also reference absorption systems and other novel technologies.

Methodology

This review applies a systematic methodology to identify, assess, and summarize the literature. Primary resources comprise peer-reviewed journal papers, conference papers, technical reports (from ASHRAE, NIST, ORNL), and some patents. Databases searched include Scopus, Web of Science, IEEE Xplore, and ScienceDirect. The search terms consist of different combinations of: __fault detection‘‘, __predictive maintenance‘‘, __refrigeration‘‘, __machine learning‘‘, and the names of specific methods.

FOUNDATIONS OF FDD IN RAC SYSTEMS

Fault Taxonomy and Characteristics

A systematic FDD strategy requires a preliminary formulation of a structured fault list. In RAC systems, faults can be classified based on the corresponding system component, their temporal characteristics, and their effects [11]: *Component-based:* The faults in the compressor (valve leakage, bearing wear), heat exchanger (fouling, blockage), expansion device (stuck valves), refrigerant circuit (undercharge, leakage), and air-side (filter clogging, duct leakage). Temporally, faults could be described as abrupt (sudden failures), incipient (gradual degradation), and intermittent (faults occurring occasionally). Depending on the time needed for the fault to aggravate, the fault can also be classified in terms of fault development as slow or fast. In terms of fault impact, faults could be either hard (leading to complete failure) or soft (performance degradation). The fault severity and detectability vary greatly. For example, refrigerant undercharge should reduce capacity and increase power, while for the compressor valve leakage, efficiency is primarily affected with subtle capacity impact [12].

Fundamental System Modeling

Physical modeling is a theoretical foundation of many FDD approaches, which can be generally categorized into three paradigms: First Principles Models: Based on conservation laws and thermodynamic principles, including lumped parameter models, distributed parameter models, and finite volume/element models [13]. First-principles models are highly interpretable, but the availability of necessary information about the underlying system may limit their use. Gray-box Models: These models mix physical principles with empirical correlations [14]. This often corresponds to the most practical trade-off between interpretability of ―white-box‖ and flexibility of ―black-box‖ models and is predominant in the field of FDD. Data-driven Models: Models are solely based on system data without the underlying physical equations [15]. They are the most flexible but also require the largest amount of training data and often lack physical interpretability.

Sensor Technologies and Data Acquisition

The effectiveness of FDD is mainly related to the choice, the number and the quality of the sensors. Nowadays, temperature (thermocouples, RTDs), pressure, electrical, flow inlets, and additional sensors (vibration accelerometers, acoustic emission sensors) are widely used on various industrial components [16].

Recent available sensors present different advantages such as micro-electromechanical systems (low-cost integrated sensors), ultrasonic flow meters (non-intrusive measurement), fiber Bragg grating sensors (distributed sensing) and computer vision (infrared thermal mapping) [17].

Concerning FDD data acquisition, sampling rates from 1 Hz to 1 kHz are often used. Then, the type of signal conditioning and noise filtering as well as the choice of embedded, fog or cloud storage architectures are crucial. For example, embedded computers can be used to preprocess the data, reducing the amount of information that must be sent to the central cloud [18].

Benchmark Datasets and Metrics

To compare FDD methodologies, standardized evaluation is of utmost importance. Several benchmark datasets are available such as: ASHRAE Project RP-1043/RP-1312 (rooftop units with induced faults) [19], Oak Ridge National Laboratory datasets (chiller and heat pump data) [20], University of California, Berkeley data (residential/commercial systems), and industrial datasets from manufacturers. Performance evaluation is done with multiple metrics:

- *Detection:* fault detection rate, false alarm rate, detection delay, precision, F1-score
- *Diagnosis:* classification accuracy, confusion matrices, specificity
- *Economic:* energy savings, maintenance cost reduction, ROI
- *Computational:* training/inference time, memory requirement

However, the absence of standardized evaluation protocols is still a limitation, with many works using custom metrics.

TRADITIONAL AND CLASSICAL ML APPROACHES

Statistical Methods

Statistical techniques were used in early data-driven FDD, including: Limit Checking: It performs a simple threshold comparison against predefined limits. Although it is computationally trivial, it has comparatively high false alarm rates and lacks diagnostic capability. Trend Analysis: It continuously monitors the slopes of the parameters or deviations from their expected trajectories. It is more sensitive than static thresholds but, since trends corresponding to normal operating condition changes and faults are not different, there is a lack of specificity. Statistical Process Control (SPC): Shewhart, Cumulative Sum (CUSUM), and Exponentially Weighted Moving Average (EWMA) control charts are used to detect significant deviation from statistics of a variable over time. The method improves the sensitivity to gradual changes in the process, but it assumes a stationary process. PCA: This is a mathematical procedure that transforms a number of correlated variables into a smaller number of uncorrelated variables called principal components. It can also be used to monitor the residuals in the space of the principal components for fault detection. Is one of the principal component directions along which the largest variance exists in the data. It is also the direction in which the residuals are most sensitive to multimodally occurring small faults.

Dynamic PCA variants include time-lagged variables, multiscale separations, and nonlinear extensions. As the PCA-based methods are based on statistical independence and correlations, they perform effective sensor fault detection, isolation, and identification. However, the PCA-based methods cannot handle nonlinear behavior monitoring and multiple operating modes.

Classical Supervised Learning

Therefore, classical supervised learning algorithms gained interest from the 2000s on:

- SVM learned optimal hyperplanes which completely separate the classes of faults. In essence, they are equivalent linear classifiers with lower-dimensional mappings of the input space. The nonlinear generalization is made possible with the so-called kernel-trick. SVM performs well even with a limited set of training data, however, its computational complexity is a linear function of the number of samples and cubically depends on the number of features. ANNs have also shown promising results for the detection and diagnosis of various faults. For instance, early multilayer perceptrons done good work for the ANN community. The limits of these approaches are related to overfitting, sensitive dependence on the initial conditions and the noise in the data, as well as to the fact that decision-making models are ―black-box‖ models.
- *Decision trees and ensemble methods:* Decision trees are used to create a hierarchical classification that can be easily interpreted. Random forests provide higher robustness by using a bagging approach to generate multiple trees. Gradient boosting machines such as XGBoost and LightGBM have achieved the best performance in tabular FDD data to date.
- *Bayesian networks:* These are used as a probabilistic graphical model to represent causal relationships. Particularly useful for leveraging expert knowledge and handling uncertainty.

Unsupervised Approaches

The previous methods require labeled fault data. The following methods can be developed using only the sensor data, and fault labels not the FDD phase labels.

- *Clustering algorithms:* Clustering algorithms such as K-means, DBSCAN, and Hierarchical clustering find natural patterns/operational modes as well as anomalies in the data. These methods can discover unknown fault patterns in an unsupervised way, but due to limited information about the classes, the proper selection of parameters and initialization can be difficult to get good results.
- *One-class classification:* One-class classification techniques such as the one-class SVM and support vector data description (SVDD) measure the distance of test points from the origin of the hyperspace created using the support vectors from the training normal examples. These methods use only the training samples from the normal state and aim to differentiate whether a new test point is within the normal state hyperspace. These techniques perform novelty detection well but give limited discrimination between which type of fault is happening.
- *Gaussian mixture models (GMM):* Probabilistic modeling of multiple operational modes as Gaussian distributions. It permits soft clustering and uncertainty quantification.

Feature Engineering

Pre-deep learning, FDD feature engineering was indispensable. Including:

- *Time-domain features:* Statistical moments, peak values, root mean square.
- *Frequency-domain features:* FFT coefficients, spectral centroids.
- *Time-frequency features:* Wavelet coefficients, short-time Fourier transform.
- *Domain-specific features:* COP, efficiency ratios, superheat, subcooling.
- Filter methods for feature selection (correlation), wrapper methods (recursive elimination), embedded methods (LASSO). The optimal feature sets differ depending on the system types and fault classes.

There are several ways to select features, such as filter methods (correlation), wrapper methods (recursive elimination), and embedded methods (LASSO). The most suitable feature sets differ from system to system and fault to fault.

Performance and Limitations

Classical ML approaches have provided the following benchmarks:

- *Simple faults:* 85-95% detection accuracy
- *Complex faults:* 60-80% detection accuracy
- Diagnosis accuracy is typically 10-20% lower than detection

The main limitations are as follows:

1. Heavy reliance on feature engineering
2. Limited capability for temporal modeling
3. Scalability to increasingly complex systems
4. Generalization to new system configurations
5. Multiple operating modes are difficult to handle

However, ML-based classical approaches continue to be relevant and effective for practical use cases when data for training is scarce, when computation needs to be minimized, and for interpreting decision rationales.

DEEP LEARNING ARCHITECTURES

Convolutional Neural Networks (CNNs)

CNNs are made compatible to RAC FDD through data transformation:

- *Image representation methods:* Gramian Angular Field (encodes time series as images), Markov Transition Field (captures state transitions), recurrence plots (visualizes phase space patterns), and raw sensor fusion (multi-channel images).
- *Architectural variants:* 1D-CNNs (temporal convolution), 2D-CNNs (applied to transformed images), multi-scale CNNs (parallel pathways), and depthwise separable CNNs (reduced parameters).

CNN-based approaches show optimal feature learning, especially for vibration and acoustic, where patterns are graphically similar to image textures. Drawbacks are computationally expensive transformation and possible loss of temporal fidelity.

Recurrent Neural Networks (RNNs)

RNNs process sequential data thanks to recurrent connections:

- *Long short-term memory (LSTM):* Solves the vanishing gradient problem using gated mechanisms. Is able to model long-term dependencies that are needed for the gradual degradation.
- *Gated recurrent units (GRU):* Introduces gating to control information flow in fewer parameters than LSTM cells. It has structural differences and computational advantages with similar performance to the LSTM.

- *Bidirectional architectures:* Propagate information in both directions allowing the network to have a context of the past and future of a certain point in the sequence at each time step.
- *Applications:* Can be used for sequence classification (simulation fault diagnosis), sequence prediction (predicting future values), anomaly detection (develop a model of normal operation).

RNNs are effective in modeling temporal data but are typically sensitive to the choice of hyperparameters and may still face challenges to model very long sequences.

Autoencoders

Autoencoders are used to learn compact representations for unsupervised anomaly detection.

- *Basic architecture:* The encoder converts the input to a compressed form to create a latent representation, then the decoder regenerates the input. This regeneration error is used as a score for anomaly detection.
- *Variational autoencoders (VAEs):* This probabilistic framework is used to learn parameters of the latent distribution. It can also generate synthetic data and estimate uncertainty.
- *Denoising autoencoders:* These are used to generate clean data from corrupted samples and increase tolerance to noise and missing data.
- *Sparse autoencoders:* These include sparsity constraints in the latent representation.
- *Applications:* Unsupervised anomaly detection (trained only on normal samples), feature learning (finding compact representations), data completion, and conditional monitoring (tracking the changes in the latent space).

Autoencoders deal with the shortage of labeled data but the selection of thresholds needs to be tuned and during transients, they may produce false alarms.

Hybrid Architectures

Some recent works have combined building blocks of Autoencoders:

- *CNN-LSTM hybrids:* In these, CNN is employed for learning spatial features and later LSTM is used for learning temporal features. Even multivariate time series in the presence of cross-sensor dependencies and temporal behavior can be effectively analyzed using this hybrid network.
- *Attention mechanisms:* These mechanisms can focus on important time steps or features while learning the representation. The recent Transformer architecture which is based solely on attention mechanisms and does not require recurrence and convolution, shows state-of-the-art performance in modeling long-range dependencies.
- *Temporal convolutional networks (TCNs):* In contrast to LSTMs, these models are able to capture long-range dependencies more efficiently using dilated convolutions.
- *Graph neural networks (GNNs):* These models system topology as a graph where nodes represent components and edges represent connections between the components. This enables the building of the system model by learning the faulty conditions of individual components. GNNs are more suited for reasoning about relationships during fault propagation and hence are suitable for fault diagnosis of multi-component systems.

Performance Advances

Deep learning has the potential to significantly increase performance in fault detection and diagnosis as well. Recent works reported in have shown the following benefits: 1) 5-15% increase in accuracy compared to classical methods, 2) reduction in false alarms, and 3) up to a 20-50% reduction in detection delay. Regarding diagnosis, discrimination between similar faults has been improved, and the simultaneous identification of multiple concurrent faults has been enabled. In general, both detection and diagnosis tasks have been enhanced in terms of generalization. New deep learning approaches have also been proposed, such as multi-task learning for the joint optimization of detection and diagnosis architectures, meta-learning for rapid learning system reconfiguration, and self-supervised learning for the exploitation of the abundance of unlabeled system operation data. Nonetheless, deep learning still poses the well-known challenges of interpretability, high computational demands, and the requirements to train on a vast amount of labeled data.

HYBRID AND EMERGING METHODOLOGIES

Physics-Informed Neural Networks (PINNs)

Physics-Informed Neural Networks incorporate prior knowledge into deep learning: Physics-Based Loss Functions: Incorporate physics into the loss function on top of the data term (e.g. conservation laws, thermodynamic relationships). Hybrid Designs: Combine data-driven and physics-based processing via parallel information streams (extracted features/encoded prior knowledge), sequential processing stages (physics-based data preprocessing) or embedded explicit/implicit constraints. Use Cases: Sampling from the physics to generate fault scenarios for data-driven models, determine extrapolation guidelines for transferring models generated in a physics-driven regime, estimate measurement errors by calculating propagated physics-tolerance. Physics-Informed Neural Networks combine the flexibility of deep learning with the interpretability and generalizability of physical models. They induce a shift of complexity from learning the dynamics to properly accounting for the physical priors.

Transfer Learning

The cost of labeled data is a major driving factor of transfer learning in the context of commercial deployment. Various transfer learning approaches are prevalent, relying on features, instances, parameters, and few-shot learning strategies. A wide range of applications such as cross-system generalization, fault type extension, and condition adaptation are also found to exist. The study has identified many commercial solutions in the literature using different transfer learning approaches. It is concluded that transfer learning is well-suited for commercial deployment considering diversity in the systems and lack of data.

Digital Twins

Digital twins are virtual replicas of a physical system, embraced widely in the fault detection and diagnostics community. Mechanically they feature high-fidelity simulations, where physics-based models like Modelica or EnergyPlus, high-performance computing generates massive volumes of operational data for a virtual replica. More commonly hybrid digital twins, are physical digital twin mathematical models, coupled with data-driven digital twins that make decisions about the physical digital twin functionality and provide decision support updates in real time. Finally, high dynamic fidelity is maintained for a computationally obtainable virtual physical system through all these methods.

Edge Computing and Federated Learning

Deployment considerations are motivating architectural innovations:

- *Edge AI:* Deploying lightweight models on IoT devices for real-time inference, reduced cloud dependency, enhanced privacy, and offline operation.
- *Model compression:* Quantization (reduced precision), pruning (removing redundant parameters), knowledge distillation (training compact models), neural architecture search (automated efficient design).
- *Federated learning:* Collaborative training across multiple systems without data centralization. Preserves privacy while leveraging diverse experiences. Challenges include system heterogeneity and synchronization.

Explainable AI (XAI)

The opaque "black box" nature hinders technician trust. Early XAI techniques helped to trust post-hoc explanations (LIME to approximate a model with interpretable local ones, SHAP, gradient-based methods). Interpretable features of the model itself involve attention (influential time steps and/ or features), prototypes (representative training examples), and rules. Domain-specific solutions provide saliency by being physically relevant, based on fault progression, or by calibrating confidence. XAI has evolved from generic techniques to solutions that are in a certain degree domain-specific and align with the mental model a technician has of the system. The primary challenge and obstacle to adoption persist -an insufficient quantity of labeled fault data to train condition monitoring systems. Data is still sparse and highly imbalanced because of the relative infrequency of faults and the typically low fault-to-normal data ratio in real-world systems.

IMPLEMENTATION CHALLENGES

Data Challenges

This imbalance is further exacerbated as the proportion of normal system operation data can exceed 99%.

- *Quality concerns:* Sensor noise, calibration drift, missing values, and synchronization errors sensor noise, calibration drift, missing values, and synchronization errors degrade performance. Robust preprocessing and noise-tolerant architectures are the prerequisite for success.
- *Multi-modal operation:* Different operating (ambient temperature, load, set points, etc.) conditions may lead the system to different operational modes complicating pattern recognition.
- *Temporal characteristics:* Number of faults demonstrate fast and slow time characteristics—real-time/low-latency requirements for fault detection for fast and early fault detection to capture subtle information indicative of faults.

Algorithmic Challenges

- *Generalization vs specialization:* While it would be ideal if a model could be designed so that it is generally applicable to every specific system (i.e., no data needs to be collected for training), in practice, there is always a dataset from which to train, and often the tension between cross-system generalization and installation-specific optimization must be somehow addressed. Context-Aware Adaptation will help to mitigate this.
- *Adaption to aging:* aging of systems also corresponds to the concept drift in data science. Over the lifetime of a piece of equipment, its operation will change slightly from when it was first installed, the characteristics of its sensory signals may vary, and it may be repaired causing a step-change in some behaviors. A long period after installation, replacement, repairs, etc., it may no longer represent the system it was designed to as much as a new system fitted in its position when the original system failed, supporting its last days of operation.
- *Real-time processing:* Efficient algorithms are necessary due to Edge device limitations. Compression, pruning, and hardware-aware design are being actively investigated by the research community.
- *Uncertainty quantification:* There is a need for reliable quantification of the model‘s uncertainty for safety-critical decisions. Bayesian deep learning and ensembles provide principled solutions at the cost of increased computational demand.

Integration Challenges

- *Legacy system integration:* The retrofitting of FDD to legacy systems having a limited number of sensors requires the development of creative utility of measurements and low-cost additional sensor modalities. Standardization Gaps: Current standardization gaps related to the absence of standardized protocols of data exchange, data formats, and application program interfaces (APIs).
- *Human-machine interaction:* Outputs must speak the language that human technicians can understand and act upon. Overly complex output interpretation configurations will also reduce the rate of adoption irrespective of the performance of the underlying algorithm.
- *Economic justification:* Rationalize ROI by taking into account implementation costs, energy savings, maintenance optimization, and downtime reduction. Lifecycle analyses indicate that returns are positive but the net present value and internal rate of return are necessary, however, these methods require the investment costs of the implemented system and its alternatives. Further, net benefits should be differentiated from total lifecycle costs.

Evaluation Methodologies

For lifecycle costs, external costs are often difficult to quantify, e. g., how to account for the societal cost of greenhouse gas emissions. For industry, the value of lost production during downtime has to be estimated or assume that the loss of production during maintenance would be transformed into paying overtime or contract workers to meet the delivery deadline. The economic lifetime of the implementation, as well as the expected technical and economic life of the system, needs to be estimated. Long-term impacts could be, for example, the increasing importance of predictive maintenance in Industry 4.0. But also, technological advancements could lead to considerable changes in costs, e. g., in the price of IoT sensors and the cost to monitor them. This particularly applies to cognitive solutions as they can contribute to automatization and saving manpower, especially when considering the cost of labor. This also leads to an additional value, which is an information surplus in the sense of Industry 4.0. The economic evaluation of this information surplus is, however, not straightforward as it depends on the process-specific value of monitored data. Raising investments costs would lead to a decreasing ROI. These problems are commonly underestimated by the lifecycle costing community such as. From a technical point of view, investments could not rise excessively, as more important performance metrics compared to computational costs. Lifecycle analyses and thereby ROI thresholds could only provide a rough hint on the upper limit on actuating-related costs out of a timeline.

Regulatory Landscape

- *Building codes:* With stricter requirements in energy codes like ASHRAE 90.1 and California Title 24, FDD could be specified based on higher detection performance levels.
- *Certification programs:* Similar to building codes, manufacturer certifications such as ENERGY STAR and Eurovent could use the availability of FDD features as product differentiation criteria.
- *Data privacy:* With cloud-based solutions, the implementation of data protection directives like GDPR and CCPA can restrict the use of FDD in certain application contexts.
- *Interoperability standards:* Industry-defined data models like Project Haystack and Brick Schema can be used for semantic annotation to enable cross-algorithm development for different systems.

CONCLUSIONS AND FUTURE DIRECTIONS

Key Findings

This review paper provides an overview of RAC FDD AI/ML evolution from classical to current state-of-the-art deep learning:

1. *Performance:* Detection accuracy increased from classical ~70-80% to ~85-95% (deep learning), with diagnosis accuracy 5-15% lower.
2. *Shift in methodology:* Evolved from manual engineered feature designing to automated learning, single-algorithm to hybrid, and solely data-driven to physics informed.
3. *Dynamic of data:* Algorithms matured drastically but data paucity still limits the application.
4. *Gap in implementation:* Laboratory successes surpass real-world applications, mainly due to lack of an effective integration path, economics, and human dynamics.
5. *Rising consensus:* Best solutions feature synergy of multiple techniques, based on the nature of the data and application, as well as limitations of the deployment scenario.

Research Gaps

- *Longitudinal studies:* There have been very few rigorous evaluations of the technology as systems degrade over years.
- *Cross-system generalization:* Little progress has been made in developing robust methods for transferring lessons learned about fault signatures and modeling strategies across different but substantially similar cyber-physical systems.
- *Human-centric design:* Prior research has shown that, in practice, too little (if any) attention has been given to the real-world requirements for successful adoption of FDD by maintainers and technicians.

- *Economic validation:* Despite more than a decade of research advances and greater investment in model-based engineering tools in industrial practice, there have been relatively few published reports of real-world applications and limited rigorous analyses available in the literature that quantify lifecycle-wide total cost of ownership and value creation for different stakeholders over the systems‘ lifecycle.
- *Standardization initiatives:* While progress has been made, several gaps remain. For example, evaluation protocols and benchmark datasets and their respective interfaces to cyber-physical systems (or their emulators and high-fidelity simulators) are critically needed to facilitate more meaningful and realistic comparative assessments.

Future Directions

Algorithmic Innovations

1. *Foundation models:* These are pre-trained on diverse plant data enabling rapid and effective fine-tuning with minimal labeled data from the new system.
2. *Causal discovery:* In addition to correlating monitored signals, understanding and potentially exploiting causal relationships.
3. *Multimodal learning:* Combining information from different types of sources (e.g., image, text, acoustic).
4. *Reinforcement learning:* Enabling adaptive control strategies, e.g., some fault impacts could be acceptable while waiting for an opportunistic repair.

Implementation Advancements

1. *Plug-and-play FDD:* Standardized software packages that can be plugged into different systems [131].
2. *Blockchain:* Providing an immutable maintenance history to train the algorithms.
3. *Augmented reality:* Visualizing in 3D faulty components and associated maintenance/repair procedures.
4. *FDD-as-a-service:* A cloud-based service that does not require a large investment in infrastructure.

System Integration

1. *Grid-interactive FDD:* Considering electricity prices and demand response signals to determine the best time to perform particular maintenance actions (such as battery recharging).
2. *Lifecycle integration:* Tracking the health and maintenance/repair history of equipment from its installation to decommissioning.
3. *Circular economy integration:* Engineering information to estimate the remaining physical life of a component to support the decision of whether to repair, replace, or refurbish it and put it back into service.

Conclusion

Using AI/ML for RAC predictive maintenance and FDD is becoming mainstream. The greatest energy-saving potential lies in integrating data-driven methods with first principles to predict performance degradation and malfunction before they occur. These hybrid approaches derive benefits from both the flexibility of data-driven methods and the physical insight of first-principles models. Data-driven methods require significantly less engineering effort than first-principles models to implement and maintain. However, first-principles models are more interpretable and transferable to novel conditions.

Smart FDD systems will be necessary to achieve future performance and energy goals, especially for complex and dynamic building systems and the urban systems into which they are integrated. The design and operation of such systems in the future will greatly benefit from connecting with—and perhaps even directly building on—the foundational ideas, methods, and technologies under research and development today.

REFERENCES

1. IEA. (2022). The Future of Cooling. International Energy Agency.
2. UNEP. (2021). Cooling Emissions and Policy Synthesis Report. United Nations Environment Programme.
3. Mobley, R. K. (2002). An Introduction to Predictive Maintenance. Butterworth-Heinemann.Elsevier. USA.
4. Selcuk, Sule. "Predictive Maintenance, Its Implementation and Latest Trends." Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, (2017). Accessed February 7, 2026. https://doi.org/10.1177/0954405415601640.
5. Katipamula, S., & Brambley, M. R. (2005). Methods for Fault Detection, Diagnostics and Prognostics for Building Systems - A Review Part. HVAC&R Research, 11(1), 3-25.
6. Breuker, M.S. & Braun, J.E., 1998a, Evaluating the performance of a fault detection and diagnostic system for vapor compression equipment. HVAC&R Res., 4 (4) (1998) 401–425. https://doi.org/10.1080/10789669.1998.10391412.
7. Beghi, A., Cecchinato, L., Corradi, M., & Rampazzo, M. (2016). A data-driven fault detection and diagnosis strategy for rooftop air conditioning units based on support vector machines. International Journal of Refrigeration, 69, 138–152
8. Kim, W., & Katipamula, S. (2018). A review of fault detection and diagnostics methods for building systems. Science and Technology for the Built Environment, 24(8), 847–869.
9. Wang, H., Chen, Y., Li, C., & Li, H. (2020). A hybrid fault diagnosis method for HVAC systems based on deep belief network and Dempster–Shafer evidence theory. Measurement, 151, 107199..
10. Shi, W., Xie, X., Chu, X., & Gadh, R. (2016). Data-driven methods for fault detection and diagnosis in HVAC systems: A review. IEEE Access, 4, 2711–2738..
11. Pritoni, M., Li, H., Hong, T., Peffer, T., & Granderson, J. (2021). A synthetic building operation dataset. Scientific Data, 8(1), Article 213, 1–16
12. Granderson, J., Lin, G., Casillas, A., Harding, A., Im, P., & Chen, Y. (2017). Curation and evaluation of ground-truth validated benchmarking datasets for fault detection and diagnostics in building systems.
13. Mirnaghi, M. S., & Haghighat, F. (2020). [Title not indexed in major open sources] Journal of Building Engineering, 31, 101383.
14. House, J. M., Vaezi-Nejad, H., & Whitcomb, J. M. (2001). An expert rule set for fault detection in air-handling units. HVAC&R Research, 7(4), 347–369. 2001
15. Schein, J., Bushby, S. T., Castro, N. S., & House, J. M. (2006). A rule-based fault detection and diagnostic method for air-handling units. ASHRAE Transactions, 112(2), 554–567.
16. Li, H., & Braun, J. E. (2007). An overall performance index for characterizing the economic impact of faults in direct-expansion cooling equipment. International Journal of Refrigeration, 30(2), 299–310.
17. Du, Z., Fan, B., Jin, X., & Chi, J. (2014). Sensor fault detection and its efficiency analysis in air-handling units using combined neural networks. Energy and Buildings, 85, 441–447.
18. Zhao, Y., Xiao, F., Wen, J., Lu, Y., & Wang, S. (2015). A data-driven fault detection and diagnosis method for building energy systems using pattern recognition. Energy Conversion and Management, 103, 180–191.
19. Swider, D. J. (2003). Fault detection and diagnosis in refrigeration systems — A modeling approach. International Journal of Refrigeration, 26(4), 434–444.
20. Yan, R., Ma, Z., Zhao, Y., & Kokogiannakis, G. (2016). A decision-tree based data-driven diagnostic strategy for air-handling units. Energy and Buildings, 133, 37–45

