IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 October 2024, accepted 14 November 2024, date of publication 20 November 2024,
date of current version 29 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3502546

TOPICAL REVIEW

Review of Artificial Intelligence Methods for Faults Monitoring, Diagnosis, and Prognosis in Hydroelectric Synchronous Generators

HELENE BECHARA[1], RONY IBRAHIM[2], (Member, IEEE), RYAD ZEMOURI[3], BACHIR KEDJAR[1], AREZKI MERKHOUF[3], (Senior Member, IEEE), ANTOINE TAHAN[2], AND KAMAL AL-HADDAD[1], (Life Fellow, IEEE)

[1]Department of Electrical Engineering, École de technologie supérieure (ÉTS), Montreal, QC H3C 1K3, Canada
[2]Department of Mechanical Engineering, École de technologie supérieure (ÉTS), Montreal, QC H3C 1K3, Canada
[3]Hydro Québec Research Institute (IREQ), Varennes, QC J3X 1S1, Canada

Corresponding author: Helene Bechara (helene.bechara.1@ens.etsmtl.ca)

ABSTRACT This scientific article aims to provide a comprehensive review of fault monitoring, diagnosis, and prognosis methods based on Artificial Intelligence (AI) for Hydroelectric Generator Units (HGUs). It presents a compilation of research studies that have utilized AI models for fault monitoring, diagnosis, and prognosis in HGUs. Additionally, it outlines the process for building an AI model in the context of fault management in HGUs and discusses the advantages and disadvantages associated with AI methods in this domain. Furthermore, the article examines the research prospects and trends of AI models for fault management in HGUs. By synthesizing existing literature and highlighting future directions, this article serves as a valuable resource for researchers and practitioners seeking to leverage AI techniques for effective fault management in HGUs.

INDEX TERMS Artificial intelligence (AI), diagnosis, hydroelectric generator unit (HGU), monitoring, prognosis.

I. INTRODUCTION

The hydroelectric machine is an important source of renewable energy and is used to generate electricity in many parts of the world. However, they are multi-failure mode systems evolving through a great number of components and complex failure mechanisms. Fault monitoring, diagnosis, and prognosis (MDP) of hydroelectric generators are essential for ensuring the reliable and efficient operation of these renewable energy sources. By identifying and diagnosing potential problems before they become serious, maintenance planning can be optimized to reduce the risk of costly repairs or replacements and the risk of downtime due to generator failure. Fault monitoring focuses on detecting and identifying faults as they occur, while fault diagnosis involves analyzing the detected faults to determine their type, severity, causes, and location. Prognosis goes a step further by predicting the degradation assessment and remaining useful life as presented in Figure 1. Fault detection is the process of monitoring the system for any signs of a fault or malfunction. Different parts of hydroelectric generators can have faults and these faults can be electrical or mechanical which they can be monitored by different techniques like electromagnetic field monitoring, temperature measurements, etc. [1], [2], [3], [4]. IEEE1129-2014 standard defined the online monitoring techniques and guidelines of large synchronous generators [5]. Degradation mechanisms were described, and all online monitoring methods and instrumentation were explained. HGU components, some monitoring sensors and faults are illustrated in Figure 2. Condition monitoring techniques based on magnetic fields for electrical machines are compared in [6] and applications of magnetic flux for faults diagnosis of wound field synchronous machines are discussed.

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Liu.

Fault diagnosis involves identifying the fault type, its severity, cause and its location. The condition monitoring and fault diagnosis techniques of electrical machines were reviewed, discussed and compared in [7] and [8].

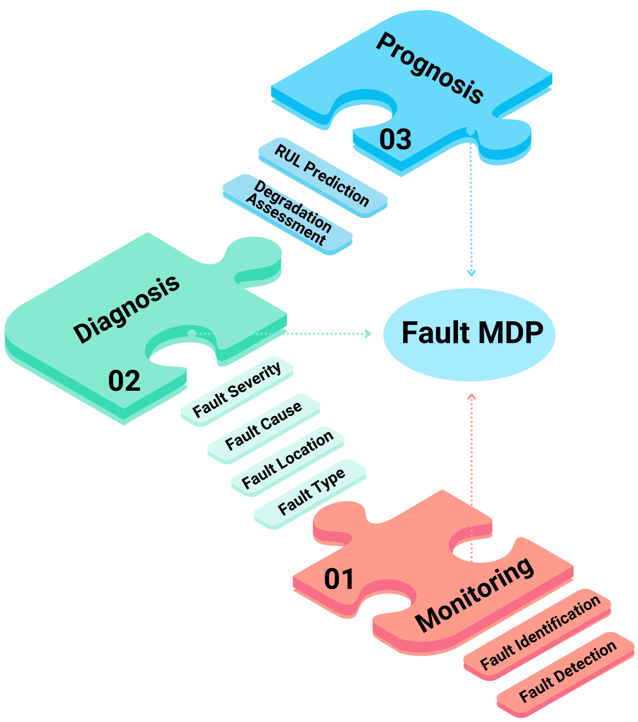


FIGURE 1. Fault monitoring,diagnosis and prognosis.

Reference [9] overviewed diagnosis methods based on electromagnetic fields (air-gap and stray fluxes) for large synchronous machines diagnosis. There are many methods based on different signal type proposed to diagnose faults in hydroelectric generator. For example, [10] proposed mathematical models based on vibration signals to assist fault diagnosis, however [11] and [12] proposed methods based on electromagnetic fields to diagnose eccentricity fault. Diagnostic tools aid the researchers to perform fault prognosis. For example, MIIDA (Methodology for Integrated Diagnostic of hydroelectric generators), a web-based application used by Hydro-Québec (Canada) for condition-based maintenance, calculates a health index for large hydroelectric generators by using data from actual diagnostic tools and inspections [13].

A prognosis provides the basis for predictive maintenance. It can be an assessment of the remaining useful life of repairable systems, a prediction of the future state of the system or a representation of degradation evolution [14]. Fault prognosis provides useful information that support decision-making and optimize the maintenance strategies. Reference [15] proposed a prognosis method based on Physics of Failure, data and expert knowledge and estimated the failure propagation by using the Petri-Net technique. The prognosis model in [16] relied on the failure mechanisms. According to degradation states and data obtained from the diagnosis tools, the failure mode is predicted. By utilizing different types of sensors to monitor the status and performance of complex systems, utilities can leverage the latest advances in MDP technology to reduce maintenance costs, minimize unscheduled outages, and prevent catastrophic failures.

Reference [17] showcases the viability of using artificial neural networks for the monitoring of electrical machines' condition. While [18] presents an in-depth review of the application of deep learning techniques in renewable energy, providing an assessment of their performance and discussing the main challenges and opportunities for further research in the field. References [19] and [20] underscore the growing acceptance of AI-based data-driven approaches in the fields of electric machine drives and fundamental sciences, serving as a catalyst for researchers to gain a profound understanding of AI applications and contribute to the ongoing advancement of these fundamental sciences.

The use of AI models and reinforcement learning for predictive maintenance is becoming increasingly popular, by combining them the predictive maintenance can be more accurate and cost-effective [21]. AI-based models can handle a large volume of data, adapt to new information, and detect anomalies that can be missed by experts or traditional methods, whereas they require big data to be trained, tested and validated which can be the main drawback when using them for fault MDP of large hydroelectric generators. As these machines are well-monitored and quickly repaired, it is difficult to find large faulty data. Moreover, fault cannot be implemented on large HGUs as they are only designed for power generation not for conducting experiments.

The aim of this paper is to present, analyze, and classify the AI based methods specifically applied for fault MDP of hydroelectric generator units. The work is divided into the following sections. Section II describes the methodology employed for the literature review. Section IV provides a compilation of research studies that have utilized AI models for fault monitoring, diagnosis, and prognosis in HGU. Section III describes the various AI techniques by explaining the techniques used for fault MDP of HGU. Section V presents the process for building an AI model. Section VI discusses the advantages and disadvantages of AI methods in this domain. Section VII outlines the research prospects and trends of AI models. Lastly, the article culminates with a comprehensive conclusion that encapsulates the key findings and insights derived from the preceding sections.

II. LITERATURE REVIEW METHODOLOGY

The literature review aims to present most of the studies that have investigated methods based on artificial intelligence for detecting, diagnosing, or prognosing faults in hydroelectric generators. The review predominantly relied on the Web of Science (WOS) and SCOPUS databases, which are widely recognized as the standard and most authoritative repositories for scientific research. Different combinations of keywords such as hydroelectric generator, hydropower, hydraulic generator, monitoring, diagnostic, prognostic, artificial intelligence and fault were used to

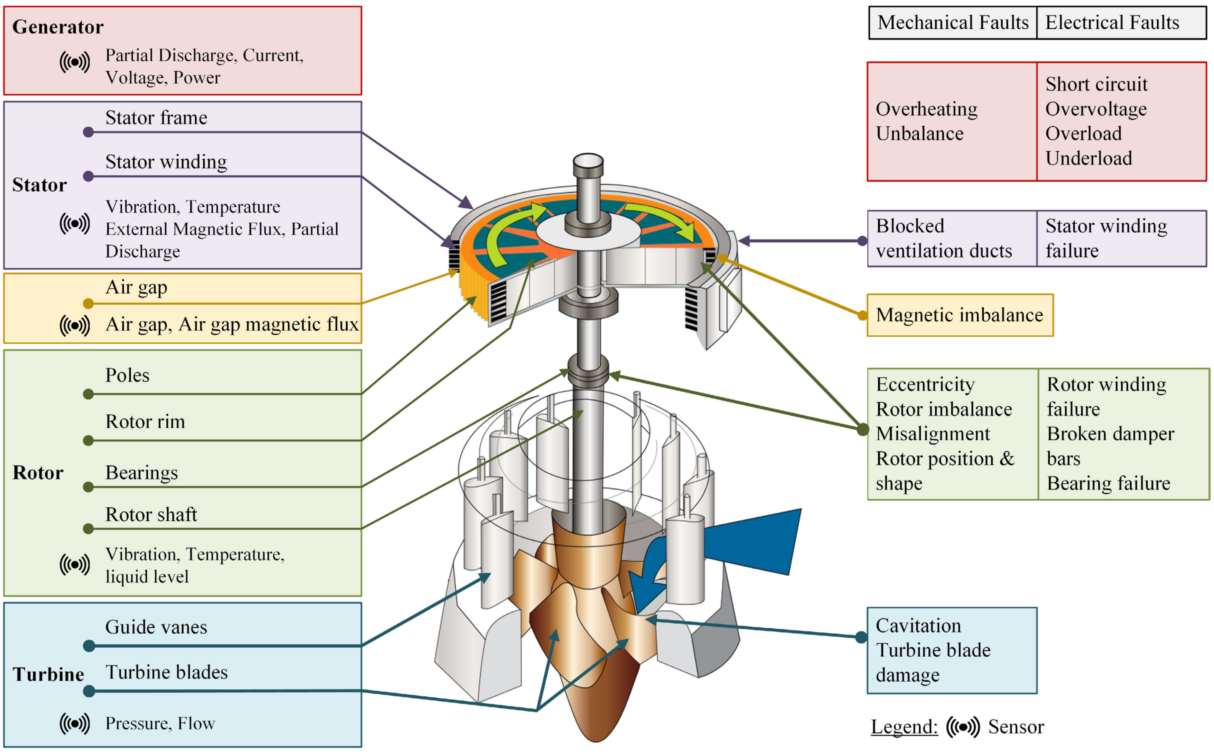


FIGURE 2. Hydroelectric generator unit fault monitoring, adapted from "https://en.wikipedia.org/wiki/Water_turbine".

capture the majority of the relevant studies, for example: hydropower AND fault AND artificial intelligence, etc. The retrieved documents were then filtered for the years 2004-2024 to highlight the evolution of research interest in applying AI-based methods to study faults in hydroelectric generators, including monitoring, diagnosis, and prognosis. During the research process, papers not written in English, notes, editorial materials, meeting abstracts, and retracted papers were excluded as they lack a sufficient level of detail. After screening the abstracts, studies that did not focus on studying faults in hydroelectric generators using artificial intelligence methods were excluded. Only studies that presented methods for monitoring, diagnosing, or prognosing faults in hydroelectric generators were considered eligible for inclusion in this review. The remaining papers were assessed for relevancy, and it was determined whether they would contribute valuable knowledge to this review. Furthermore, studies for which full texts were not available were also excluded. As a result, 35 research papers, suitable for full-text reading, were included and interpreted in this review.

Figure 3 provides a comprehensive overview of the publication trajectory of research papers in the field of fault analysis in hydroelectric generators with artificial intelligence from 2004 to 2024. The trajectory is measured by the cumulative number of publications, which serves as a quantitative indicator of the research output in this specific domain.

The figure is instrumental in discerning the publication trends and assessing the growth and development of scientific contributions in the field of fault analysis with artificial intelligence applied to hydroelectric generators.

During the period spanning from 2004 to 2015, researchers primarily directed their focus towards the utilization of artificial intelligence (AI) for the purpose of fault diagnosis. Throughout this timeframe, monitoring practices predominantly relied on methodologies that did not incorporate AI techniques. However, since 2015, there has been a discernible increase in the prevalence and application of AI techniques among researchers. These techniques are now not only employed for fault diagnosis but also for prognosis and monitoring tasks. The surge in AI adoption can be attributed to its inherent reliability and its ability to analyze large datasets, thereby facilitating the identification of common patterns across disparate data types and enabling the prediction of future values.

The discernible and sustained growth pattern, particularly observed between the years 2018 and 2024, serves as a clear indication of the heightened interest and emphasis placed on harnessing the capabilities of AI-based methods for fault monitoring, diagnosis, and prognosis within the domain of hydroelectric generators.

Consequently, the application of artificial intelligence in fault analysis for hydroelectric generators has garnered substantial global attention in recent years. This attention is primarily driven by the diverse range of applications offered by these methodologies, leading to an increasingly prominent and rapidly developing area of research.

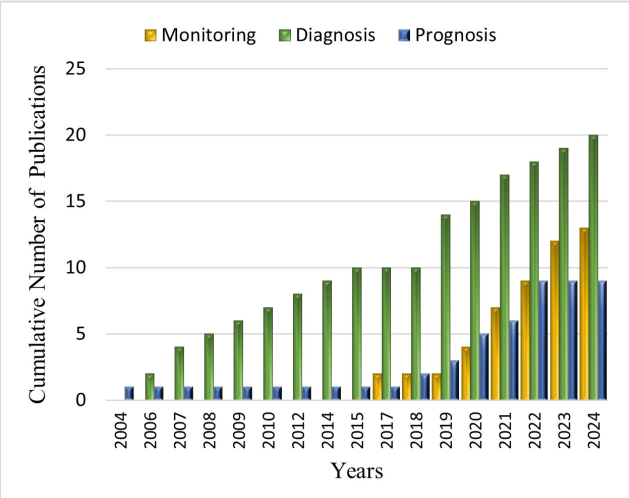


FIGURE 3. Cumulative number of publications over time.

III. OVERVIEW OF AI TECHNIQUES

AI is becoming increasingly prevalent in today's world and is seen as a crucial part of Industry 4.0, the fourth industrial revolution; as it has the potential to handle big data, increase efficiency and reduce costs [22]. It is a rapidly developing field of technology that allows machines to think and act like human expert. AI techniques are used to create autonomous systems that can learn from their environment and make decisions based on the data they receive. They are intelligent systems that are used to cluster data, recognize patterns, make predictions, and solve complex problems [23]. Figure 5 presents some of the most common AI-based methods including machine learning, deep learning, natural language processing (NLP), expert systems and others. Machine learning is a specific field of AI that focuses on the development of computer programs that can learn from data and has many subset techniques such as supervised and unsupervised learning, support vector machines and others [24]. The following AI-based techniques are mostly used for fault MDP of hydroelectric generators.

A. EXPERT SYSTEM

Expert systems were developed in 1970s by extracting knowledge from human experts and applying it to a computer program. Expert systems are computer programs that employ artificial intelligence to imitate the decision-making capacity of a human specialist. Based on knowledge processing, they can handle qualitative and quantitative data. Expert systems are used to detect, classify, diagnose and predict problems in electrical machines, as well as to aid in their repair and maintenance [25], [26], [27].

B. MACHINE LEARNING

Machine learning is a subset of artificial intelligence that enables computers to learn from data and experience. When attempting to solve a problem, data scientists emphasize that there is not a single algorithm that is the best for all situations. The type of algorithm used depends on the nature of the problem, the number of variables, and the model that would be most suitable. Machine learning has become increasingly popular in recent years because of its capability to quickly and accurately process large amounts of data. It can be used to identify and classify faults in electrical machines, predict their performance, and optimize their performance [28], [29].

Machine learning without guidance, known as unsupervised learning, is a branch of machine learning that relies on data to make decisions without direction. It utilizes algorithms to identify patterns and clusters in data, allowing for the detection of subtle, hidden patterns. Unsupervised learning has applications in diverse areas such as anomaly detection and data segmentation, clustering. This capability holds the potential for automating various tasks by analyzing vast amounts of information, thereby discovering intricate correlations [30].

1) NEURAL NETWORK

A Neural Network is a type of machine learning algorithm that uses interconnected nodes, or neurons, arranged in layers. Each neuron is connected to other neurons in the network and is responsible for processing information. The ANN is trained by providing it with a set of input data and a set of desired output data. The weights of the connections between the neurons are then adjusted using a process called backpropagation to reduce the difference between the desired output and the actual output, thus improving the accuracy of the ANN. To ensure accurate predictions, training should be done carefully to avoid overfitting. ANN is used for pattern recognition, classification, fault detection and prediction [31], [32].

Adaptive NeuroFuzzy Interference System (ANFIS) is a type of artificial intelligence technique that combines the learning ability of neural networks with the fuzzy logic of fuzzy systems. The ANFIS design consists of two steps: designing the premise parameters and training the consequent parameters. A hybrid learning algorithm is commonly used to train the ANFIS. Reference [33] proposed this technique and developed the model architecture. This technique is a powerful tool for fault detection, diagnosis and prediction [34], [35], [36].

Long Short-Term Memory (LSTM) architecture, developed by Hochreiter and Schmidhuber [37], is a type of Recurrent Neural Network (RNN) designed to address the vanishing gradient problem associated with traditional RNNs. LSTM is particularly effective at capturing long-term dependencies, making it well-suited for sequence prediction tasks. Unlike conventional neural networks, LSTMs feature feedback connections that enable them to process entire sequences of data rather than just individual points, enhancing their ability to recognize patterns in sequential information such as time series, text, and speech.

Multi-layer feed-forward neural networks (MFNN) involve a neural network that is aware of the desired output. The weights are adjusted to minimize the difference between the calculated output and the target output. MFNNs, particularly those trained with back-propagation algorithms, are among the most widely used. They can adapt independently without human's intervention. Neurons function as nonlinear devices, which means the entire network exhibits non linearity. This property is crucial when the relationship between input and output is inherently nonlinear. In supervised training, each example consists of a specific input and its corresponding desired output. The network processes examples from the training set, adjusting weights to reduce the discrepancy between the expected and actual outputs. This process continues until the network stabilizes, effectively creating a mapping between inputs and outputs for the given problem. MFNN demonstrate robustness, maintaining performance even as noise levels increase. However, ANNs also have some drawbacks such as for certain problems, convergence can be slow, and the high number of weights in an ANN often leads to lengthy training times [38].

2) AUTOENCODER

Autoencoders are a type of neural network used in unsupervised machine learning. It is made up of two components; an encoder that compresses input data into a hidden representation and a decoder that uses this knowledge to reconstruct the original data. They are widely used for tasks such as compression, denoising, feature extraction, and dimensionality reduction. Autoencoders can also help with the reconstruction of missing data points, making them an invaluable tool [39].

3) SUPPORT VECTOR MACHINE

SVM is a supervised machine learning algorithm that can be used for both classification and regression tasks. The algorithm creates a hyperplane or set of hyperplanes in a high-dimensional space to which data are mapped and classified. It is a powerful and versatile algorithm that can classify data even when they are not linearly separable [40]. SVMs can identify relationships between different variables, identify patterns in data that are not easily detected by other algorithms and identify anomalies in data. SVMs can be used to identify fault features that aid in diagnosing faults [41].

4) RANDOM FOREST

Random forest is a machine learning technique used for both classification and regression problems that constructs multiple decision trees to perform data mining tasks and produce a more accurate outcome. Random Forest is also known for its robustness to overfitting, accuracy, stability, and ease of user. It is capable of processing a wide range of descriptors at the same time while disregarding redundant or irrelevant ones [42].

IV. AI-BASED TECHNIQUES FOR FAULTS MDP

The following paragraph presents a compilation of research studies that have utilized AI models for the purposes of monitoring, diagnosing, or prognosing faults in HGUs. Table 1 provides an all-inclusive list of these studies, including their respective publication dates, the fault types investigated (electrical or mechanical), the types of input used (invasive or non-invasive), the specific input types utilized, the feature extraction techniques applied, the AI models employed for fault MDP, and the methods used for evaluating the models. The abbreviations of the feature extraction and AI-based methods mentioned in Table 1 are presented in Table 2 and Table 3, respectively, in the Appendix section.

As shown in Table 1, most studies focus on signals related to vibration, power, partial discharge, air gap flux, and stray flux, which highlights a considerable interest in these signal types within the research landscape and indicates their significance in the MDP processes of hydroelectric generators.

Upon analyzing the references, it is evident that a subset of studies employed non-AI techniques for feature extraction, while others opted for AI methods. To provide a comprehensive overview, Figure 4 presents a categorization of the feature extraction techniques into two distinct groups: non-AI based and AI-based. Each technique is explicitly linked to its respective reference. Furthermore, it is noteworthy that although all the MDP models utilized in the studies are based on AI, each method falls into a specific category. To facilitate a more comprehensive analysis of these methods, they have been distributed and organized in Figure 5 according to their respective categories, with each category associated with the corresponding reference.

By analyzing the data, it is evident that machine learning algorithms were predominantly used (84%) among the mentioned methods. Specifically, NN were utilized in 29% of the references, while SVM were employed in 13% of the cases. AE and PCA were each used in 11% of the references. Furthermore, KELM was utilized 6% among the used methods, while other methods accounted for 2% each. In addition to machine learning approaches, AI methods that do not fall under the machine learning category like ES, BN, and SSA were each employed in 4% among the used methods. Furthermore, AMWGO method and WFST were each used in 2% of the references.

A. ML-BASED TECHNIQUES

1) NN-BASED TECHNIQUES

NN have been widely adopted for fault diagnosis, prognosis, and monitoring HGUs. Their effectiveness in addressing complex pattern recognition challenges through the approximation of non-linear mapping relationships, along with their parallel processing capabilities for knowledge acquisition, has contributed to their popularity [49]. In their study, [49] utilized granular computing techniques to reduce the dimensionality of input data for fault diagnosis. Specifically

TABLE 1. Review of the papers using AI-based techniques for hydroelectric generator faults MDP.

Ref	Year	Fault Type		MDP	MDP Type		Input Signal	Feature Extraction	Method	Model Evaluation
		Elec.	Mech.		Inv.	Non-Inv.				
[43]	2004	-	✓	Prognosis	-	✓	Vibration	WPT	SOFM	Pearson correlation
[44], [45]	2006	-	✓	Diagnosis	✓	-	Multiple	GA-NLPCANN	BNN	Linear regression analysis
[46]	2007	-	✓	Diagnosis	✓	-	Vibration or Pressure	RWE	RBFNN	Other data type analysis
[47]	2007	-	✓	Diagnosis	-	✓	Vibration	WD	LS-SVM	Comparison with RBFNN's output
[48]	2008	-	✓	Diagnosis	-	✓	Vibration	-	Multi-NN and D-S evidence theory	Basic probability analysis
[49]	2009	-	✓	Diagnosis	✓	-	Vibration, Speed, Load and Flow	GC	NN	Comparing with threshold criterion
[50]	2010	-	✓	Diagnosis	✓	-	Multiple	-	ES	Trend of success
[51]	2012	✓	✓	Diagnosis	-	✓	Vibration	RST	SVM	Reliability
[52]	2014	-	✓	Diagnosis	-	✓	Vibration and Speed	SFVCC	SVM	Identification rates
[53]	2015	-	✓	Diagnosis	✓	-	Multiple	MNCWCP	RDM-WNN and WFST	Accuracy
[54]	2017	✓	-	Monitoring	-	✓	Partial Discharge	Projection in a normalized PRPD map	MFNN	K-fold cross-validation
[55]	2017	ns*	ns*	Monitoring	✓	-	Multiple	-	PCA	Precision and recall
[56]	2018	-	✓	Prognosis	-	✓	Vibration	FEEMD	SSA and KELM	RMSE and MAE
[57]	2019	-	✓	Diagnosis	✓	-	Multiple	MSWAD	LGPCA-RF	Precision
[58]	2019	-	✓	Diagnosis	-	✓	Vibration	GST	QPSO-SVM	Accuracy
[59]	2019	-	✓	Diagnosis	-	✓	Vibration	-	1-D CNN-GRU	Accuracy, precision, recall and F1-score
[60]	2019	ns*	ns*	Prognosis	-	✓	Vibration	VMD, SSA, PSR	KELM, AMGWO	RMSE, MAE, and MAPE
[61]	2019	✓	✓	Diagnosis	-	✓	Vibration	-	BN and ES	Conditional probability analysis
[62]	2019	ns*	ns*	Prognosis	-	✓	Vibration	EWT, Entropy-based reconstruction strategy, GSO	KELM and MOSSA	RMSE, MAE, MAPE and Pearson correlation
[63]	2020	✓	-	Monitoring	✓	-	Temperature	-	MWPCA	99.73% limits
[64]	2020	ns*	ns*	Monitoring	✓	-	Generated Power	-	IF	Average, max, min values, and standard deviations of AUC
[65]	2020	ns*	ns*	Prognosis	✓	-	Multiple	PCC, GCD, MIC	GPR, MD	Average relative error
[66]	2020	✓	-	Diagnosis	-	✓	Partial Discharge	EK	CVAE	Accuracy
[67]	2021	✓	-	Diagnosis	-	✓	Partial Discharge	EK	CVAE	Expert knowledge and confusion matrix
[68]	2021	✓	✓	Monitoring and Diagnostic	✓	-	Multiple	-	MWPCA-BN	SPE and SEN
[69]	2021	ns*	ns*	Monitoring	✓	-	Generated Power	HSIC-KNN-FS	RRCF	Average, max, min values and standard deviations of AUC
[70]	2021	✓	-	Prognosis	✓	-	Temperature	-	MWPCA	Case study analysis
[71]	2021	-	✓	Diagnosis	-	✓	Vibration	VMD	SVM	Accuracy
[72]	2021	✓	-	Monitoring	-	✓	Stray Flux	DWT and STWE	VAE	-
[73]	2022	ns*	ns*	Prognosis	✓	-	Multiple	-	LSTM-weight	Monitoring data analysis
[74]	2022	✓	-	Prognosis	-	✓	Stator Terminal Voltage, Stator winding Current	-	ANFIS	RMSE

TABLE 1. ***(Continued.)*** **Review of the papers using AI-based techniques for hydroelectric generator faults MDP.**

[75]	2022	-	✓	Prognosis	✓	-	Multiple	-	● DNNLR - LSTM with AE	Precision, recall and F1-score
[76]	2022	-	✓	Monitoring	-	✓	Vibration	-	● VAE	**99.73% limits**
[77]	2022	✓	-	Monitoring	-	✓	Stray Flux	● FFT	● VAE	Euclidean distance in latent space
[78]	2022	✓	-	Diagnosis	✓	-	Air-gap magnetic field	● FFT, DWT, TS-FRESH	● LR, KNN, SVM, linear SVM, XG-Boost DTF, MLP	Accuracy, SEN, SPE, precision and F1-score
[79]	2023	✓	-	Monitoring	-	✓	Vibration	● RMS	● VAE-SDL	Squared Prediction Error and Euclidian distance
[80]	2023	✓	-	Monitoring	✓	-	Partial discharge	● Canny edge filter, principle boundary, rough contour, amplitude and phase histogram	● U-Net, CNN	Accuracy, Recall, specificity, precision and F1-score
[81]	2023	✓	-	Monitoring and Diagnosis	-	✓	Vibration	● RMS	● VAE	Squared Prediction Error, Euclidian distance, and F_1*score*
[82]	2023	✓	✓	Diagnosis	-	✓	Vibration	● RMS	● VAE	Squared Prediction Error
[83]	2024	✓	-	Monitoring and Diagnosis	-	✓	Stray Flux	● RMS	● VAEC	Squared Prediction Error,
[84]	2024	✓	✓	Diagnosis	-	✓	Vibration	● RMS	● VAEC and Desirability	F1 score and Accuracy

ns*:not specified

targeting faults such as uncentering, movement rubbing, uneven rotor quality, tail water pipe eccentricity, and vortex formation, this approach demonstrated improved efficiency in fault diagnosis. Similarly, in [59], a novel artificial neural network model was proposed, resulting in significant improvements in the accuracy of fault diagnosis specifically tailored for HGUs. The authors utilized a 1-D CNN-GRU technique to diagnose various HGU faults, including eccentric draft tube surges, thrust bearing unevenness, hydraulic disequilibrium, unbalance, and mixed faults. The technique demonstrated robustness in the presence of changing conditions, enabling real-time diagnosis and exhibiting higher reliability compared to other AI-based methods. It is important to note that in cases where new faults arise, retraining the model with updated data becomes crucial to enhance classification performance and improve the diagnostic capabilities of the system. In their study, [48] incorporated the output of a local neural network diagnosis system as input at the decision level. This integration method resulted in enhanced precision for diagnosing faults associated with unbalanced rotor quality, misalignment of the generator's axis, rubbing between the stator and rotor, and a thinner main axis. On the other hand, RBFNN possesses local approximation characteristics and outperforms conventional BP (Backpropagation) neural networks in terms of learning rate, pattern recognition, and classification abilities. By training an RBFNN, it becomes possible to determine the type and severity degree of faults as the output. Reference [46] employed this approach to diagnose mechanical faults such as mass unbalance of the rotor, vortex in the draft tube, unsymmetrical entry of guide vanes, large ellipticity of the stator, rotor misalignment, and looseness of the stator core. Moreover, SOFM, a method categorized under NN, exhibits pattern recognition capabilities in fault diagnosis. Its objective is to classify input vectors belonging to different categories accurately. [80] employed the U-Net to extract the gap signature from other PD sources, producing an output image that serves as input for a convolutional neural network (CNN) model, which is then followed by a classifier to identify the PD sources automatically. Subsequently, a decision-making method is proposed to determine the optimal output category by considering the posterior and prior probabilities estimated by various individual models. In [43], the authors classified inputs into eight distinct categories, including shaft disequilibrium, shaft asymmetry, shaft scrape, vane looseness, shaft abrasion, shaft defection, shaft crack, and vane fracture. In the context of anomaly detection, LSTM neural network was utilized to analyze variables including bearing temperatures and vibrations. The LSTM model was employed to predict the temperature one hour ahead, taking into account the rate of variation in bearing temperature. The model showcased its ability to make accurate predictions of values closely associated with failures in the temperature of generator bearings [53]. In their work, [75] proposed two predictive maintenance techniques for fault detection: a deep Neural Network with Logistic

Regression and an LSTM model with an Autoencoder. These techniques leveraged the analysis of temperature, pressure, and voltage data to detect anomalies and classify faults, such as generator bearing faults, generator asymmetries, and forced pressures in the turbine. To enhance the performance of the models, the data underwent feature engineering that improved the training process. The LSTM model exhibited accurate predictions of bearing temperature, and the Deep Neural Network outperformed a simple neural network in terms of accuracy. As a result, these techniques hold promise for detecting other failures. However, further research is required to establish appropriate thresholds for failure alarms. Another study conducted by [73] employed generator apparent power, bearing hydraulic lubrication unit inflow, and bearing vibration as input variables for an LSTM-weight model. The study demonstrated that this technique possesses the capability to detect subtle changes, unlike traditional detection models, and exhibits a high level of prediction ability. In the realm of NN, [44] and [45] utilized various machine parameters, including load, flux, operating status, pressure fluctuation, frequency, vibration, throw, amplitude, and temperature changing parameters of the upper bearing, thrust bearing, and turbine bearing, in conjunction with BNN for diagnosing faults such as eccentric vortex band, unbalance, and misalignment. Reference [54] presented an innovative methodology in which MFNN was employed for automatic classification of partial discharge (PD) in order to assess the insulation condition of stator windings. This approach incorporated advanced noise filtering techniques, the extraction of novel features using image projection, and the training of multiple neural networks to achieve accurate PD classification. Furthermore, [74] employed ANFIS, a widely used prediction technique, to forecast the faulty temperature zone for prognosticating stator winding failure. Given the success of this technique in prognosticating stator insulation, it can also be applied to prognosticate insulation failure in other components, such as rotor winding insulation.

2) AE-BASED TECHNIQUES

In the field of machine learning (ML), alternative techniques have been explored for fault classification and monitoring. AE has been utilized primarily for fault monitoring, diagnosis, and occasionally prognosis. The application of VAE technique in the early detection of failure modes in large hydroelectric generators has gained attention due to its ability to reduce dimensionality and accurately diagnose faults in high-dimensional data. Reference [72] utilized the discrete wavelet transform (DWT) to analyze the stray flux signal, then applied Short-Time Wavelet Entropy (STWE) to extract features from the resulting subbands. Subsequently, a variational autoencoder (VAE) was employed in an unsupervised learning framework to organize the STWE signatures derived from stray flux measurements. The analysis of the latent space showed a significant correlation between specific trajectories in this reduced space and an increase in WE. Reference [76] employed VAE to monitor Rotor Inter turn Short Circuits (RITSC) using vibration signals. The study confirmed that VAE exhibits high sensitivity to fault occurrences and can detect faults at their early stages. Moreover, [77] utilized VAE to project stray flux measurements into a 2D space, enabling the monitoring of RITSC in large hydroelectric generators. This technique effectively distinguished healthy signals from faulty ones and clustered them based on the severity of the fault. VAE, a variant of the classical autoencoder, has demonstrated high accuracy in fault detection and dimensionality reduction. To extend the capabilities of the VAE, convolution layers are added to both the encoder and the decoder components, resulting in the CVAE. CVAE is another approach employed for data projection into a 2D-visualization latent space. The input vectors are encoded and represented in this 2D space, facilitating visual analysis of the spatial distribution of the training dataset. Reference [83] employed the CVAE to project stray flux signals into a 2D space for monitoring interturn short circuits in large hydrogenerators. The CVAE effectively clustered the signals into multiple groups, reflecting the severity of the faults present in the signals. Additionally, it was shown that the CVAE is robust against external noise and demonstrates greater sensitivity in fault detection compared to the standard method (RMS). References [66] and [67] utilized CVAE in conjunction with classifiers to diagnose stator winding insulation failure by classifying Partial Discharge Analysis (PDA) patterns. CVAE aided classification and data visualization by projecting data into a 2D space while preserving valuable features. However, it was observed that expert knowledge was necessary for feature extraction and identification of the source of the partial discharge, which significantly impacted the classification accuracy. Furthermore, the VAE has been employed to detect various levels of severity of the RITSC in large hydrogenerators by analyzing vibration signals, proving its capacity for fault classification within a 3D, user-friendly space, as demonstrated in [81]. This technique has also been extended to two other fault types, namely Static Eccentricity (SE) and Broken Damper Bar (BDB), with results confirming its effectiveness in fault classification, as noted in [82]. The method has been further enhanced by integrating it with a Sparse Dictionary Learning algorithm, as developed in [79], enabling therefore earlier fault detection without false alarms. Additionally, a new term based on the Desirability function was incorporated into the CVAE model's cost function in [84], standardizing the technique across two different large hydrogenerator designs. This improvement demonstrated the model's potential to classify the same fault in two different machines within the same geographical region.

3) SVM-BASED TECHNIQUES

SVM is a widely used technique in fault diagnosis. It is a supervised learning method suitable for pattern classification problems and has found applications in various domains

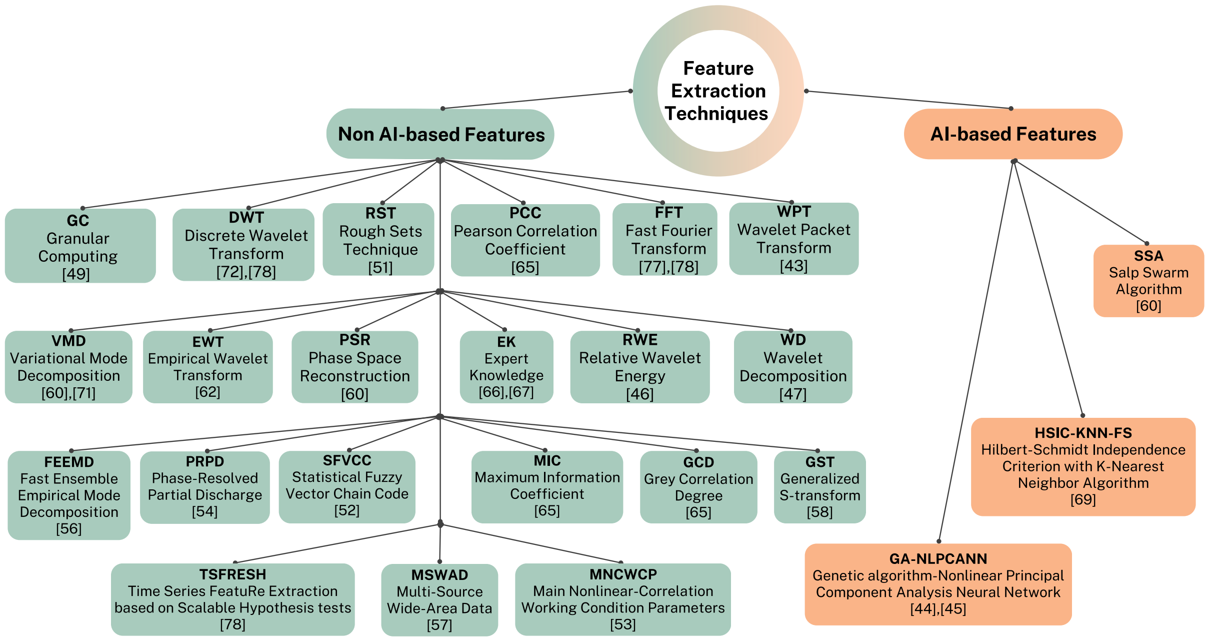


FIGURE 4. Feature extraction techniques used for fault MDP in hydroelectric generators.

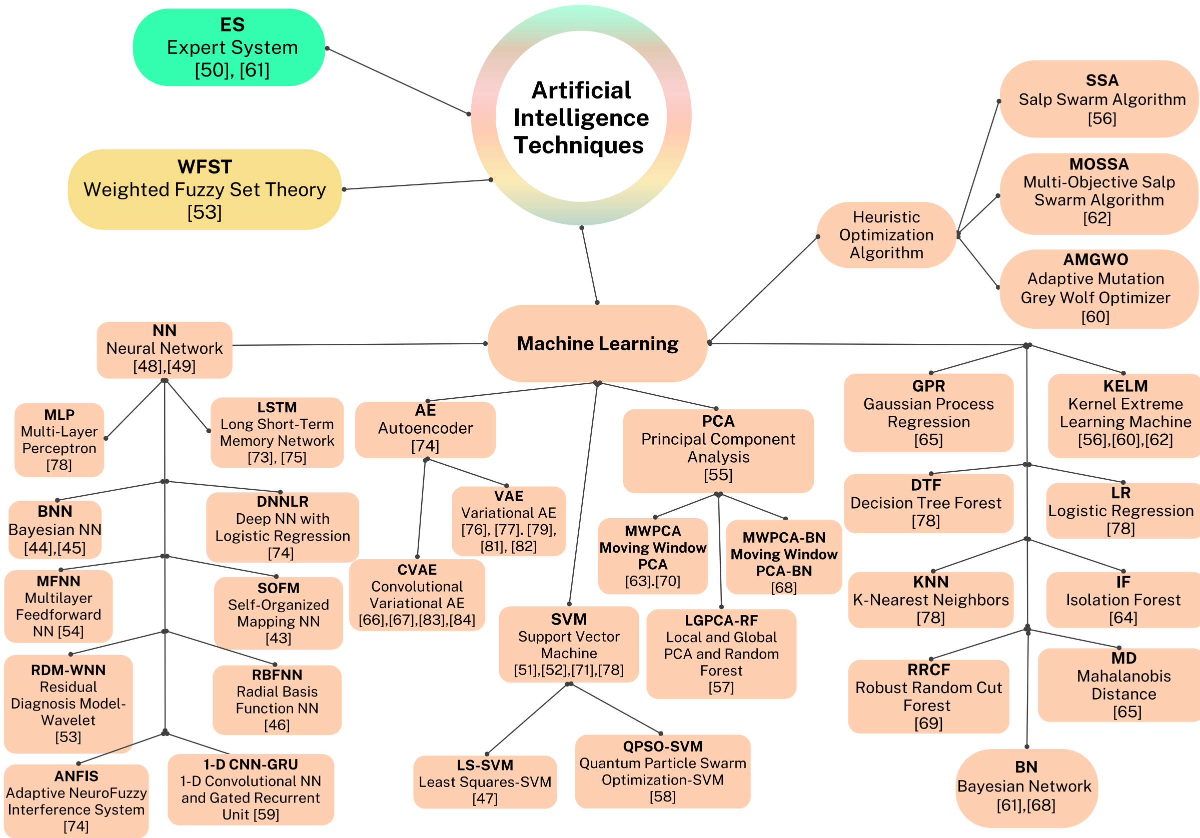


FIGURE 5. AI-based techniques used for fault MDP in hydroelectric generators.

such as target classification, pattern recognition, and fault diagnosis. Reference [52] proposed an innovative method for determining the correlation between vibration amplitude and rotating speed in hydroelectric generator units. They utilized a statistical fuzzy vector chain code to extract shape features from the vibration-speed curve and employed an SVM to accurately identify the type of vibration-speed curve, improving the efficiency and effectiveness of fault type identification compared to previous methods. While [71] introduced a fault diagnosis method using vibration signals from bearings. They decomposed the vibration signals into intrinsic mode functions using VMD. The calculated singular values of the resulting modal components, representing the energy characteristics of the maximum frequency, were used as feature vectors input to an SVM for fault diagnosis and recognition. Furthermore, LS-SVM was employed by [47] to diagnose faults in HGU, such as low-frequency vortex of the draft tube, asymmetry aperture of runner blade, and nonuniform air space between the rotor and stator. They utilized a multiclass classifier based on SVM for fault type identification, emphasizing the importance of optimizing model parameters for improved accuracy. On the other hand, [51] proposed a novel method for fault diagnosis in HGU by combining rough sets and SVM. Vibration signals were used, and fault patterns were extracted from overlapped regions. Rough set techniques were employed to define upper and lower approximations for each fault class. Moreover, [58] diagnosed mechanical faults like vortex belt eccentricity, unit axis misalignment, and dynamic-static rubbing using the QPSO-SVM approach. They extracted the energy eigenvector as a feature to reflect the spectrum characteristics, which was used as input to the diagnosis model. The QPSO technique enhanced the SVM's classification accuracy, and the model was found to be robust to noise. Increasing the number of learning samples could further improve SVM performance.

4) PCA-BASED TECHNIQUES

PCA is a widely used method for monitoring, diagnosis, and prognosis of industrial processes. It is a data-driven approach for anomaly detection, particularly in processes with a large number of variables. The basic idea of PCA is to remove noise and eliminate correlations between process variables. It achieves this by constructing a principal component subspace that contains the most important information of the original dataset, as well as a residual subspace that contains noise and unimportant information [55]. In their work, [55] proposed an improved PCA algorithm for detecting anomalies in HGUs. The approach involved identifying the operational conditions of the HGUs and using adaptive methods to update the PCA model. By considering variables such as the X and Y swing of the upper guide and the vertical and horizontal vibrations of the upper bracket, the enhanced method achieved higher precision and satisfactory detection rates for anomalies compared to traditional approaches. Furthermore, [57] diagnosed faults in HGUs, including rotor imbalance, rotor misalignment, spindle bending, dynamic and static rubbing, draft tube eccentric vortex, and hydraulic imbalance. They extracted features from multi-source wide-area data, including time and frequency domain features, air gap, magnetic pull force, and axis orbit shape features. The diagnosis model consisted of two steps: first, the dimensionality of the features was reduced using LGPCA, and then the reduced feature set was classified using the Random Forest algorithm for fault diagnosis. On the other hand, [63] proposed an adaptive variation of MWPCA for the early detection of aging in stator winding electrical insulation of HGUs using temperature data. The findings from simulated data demonstrated the capability of this technique in identifying faults at an early stage, thereby enhancing maintenance planning. Additionally, [70] studied the prognosis of stator winding insulation failure using MWPCA. The authors employed this technique to estimate the system failure date and determine the remaining useful life. The results showed that this technique is a powerful tool for fault prognosis, as it is more accurate and faster than traditional methods.

5) OTHER ML-BASED TECHNIQUES

Various machine learning models, such as KELM, GPR, DTF, and others, are mainly employed for prognosis and some for diagnosis, and monitoring tasks. For instance, the KELM incorporates a kernel function into the ELM to achieve a least-squares optimal solution, leading to improved generalization performance and stability compared to basic ELM. Moreover, KELM has the advantage of multi-output capability over traditional single-output Support Vector Regression (SVR), resulting in reduced training time. Therefore, KELM is considered more suitable for vibration tendency prediction [56]. In the work presented by [56], the authors aimed to predict the vibration tendency, which serves as an indicator of the health status and stability of HGUs. They utilized two AI-based techniques, SSA and KELM, to enhance prediction results. By employing feature extraction through FEEMD in conjunction with SSA and KELM, the proposed FEEMD-KELM approach outperformed predictions made by NN, SVR, and traditional KELM methods. On the other hand, [64] proposed a fault detection technique based on IF to automatically build models using normal data and successfully identify faults, even in the presence of non-linear correlations. The effectiveness of the method was evaluated by examining the generated power under both healthy and faulty conditions. Addressing the challenge of evaluating operational status and predicting failures in HGUs, [65] leveraged real-time monitoring data. They employed techniques such as PCC, MIC, and GCD to select appropriate operational state parameters from the extensive HGU system data. These parameters were then used to construct an input eigenvector, and a health assessment model for HGUs was established based on the GPR framework. Additionally, a condition monitoring directive based on the MD was designed. In [69], the authors explored a fault detection method for hydroelectric generators employing RRCF and conducted feature selection based on HSIC-KNN-FS. Their proposed method effectively addressed the challenges of scarce fault data and non-linear correlations in HGU data by selectively retaining features with distinct characteristics while eliminating those with similar characteristics. While in [78], the authors investigated signal processing techniques combined with different machine learning classifiers, such as LR, KNN, SVM, XGBoost DTF and MLP, for the detection of interturn short circuit faults in hydroelectric generators. The classifiers were trained on a dataset of spectral and waveform features extracted from air-gap signals. The proposed approach achieved an 84.5% detection rate for interturn short circuit faults with a 92.7% accuracy in fault detection.

B. OTHER AI-BASED TECHNIQUES

Various AI techniques, like BN, are employed in the field of fault diagnosis, and in some cases, for monitoring purposes as well. In the study conducted by [68], a hybrid framework combining MWPCA and BN is proposed for automated fault detection and diagnosis in complex systems. The framework utilizes various sensors, including vibration and temperature sensors, to detect electrical or mechanical faults. The results demonstrate the capability of the framework in successfully identifying and diagnosing multiple simulated failures in a hydroelectric generator.

On the other hand, AMGWO and SSA algorithms are employed for prognosis tasks. Reference [60] present a hybrid approach that combines VMD, SSA, and PSR techniques with KELM and AMGWO for accurate vibration tendency forecasting in HGUs. VMD decomposes the monitored vibration signal into components with different frequency scales, SSA extracts characteristic trends from nonstationary subseries, and PSR generates inputs and outputs for the prediction models. The study demonstrates the effectiveness of KELM for vibration tendency forecasting, and proposes the use of AMGWO to significantly improve the forecasting model. In their work, [62] put forward an intelligent approach for forecasting vibration tendencies in HGUs, with the objective of attaining a balance between stability and accuracy. The method leverages KELM and MOSSA to achieve this goal. The method involves several steps, including decomposing raw sensor signals using EWT, refactoring modes with a sample entropy-based reconstruction strategy, selecting important input features through GSO, and predicting refactored modes using KELM. The parameters of GSO and KELM are simultaneously optimized using the MOSSA.

An ES mainly used for diagnosis and it has been successfully applied in hydroelectric power plant maintenance. It has demonstrated high reliability and accuracy in fault diagnostics, resulting in effective maintenance planning and cost reduction. However, a disadvantage of expert systems is their reliance on predefined rules stored in the knowledge base (KB), which requires updating to detect new failure modes [50]. In their work, [50] developed an expert system for real-time fault diagnosis in complex systems, utilizing machine variables such as temperature, pressure, and flow. The system, integrated into an intelligent maintenance system, aids in maintenance planning and enables the transformation from time-based maintenance to condition-based maintenance. Moreover, [61] propose a fault diagnosis expert system that aims to facilitate the transition from time-based maintenance to condition-based maintenance. This intelligent tool integrates expert experiences and Bayesian inferences, offering advantages such as a comprehensive collection of expert knowledge, accurate simulation of expert thinking, and precise fault diagnosis. In their methodology, the authors thoroughly analyze hydraulic, electrical, and mechanical faults. They establish a precise Bayesian network and utilize a Noisy-Or modeling approach within the fault diagnosis expert system to effectively diagnose these faults.

Lastly, WFST is employed for quantitative fault status diagnosis in HGUs. Reference [53] successfully diagnose upper guide swing and pressure fluctuation faults in HGUs by using machine performance parameters and working condition parameters with RDM-WNN and weighted fuzzy set theory. The output of the AI-based RDM-WNN model is used as input for the weighted fuzzy set theory, enabling accurate fault detection even under load variation.

V. AI-BASED MODELS PROCESS

This section provides a comprehensive overview of the essential steps involved in utilizing AI-based models for fault monitoring, diagnosis, and prognosis in hydroelectric generators.

Upon thorough analysis of the references in paragraph IV, it becomes evident that the utilization of AI models in fault monitoring, diagnosis, and prognosis for HGUs entails a series of distinct steps, as depicted in Figure 6. The process initiates with the collection of relevant data from diverse sources, thereby forming a comprehensive dataset. Subsequently, feature extraction techniques are employed to extract meaningful and informative features from the collected data. The objective is to obtain robust inputs that enhance the AI model's capacity. Following the training and deployment of the model, its output is subjected to a rigorous evaluation process. It serves as a pivotal stage in validating the model's reliability and determining its applicability in real-world scenarios.

A. DATA COLLECTION

Data collection is an important part of AI-based methods, as it is the foundation of any model. Data acquisition involves discovering, augmenting, or generating new datasets. Some online tools, such as Google Dataset Search, are available to assist researchers in finding publicly available data on the internet. However, benchmark and datasets for large hydroelectric generator, such as external magnetic fields measurements, are still hard to find. Hence, the first question to ask is 'What data is available that can provide useful information and contribute to fault MDP?'. Sometimes a lack of data leads to the need to launch new measurement campaigns or to augment the existing datasets. Data quality checks are essential, as they affect the model accuracy [85]. Hence, to increase the data quality, one should clean the data, which will be discussed in the next paragraph. Moreover, one should also consider the data privacy and security implications of the data. Data can be confidential, and the privacy of data providers should be properly protected.

Machine learning can perform well with smaller datasets and often requires less data for training, whereas deep learning typically requires large amount of data to achieve high performance due to the complexity of its models.

B. DATA PREPROCESSING

The purpose of data preprocessing and its need is defined according to the collected data. So firstly, one should determine the reason for doing this step, data cleaning, data calculation or other. Moreover, data quality should be checked because sensors are electronic devices that are susceptible to faults, interruptions, or saturation. In addition, the measurements of sensors can be affected by the noise of the hydroelectric generators' industrial environment [86]. As explained in the book [87], data cleaning can be done by quantitative error detection which consists of detecting the values that are so different than the others, called 'outlier values'. Or by qualitative data cleaning which requires domain expertise and can be automated by detection algorithms to reduce data cleaning time. Furthermore, data transformation and data duplication should also be performed, if necessary. Data transformation is the process of changing the data format or normalizing it. If the transformation does not require external knowledge of the data, it is called syntactic transformation, and if it requires an understanding of the data, it is called semantic transformation. Data duplication consists of creating metrics to measure the similarity between two records. Data cleaning techniques should be wisely chosen as they depend on data type and size [88]. One should be cautious in building his datasets as the accuracy of the AI model greatly relies on data quality. One should also ensure that the data is properly constructed and that any outliers or anomalies are identified and addressed, as the AI model's accuracy greatly relies on its inputs.

Machine learning typically works with structured, tabular data, while deep learning directly handles unstructured data types.

C. FEATURE EXTRACTION

The quality of a model's output is influenced by data preprocessing and information redundancy. Feature extraction is a way of decreasing the number of features in a dataset while keeping the most important information. It is used to make a dataset simpler to analyze and to reduce its complexity. When selecting a feature extraction technique, the purpose of the extraction and the type of input should be considered. If the data is non-stationary, time-frequency techniques are recommended, and if frequency features are desired, the FFT method is usually used for spectral analysis [89]. Feature extraction can be used to reduce the dimensionality of a dataset, reduce the noise in a dataset, and improve the accuracy of a model. After building the database, it is split into three sets: training, testing and validation, to train, test and validate the model.

Machine learning requires significant manual feature engineering to select, extract, and transform relevant features, whereas deep learning automatically learns features from raw data.

D. MODEL TRAINING

The selection of the right model is a critical step in the machine learning process, as it determines the model's performance and accuracy. The choice of model depends on the complexity of the problem, the size and structure of the data, the computational resources available, and the desired output and level of accuracy. Some models classify data which are used to diagnose or monitor faults, other model predict values and are used to calculate future values to predict faults and system behavior. The process involves the selection of a set of hyperparameters, the training of the model with those hyperparameters and the training set, and the evaluation of the model's performance. The process is repeated until the hyperparameters are optimized [90].

Machine learning generally requires less time for training due to simpler models and fewer parameters, making it suitable for quick iterations. Whereas, deep learning involves longer training times because of the complexity of the models and the large volume of required data.

E. MODEL EVALUATION

Finally, the model undergoes testing using previously unseen data from the validation set to assess its performance. Performance metrics such as accuracy, precision, recall, and others can be used to evaluate the model's effectiveness [91]. In comprehensive reviews [92], [93], researchers have compiled and described different evaluation metrics that are also used for evaluating the performance of the model like the MAPE, MSE, F1-score, etc. Additionally, assessing the robustness of the model is also crucial, taking into account various factors that can introduce uncertainties in the deterioration process, the absence of run-to-failure data, sensor noise, unknown environmental and operating conditions, as well as engineering variations. These factors have the potential to impact the AI model and compromise its robustness [94].

Machine learning commonly employs standard metrics like accuracy, precision, recall, F1-score, and ROC-AUC for classification tasks, as well as Mean Absolute Error (MAE) and Mean Squared Error (MSE) for regression tasks. Whereas, deep learning models may use these standard metrics alongside additional specialized metrics, such as Intersection over Union (IoU) for image segmentation.

VI. ADVANTAGES & DISADVANTAGES OF AI METHODS

AI is a rapidly growing field of technology that has the potential to revolutionize the fault MDP approaches. Nevertheless, while AI has many potential benefits when

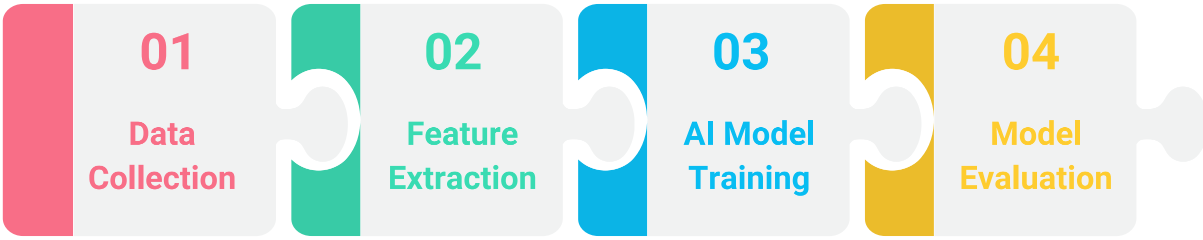


FIGURE 6. Model development process.

developed with a hydroelectric generator system, one may still encounter some challenges that should be considered before implementation.

A. BENEFITS OF USING AI FOR HGU'S FAULT MDP

1) **Increase efficiency and accuracy:**, A hydroelectric generator unit is a complex Multiphysics system with multiple components of different types, making data analysis difficult. Artificial intelligence (AI) can process quickly and accurately all kinds of input such as magnetic flux, vibration, and temperature, while uncovering fault-related signatures, trends and patterns, as well as identifying correlations and dissimilarities between data from datasets that are hard to detect with the human eye or traditional methods [95].
2) **Handling multimodal data:** Obtaining all the necessary information for fault MDP can be difficult when relying on just one type of input, which leads to the need for using different types of data such as images, records, text, etc. Though analyzing multimodal data with traditional methods is complicated, it is achievable with some AI-based models like multi-branch deep neural networks. Hence, one of the benefits of AI is that it can handle numerical and non-numerical data, extract features, and process them quickly and accurately [96].
3) **Handling Nonlinearity:** Some machine systems may exhibit nonlinear behavior, where outputs do not vary linearly with inputs. AI techniques, particularly Neural Networks, are well-suited for modeling these complex relationships, facilitating more accurate fault detection and diagnosis. This capability reduces the reliance on human expertise and minimizes the risk of oversight [97].

B. CHALLENGES OF USING AI FOR FAULT MDP IN HGU

1) **Lack of faulty data:** Analysis of the studies discussed shows that AI models have high accuracy when trained in a supervised setting with labelled data representing both normal and faulty conditions. In addition, to achieve high performance and strong predictive capabilities, the model needs to be trained on a balanced dataset, as it will be more likely to learn the conditions that are more frequently encountered [98]. However, this is difficult to apply in real-world industrial systems such as hydroelectric generator units for two main reasons. Firstly, as the power plant is expected to generate stable electricity for the grid, the HGU is designed to always operate optimally. Therefore, most of the provided measurements are labelled as normal, "healthy" conditions. Even when a fault occurs, the maintenance team usually quickly schedules corrective maintenance to fix or replace the equipment, as the generator's shutdown has many drawbacks, such as economic loss. Having a quick maintenance action limits the amount of faulty data, which leads to a lack of available faulty samples [99]. Moreover, the equipment is designed for industrial purposes and cannot be used for testing, so it is not possible to introduce faults to generate faulty bench tests. Secondly, the process of labelling data is a very time-consuming endeavour, as it necessitates the knowledge of a domain expert to recognize the fault condition and assign the data the appropriate label.
2) **Maintenance cost:** To ensure AI models remain reliable, model calibration strategies must be implemented when performance drops or unexpected or new events occur. The maintenance required can vary depending on the error tolerance, and customer requirements, and can range from simple patching and troubleshooting, model upgrade, to model reconstruction and full model retraining [100].
3) **Architecture selection:** Selecting an appropriate architecture for a specific fault diagnosis task presents significant challenges. The performance of different architectures can vary based on the complexity of the relationships within the data; and an incorrect choice, may result in suboptimal performance.
Moreover, AI models often necessitate meticulous tuning of hyperparameters, such as the learning rate and the number of layers. This process is time-consuming and typically requires considerable expertise. Manual selection of hyperparameters is particularly challenging, and estimating the optimal values can be both labor-intensive and demanding [101].

VII. RESEARCH PROSPECTS AND TRENDS

The monitoring, diagnosis, and prognosis of faults in hydroelectric generators are of paramount importance in ensuring the dependable and efficient operation of hydroelectric

TABLE 2. Abbreviations of feature extraction methods.

Abbreviation	Method	Reference
DWT	Discrete Wavelet Transform	[78]
EWT	Empirical Wavelet Transform	[62]
EK	Expert Knowledge	[66], [67]
FE	Feature Engineering	[75]
FEEMD	Fast Ensemble Empirical Mode Decomposition	[56]
FFT	Fast Fourier Transform	[77], [78]
GA-NLPCANN	Genetic algorithm-Nonlinear Principal Component Analysis Neural Network	[44], [45]
GC	Granular Computing	[49]
GCD	Grey Correlation Degree	[65]
GSO	Gram-Schmidt Orthogonal process	[62]
GST	Generalized S-transform	[58]
HSIC-KNN-FS	Hilbert-Schmidt Independence Criterion with K-Nearest Neighbor Algorithm	[69]
MIC	Maximum Information Coefficient	[65]
MNCWCP	Main Nonlinear-Correlation Working Condition Parameters	[53]
MSWAD	Multi-Source Wide-Area Data	[57]
PCC	Pearson Correlation Coefficient	[65]
PRPD	Phase-Resolved Partial Discharge	[54]
PSR	Phase Space Reconstruction	[60]
RST	Rough Sets Technique	[51]
RWE	Relative Wavelet Energy	[46]
SFVCC	Statistical Fuzzy Vector Chain Code	[52]
SSA	Salp Swarm Algorithm	[60]
TSFRESH	Time Series FeatuRe Extraction based on Scalable Hypothesis tests	[78]
VMD	Variational Mode Decomposition	[60], [71]
WD	Wavelet Decomposition	[47]
WPT	Wavelet Packet Transform	[43]

power plants. This field of research has demonstrated enduring commitment over an extended period, continuously advancing in parallel with technological innovations and data analytics [102], [103]. This article delineates significant research prospects and emerging trends related to the diagnosis and prognosis of hydroelectric generator faults.

A. SENSOR TECHNOLOGY AND IOT INTEGRATION

A central research focus pertains to the development and deployment of advanced sensors capable of providing real-time data on various aspects of hydroelectric generator performance. These aspects encompass critical parameters such as temperature, vibration, pressure, and electrical characteristics. Such data plays a pivotal role in the identification and diagnosis of faults [104]. Furthermore, there is a burgeoning anticipation of increased integration of Internet of Things (IoT) technology, facilitating remote monitoring. This integration affords immediate oversight of hydroelectric generators and their associated systems, thereby streamlining responses to emergent issues. The advent of the Industry 4.0 paradigm has opened doors to IoT technologies, enabling elevated levels of automation and productivity [105].

B. DATA ANALYTICS AND MACHINE LEARNING

The significance of data analytics and machine learning algorithms continues to ascend in the domain of fault diagnosis and prognosis. Researchers are dedicated to crafting predictive models adept at scrutinizing extensive datasets to reveal patterns and anomalies indicative of faults. These models significantly contribute to early fault detection and prediction. In conjunction with the introduction of the Industry 4.0 concept, there is heightened interest in artificial intelligence-based fault analysis, which has garnered community involvement in the development of intelligent fault diagnosis and prognosis (IFDP) models for rotating machinery [106].

C. CONDITION-BASED MAINTENANCE (CBM) AND PREDICTIVE MAINTENANCE

CBM is poised to be steered by technological advancements, data analytics, and an unwavering commitment to optimize the reliability and performance of these critical power generation assets. The amalgamation of intelligent diagnosis systems, leveraging the collective knowledge and expertise of multiple specialists, promises to transcend the capabilities of individual experts. This system facilitates the expeditious

TABLE 3. Abbreviations of AI-based methods.

Abbreviation	Method	Reference
AMGWO	Adaptive Mutation Grey Wolf Optimizer	[60]
ANFIS	Adaptive NeuroFuzzy Interference System	[74]
AE	Autoencoder	[75]
BN	Bayesian Network	[61]
BNN	Bayesian Neural Network	[44], [45], [61]
CVAE	Convolutional Variational AutoEncoder	[66], [67]
DNNLR	Deep Neural Network with Logistic Regression	[75]
DTF	Decision Tree Forest	[78]
ES	Expert System	[50], [61]
GPR	Gaussian Process Regression	[65]
IF	Isolation Forest	[64]
KELM	Kernel Extreme Learning Machine	[60], [56], [62]
KNN	K-Nearest Neighbors	[78]
LGPCA-RF	Local and Global Principal Component Analysis and Random Forest	[57]
LR	Logistic Regression	[78]
LSTM	Long Short-Term Memory Network	[73], [75]
LS-SVM	Least Squares Support Vector Machines	[47]
MD	Mahalanobis Distance	[65]
MFNN	Multilayer Feedforward Neural Network	[54]
MLP	Multi-Layer Perceptron	[78]
MOSSA	Multi-Objective Salp Swarm Algorithm	[62]
MWPCA	Moving Window Principal Component Analysis	[63], [70]
MWPCA-BN	Moving Window Principal Component Analysis and Bayesian Network	[68]
NN	Neural Network	[48], [49]
PCA	Principal Component Analysis	[55]
QPSO-SVM	Quantum Particle Swarm Optimization-Support Vector Machine	[58]
RDM-WNN	Residual Diagnosis Model-Wavelet Neutral Network	[53]
RBFNN	Radial Basis Function Neural Network	[46]
RRCF	Robust Random Cut Forest	[69]
SOFM	Self-Organized Mapping Neural Network	[43]
SSA	Salp Swarm Algorithm	[56]
SVM	Support Vector Machine	[51], [52], [71], [78]
VAE	Variational AutoEncoder	[76], [77]
WFST	Weighted Fuzzy Set Theory	[53]
1-D CNN-GRU	1-D Convolutional Neural Network and Gated Recurrent Unit	[59]

and reliable diagnosis of multiple faults, handling complex processes and abrupt abnormal faults. Predictive maintenance within the framework of an intelligent-control-maintenance-management system (ICMMS) effectively harnesses a trove of data from control, maintenance, and technical management domains to execute timely and precise maintenance. This article introduces the ICMMS platform for hydroelectric generating units, with special emphasis on its maintenance functions [107], [108]. A cognitive mechanism has been formulated and empirically examined, exhibiting the ability to simultaneously track alterations in the dataset and the performance of predictive models. This mechanism perpetually refines the predictive models. Consequently, the method introduced herein can serve as an augmentative component within the decision support system of manufacturing facilities that operate injection molding machines. Its primary aim is to mitigate the occurrence of production failures and reduce machine downtime [109].

VIII. CONCLUSION

In conclusion, this review highlights the dynamic nature of research in fault monitoring, diagnosing, and prognosis (MDP) in hydroelectric generators. Nowadays, one can observe that researchers are increasingly utilizing advanced AI techniques to enhance and strengthen these processes, aiming to achieve objectives such as fault diagnosis, RUL estimation, and degradation assessment more effectively than traditional non-AI-based methods. By harnessing the potential of AI, the research community can leverage data-driven approaches and machine learning algorithms to improve the reliability and accuracy of fault MDP in hydroelectric generator systems.

Furthermore, this ongoing evolution underscores the significance of interdisciplinary collaborations. The successful development and implementation of advanced fault MDP systems require expertise from diverse fields, including engineering and data science. By integrating knowledge and skills from these disciplines, researchers can effectively address the complex challenges associated with hydroelectric generator systems, ensuring their sustained reliability and security in the face of advancing technologies and emerging threats.

In summary, the integration of advanced AI techniques in fault monitoring, diagnosing, and prognosis in hydroelectric generators holds great promise for enhancing the efficiency, reliability, and sustainability of hydroelectric power generation. Accurate AI-based diagnosis has since become a strategic goal to supply reliable energy to meet the growing demand for the electrification and transportation, set as global goal to fight against Green house gaz emission. This review underscores the importance of interdisciplinary collaborations and emphasizes the continued need for ongoing research and innovation to meet the evolving demands of the utilities.

APPENDIX. ACRONYMES

See Tables 2 and 3.

REFERENCES

[1] S. Nandi, H. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Apr. 2005.

[2] B. M. Babić, S. D. Milić, and A. Ž. Rakić, "Fault detection algorithm used in a magnetic monitoring system of the hydrogenerator," *IET Electr. Power Appl.*, vol. 11, no. 1, pp. 63–71, Jan. 2017.

[3] I. Sadeghi, H. Ehya, J. Faiz, and A. A. S. Akmal, "Online condition monitoring of large synchronous generator under short circuit fault—A review," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Feb. 2018, pp. 1843–1848.

[4] H. C. Dirani, A. Merkhouf, B. Kedjar, A.-M. Giroux, and K. Al-Haddad, "Rotor interturn short circuit impact on large hydrogenerator magnetic quantities," *IEEE Trans. Ind. Appl.*, vol. 54, no. 4, pp. 3702–3711, Jul. 2018.

[5] *IEEE Guide for Online Monitoring of Large Synchronous Generators (10 MVA and Above)*, Standard 1129-2014, 2014, p. 62.

[6] I. Zamudio-Ramirez, R. A. Osornio-Rios, J. A. Antonino-Daviu, H. Razik, and R. d. J. Romero-Troncoso, "Magnetic flux analysis for the condition monitoring of electric machines: A review," *IEEE Trans. Ind. Informat.*, vol. 18, no. 5, pp. 2895–2908, May 2022.

[7] M. Riera-Guasp, J. A. Antonino-Daviu, and G.-A. Capolino, "Advances in electrical machine, power electronic, and drive condition monitoring and fault detection: State of the art," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1746–1759, Mar. 2015.

[8] N. Djagarov, G. Enchev, Z. Grozdev, and J. Djagarova, "Overview of diagnostic methods for rotating electrical machines," in *Proc. 7th Int. Conf. Energy Efficiency Agricult. Eng. (EE&AE)*, Nov. 2020, pp. 1–8.

[9] B. Kedjar, A. Merkhouf, and K. Al-Haddad, "Large synchronous machines diagnosis based on air-gap and stray fluxes—An overview," in *Proc. Int. Conf. Electr. Mach. (ICEM)*, Aug. 2020, pp. 1384–1389.

[10] G. C. Brito Junior, R. D. Machado, and A. Chaves Neto, "Using simplified models to assist fault detection and diagnosis in large hydrogenerators," *Int. J. Rotating Machinery*, vol. 2017, pp. 1–18, Jan. 2017.

[11] H. Ehya, A. Nysveen, and R. Nilssen, "A practical approach for static eccentricity fault diagnosis of hydro-generators," in *Proc. Int. Conf. Electr. Mach. (ICEM)*, vol. 1, Aug. 2020, pp. 2569–2574.

[12] H. Ehya, A. Nysveen, R. Nilssen, and Y. Liu, "Static and dynamic eccentricity fault diagnosis of large salient pole synchronous generators by means of external magnetic field," *IET Electr. Power Appl.*, vol. 15, no. 7, pp. 890–902, Jul. 2021.

[13] M. Lévesque, N. Amyot, C. Hudon, M. BÃélec, and O. Blancke, "Improvement of a hydrogenerator prognostic model by using partial discharge measurement analysis," in *Proc. Annu. Conf. PHM Soc.*, 2017, p. 9.

[14] N. Amyot, C. Hudon, M. Levesque, M. Bélec, F. Brabant, and C. St-Louis, "Development of a hydrogenerator prognosis approach," in *Proc. 45th Int. Conf. Large High Voltage Electr. Syst.*, Aug. 24, 2014, Paper SC A1–202.

[15] O. Blancke, A. Combette, N. Amyot, D. Komljenovic, M. Lévesque, C. Hudon, A. Tahan, and N. Zerhouni, "A predictive maintenance approach for complex equipment based on a failure mechanisms propagation model," in *Proc. Annu. Conf. PHM Soc.*, 2019, p. 15.

[16] N. Amyot, C. Hudon, M. Lévesque, M. Bélec, F. Brabant, and F.-X. Frenette, "A new prognostic approach for hydro-generator stator windings," in *Proc. Annu. Conf. PHM Soc.*, 2013, p. 6.

[17] J. Penman, "Feasibility of using unsupervised learning, artificial neural networks for the condition monitoring of electrical machines," *IEE Proc.-Electr. Power Appl.*, vol. 141, no. 6, p. 317, 1994.

[18] S. Shamshirband, T. Rabczuk, and K.-W. Chau, "A survey of deep learning techniques: Application in wind and solar energy resources," *IEEE Access*, vol. 7, pp. 164650–164666, 2019.

[19] S. Zhang, "Artificial intelligence in electric machine drives: Advances and trends," 2021, *arXiv:2110.05403*.

[20] Y. Xu et al., "Artificial intelligence: A powerful paradigm for scientific research," *Innovation*, vol. 2, no. 4, 2021, Art. no. 100179.

[21] R. Siraskar, S. Kumar, S. Patil, A. Bongale, and K. Kotecha, "Reinforcement learning for predictive maintenance: A systematic technical review," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 12885–12947, Nov. 2023.

[22] I. H. Sarker, "AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems," *Social Netw. Comput. Sci.*, vol. 3, no. 2, p. 158, Mar. 2022.

[23] S. C. Tan, A. V. Y. Lee, and M. Lee, "A systematic review of artificial intelligence techniques for collaborative learning over the past two decades," *Comput. Educ., Artif. Intell.*, vol. 3, Jan. 2022, Art. no. 100097.

[24] S. Varshney, R. Jigyasu, A. Sharma, and L. Mathew, "Review of various artificial intelligence techniques and its applications," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 594, no. 1, Aug. 2019, Art. no. 012023.

[25] C. F. Tan, L. S. Wahidin, S. N. Khalil, N. Tamaldin, J. Hu, and G. W. M. Rauterberg, "The application of expert system: A review of research and applications," *ARPN J. Eng. Appl. Sci.*, vol. 11, no. 4, 2016.

[26] C. A. Protopapas, S. D. Kaminaris, A. V. Machias, and B. C. Papadias, "An expert system for fault repairing and maintenance of electric machines," *IEEE Trans. Energy Convers.*, vol. 5, no. 1, pp. 79–83, Mar. 1990.

[27] V. Kontargyri, C. Tsirekis, S. Kabanaros, A. Sakarellos, A. Moronis, N. Kolliopoulos, S. Kaminaris, H. Laboratory, and I. P. Str, "An expert system for fault diagnosis, repairing and maintenance of electrical machines," in *Proc. 6th WSEAS Int. Conf. Appl. Electr. Eng.*, 2007, pp. 755–758.

[28] B. Mahesh, "Machine learning algorithms—A review," *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, 2020.

[29] K. Kudelina, T. Vaimann, B. Asad, A. Rassõlkin, A. Kallaste, and G. Demidova, "Trends and challenges in intelligent condition monitoring of electrical machines using machine learning," *Appl. Sci.*, vol. 11, no. 6, p. 2761, Mar. 2021.

[30] A. D. Dongare, R. R. Kharde, and A. D. Kachare, "Introduction to artificial neural network," *Int. J. Eng. Innov. Technol.*, vol. 2, no. 1, 2012.

[31] R. Dastres and M. Soori, "Artificial neural network systems," *Int. J. Imag. Robot.*, vol. 21, no. 2, pp. 13–25, 2021.

[32] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, Nov. 2018, Art. no. e00938.

[33] J.-S.-R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 3, pp. 665–685, May 1993.

[34] D. Karaboga and E. Kaya, "Adaptive network based fuzzy inference system (ANFIS) training approaches: A comprehensive survey," *Artif. Intell. Rev.*, vol. 52, no. 4, pp. 2263–2293, Dec. 2019.

[35] J. Zhang, W. Ma, and L. Ma, "A fault diagnosis method based on ANFIS and bearing fault diagnosis," in *Proc. Int. Conf. Inf. Sci., Electron. Electr. Eng.*, vol. 2, Apr. 2014, pp. 1274–1278.

[36] Y. Lei, Z. He, Y. Zi, and Q. Hu, ''Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs,'' *Mech. Syst. Signal Process.*, vol. 21, no. 5, pp. 2280–2294, Jul. 2007.

[37] S. Hochreiter and J. Schmidhuber, ''Long short-term memory,'' *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[38] D. Svozil, V. Kvasnicka, and J. Pospichal, ''Introduction to multi-layer feed-forward neural networks,'' *Chemometric Intell. Lab. Syst.*, vol. 39, no. 1, pp. 43–62, Nov. 1997.

[39] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou, ''Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion,'' *J. Mach. Learn. Res.*, vol. 11, no. 12, 2010.

[40] C. Cortes and V. Vapnik, ''Support-vector networks,'' *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[41] W. Fengqi and G. Meng, ''Compound rub malfunctions feature extraction based on full-spectrum cascade analysis and SVM,'' *Mech. Syst. Signal Process.*, vol. 20, no. 8, pp. 2007–2021, Nov. 2006.

[42] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, ''Random forest: A classification and regression tool for compound classification and QSAR modeling,'' *J. Chem. Inf. Comput. Sci.*, vol. 43, no. 6, pp. 1947–1958, Nov. 2003.

[43] D. Sun, T.-S. Huang, G. Li, F.-X. Sun, and B.-S. Qing, ''The fault tendency analysis of hydro-generator based on WNN,'' in *Proc. Int. Conf. Mach. Learn. Cybern.*, Jul. 2004, pp. 3090–3094.

[44] J. Qiaoling, Q. Weimin, and C. Weiyou, ''Fault diagnosis of hydro-generator unit via GA-nonlinear principal component analysis neural network,'' in *Proc. Chin. Control Conf.*, Jul. 2006, pp. 468–472.

[45] Q.-L. Ji, W.-M. Qi, W.-Y. Cai, and Y.-C. Cheng, ''Study of fault diagnosis of hydro-generator unit via Ga nonlinear principal component analysis neural network and Bayesian neural networks,'' in *Proc. Int. Conf. Mach. Learn. Cybern.*, 2006, pp. 319–324.

[46] Z. Liu, J. Zhou, M. Zou, Y. Zhang, and L. Zhan, ''A new method for intelligent fault diagnosis of hydroelectric generating unit,'' in *Proc. IEEE Int. Conf. Control Autom.*, May 2007, pp. 1638–1642.

[47] Z. Min, Z. Jianzhong, Z. Yongchuan, and L. Zhong, ''Fault diagnosis of hydroturbine generating units based on least squares support vector machines,'' in *Proc. IEEE Int. Conf. Control Autom.*, May 2007, pp. 152–156.

[48] J. Li and H. Wang, ''A study on fault diagnosis of hydroelectric generator based on D-S evidence theory,'' in *Proc. Int. Conf. Electr. Mach. Syst.*, 2008.

[49] K. Xie, J. Xie, L. Du, and X. Xu, ''Granular computing and neural network integrated algorithm applied in fault diagnosis,'' in *Proc. 6th Int. Conf. Fuzzy Syst. Knowl. Discovery*, Aug. 2009, pp. 188–191.

[50] E. J. Amaya and A. J. Alvares, ''SIMPREBAL: An expert system for real-time fault diagnosis of hydrogenerators machinery,'' in *Proc. IEEE 15th Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2010, pp. 1–8.

[51] X. Zhang, J. Zhou, J. Guo, Q. Zou, and Z. Huang, ''Vibrant fault diagnosis for hydroelectric generator units with a new combination of rough sets and support vector machine,'' *Expert Syst. Appl.*, vol. 39, no. 3, pp. 2621–2628, Feb. 2012.

[52] H. Xiao, J. Zhou, J. Xiao, W. Fu, X. Xia, and W. Zhang, ''Identification of vibration–speed curve for hydroelectric generator unit using statistical fuzzy vector chain code and support vector machine,'' *Inst. Mech. Eng, O, J. Risk Rel.*, vol. 228, no. 3, pp. 291–300, Jun. 2014.

[53] W. Zhu, J. Zhou, C. Li, and X. Xue, ''Residual diagnosis model based on wavelet neutral network and its application to hydroelectric generator unit,'' in *Proc. IEEE 12th Int. Conf. Netw., Sens. Control*, Apr. 2015, pp. 23–27.

[54] R. M. S. D. Oliveira, R. C. F. Araújo, F. J. B. Barros, A. P. Segundo, R. F. Zampolo, W. Fonseca, V. Dmitriev, and F. S. Brasil, ''A system based on artificial neural networks for automatic classification of hydro-generator stator windings partial discharges,'' *J. Microw., Optoelectron. Electromagn. Appl.*, vol. 16, no. 3, pp. 628–645, Sep. 2017.

[55] F. Chu, ''An improved PCA algorithm for anomaly detection of hydropower units,'' in *Proc. IEEE 2nd Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Apr. 2017, pp. 494–498.

[56] Y. Shan, J. Zhou, W. Jiang, J. Liu, Y. Xu, and Y. Zhao, ''Vibration tendency prediction of hydroelectric generator unit based on fast ensemble empirical mode decomposition and kernel extreme learning machine with parameters optimization,'' in *Proc. 11th Int. Symp. Comput. Intell. Design (ISCID)*, Dec. 2018, pp. 287–290.

[57] J. Dang, P. Fan, R. Jia, J. Wei, and J. Li, ''A generator unit fault classification approach based on multi-source wide-area feature extraction,'' in *Proc. IEEE 8th Int. Conf. Adv. Power Syst. Autom. Protection (APAP)*, Oct. 2019, pp. 888–892.

[58] J. Luo, X. Wang, and Y. Xu, ''Vibration fault diagnosis for hydroelectric generating unit based on generalized S-transform and QPSO-SVM,'' in *Proc. IEEE Sustain. Power Energy Conf. (iSPEC)*, Nov. 2019, pp. 2133–2137.

[59] G.-P. Liao, W. Gao, G.-J. Yang, and M.-F. Guo, ''Hydroelectric generating unit fault diagnosis using 1-D convolutional neural network and gated recurrent unit in small hydro,'' *IEEE Sensors J.*, vol. 19, no. 20, pp. 9352–9363, Oct. 2019.

[60] W. Fu, K. Wang, J. Tan, and K. Shao, ''Vibration tendency prediction approach for hydropower generator fused with multiscale dominant ingredient chaotic analysis, adaptive mutation grey wolf optimizer, and KELM,'' *Complexity*, vol. 2020, pp. 1–20, Feb. 2020.

[61] B. Xu, H. Li, W. Pang, D. Chen, Y. Tian, X. Lei, X. Gao, C. Wu, and E. Patelli, ''Bayesian network approach to fault diagnosis of a hydroelectric generation system,'' *Energy Sci. Eng.*, vol. 7, no. 5, pp. 1669–1677, Oct. 2019.

[62] K.-B. Zhou, J.-Y. Zhang, Y. Shan, M.-F. Ge, Z.-Y. Ge, and G.-N. Cao, ''A hybrid multi-objective optimization model for vibration tendency prediction of hydropower generators,'' *Sensors*, vol. 19, no. 9, p. 2055, May 2019.

[63] M. Michalski, ''Applying moving window principal component analysis (MWPCA) for fault detection in hydrogenerator,'' in *Proc. 30th Eur. Saf. Rel. Conf. 15th Probabilistic Saf. Assessment Manage. Conf.*, 2020, pp. 381–386.

[64] Y. Hara, Y. Fukuyama, K. Murakami, T. Iizaka, and T. Matsui, ''Fault detection of hydroelectric generators using isolation forest,'' in *Proc. 59th Annu. Conf. Soc. Instrum. Control Eng. Jpn. (SICE)*, Sep. 2020, pp. 864–869.

[65] H. Huang, A. Qin, H. Mao, J. Fu, Z. Huang, Y. Yang, X. Li, and H. Huang, ''The prediction method on the early failure of hydropower units based on Gaussian process regression driven by monitoring data,'' *Appl. Sci.*, vol. 11, no. 1, p. 153, Dec. 2020.

[66] R. Zemouri, M. Lévesque, N. Amyot, C. Hudon, O. Kokoko, and S. A. Tahan, ''Deep convolutional variational autoencoder as a 2D-visualization tool for partial discharge source classification in hydrogenerators,'' *IEEE Access*, vol. 8, pp. 5438–5454, 2020.

[67] O. Kokoko, C. Hudon, M. Levesque, N. Amyot, and R. Zemouri, ''Comparison of an automatic classification of partial dischage patterns for large hydrogenerator,'' in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2021, pp. 1–6.

[68] A. H. D. A. Melani, M. A. D. C. Michalski, R. F. da Silva, and G. F. M. de Souza, ''A framework to automate fault detection and diagnosis based on moving window principal component analysis and Bayesian network,'' *Rel. Eng. Syst. Saf.*, vol. 215, Nov. 2021, Art. no. 107837.

[69] Y. Hara, Y. Fukuyama, K. Arai, Y. Shimasaki, Y. Osada, K. Murakami, T. Iizaka, and T. Matsui, ''Fault detection of hydroelectric generators by robust random cut forest with feature selection using Hilbert–Schmidt independence criterion,'' in *Proc. IEEE Int. Conf. Smart Internet Things (SmartIoT)*, Aug. 2021, pp. 136–143.

[70] M. A. de Carvalho Michalski, R. Favarao da Silva, A. H. de Andrade Melani, and G. Francisco Martha de Souza, ''Applying principal component analysis for multi-parameter failure prognosis and determination of remaining useful life,'' in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, May 2021, pp. 1–6.

[71] X. Tang, B. Hu, and H. Wen, ''Fault diagnosis of hydraulic generator bearing by VMD-based feature extraction and classification,'' *Iranian J. Sci. Technol., Trans. Electr. Eng.*, vol. 45, no. 4, pp. 1227–1237, Dec. 2021.

[72] R. Zemouri, S. Bernier, O. Kokoko, and A. Merkhouf, ''Hydro-generators fault diagnosis with short-time-wavelet-entropy and variational autoencoder,'' *IOP Conf. Series: Mater. Sci. Eng.*, vol. 1207, no. 1, Nov. 2021, Art. no. 012009.

[73] J. Huang, B. Guo, S. Dian, X. Li, and W. Lai, ''Fault diagnosis of hydropower generator based on LSTM- weight,'' in *Proc. 7th Int. Conf. Autom., Control Robot. Eng. (CACRE)*, Jul. 2022, pp. 369–373.

[74] S. P. Dahal, M. Dahal, and B. Silwal, ''Historical data based monitoring of hydro generator using machine learning,'' in *Proc. 14th Int. Conf. Softw., Knowl., Inf. Manage. Appl. (SKIMA)*, Dec. 2022, pp. 61–66.

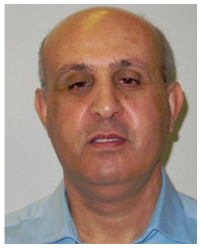
[75] V. Velasquez and W. Flores, "Machine learning approach for predictive maintenance in hydroelectric power plants," in *Proc. IEEE Biennial Congr. Argentina (ARGENCON)*, Sep. 2022, pp. 1–6.
[76] R. Ibrahim, R. Zemouri, A. Tahan, F. Lafleur, B. Kedjar, A. Merkhouf, and K. Al-Haddad, "Anomaly detection for large hydrogenerators using the variational autoencoder based on vibration signals," in *Proc. Int. Conf. Electr. Mach. (ICEM)*, Sep. 2022, pp. 1609–1615.
[77] H. Bechara, R. Zemouri, A. Tahan, B. Kedjar, A. Merkhouf, and K. Al-Haddad, "Non-invasive anomaly diagnosis for hydro electrical generators rotor inter-turn short-circuit detection using stray flux and the VAE," in *Proc. Int. Conf. Electr. Mach. (ICEM)*, Sep. 2022, pp. 1855–1861.
[78] H. Ehya, T. N. Skreien, and A. Nysveen, "Intelligent data-driven diagnosis of incipient interturn short circuit fault in field winding of salient pole synchronous generators," *IEEE Trans. Ind. Informat.*, vol. 18, no. 5, pp. 3286–3294, May 2022.
[79] R. Zemouri, R. Ibrahim, and A. Tahan, "Hydrogenerator early fault detection: Sparse dictionary learning jointly with the variational autoencoder," *Eng. Appl. Artif. Intell.*, vol. 120, Apr. 2023, Art. no. 105859.
[80] R. Zemouri and M. Lévesque, "Ensemble deep learning model for phase-resolved partial discharge diagnosis in hydrogenerators," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 30, no. 5, pp. 2394–2401, Oct. 2023.
[81] R. Ibrahim, R. Zemouri, B. Kedjar, A. Merkhouf, A. Tahan, K. Al-Haddad, and F. Lafleur, "Non-invasive detection of rotor inter-turn short circuit of a hydrogenerator using AI-based variational autoencoder," *IEEE Trans. Ind. Appl.*, vol. 60, no. 1, pp. 28–37, Jan. 2024.
[82] R. Ibrahim, R. Zemouri, B. Kedjar, A. Tahan, A. Merkhouf, and K. Al-Haddad, "Rotor fault diagnosis in a hydrogenerator based on the stator vibration and the variational autoencoder," in *Proc. IEEE Int. Electr. Mach. Drives Conf. (IEMDC)*, May 2023, pp. 1–7.
[83] H. Bechara, R. Zemouri, B. Kedjar, A. Merkhouf, K. Al-Haddad, and A. Tahan, "Non-invasive detection of rotor inter-turn short circuit in large hydrogenerators by using stray flux measurement combined with convolutional variational autoencoder analysis (CVAE)," *IEEE Trans. Ind. Appl.*, vol. 60, no. 1, pp. 196–205, Jan. 2024.
[84] R. Ibrahim, R. Zemouri, A. Tahan, B. Kedjar, A. Merkhouf, and K. Al-Haddad, "Fault detection based on vibration measurements and variational autoencoder-desirability function," *IEEE Open J. Ind. Appl.*, vol. 5, pp. 106–116, 2024.
[85] S. E. Whang, Y. Roh, H. Song, and J.-G. Lee, "Data collection and quality challenges in deep learning: A data-centric AI perspective," *VLDB J.*, vol. 32, no. 4, pp. 791–813, Jul. 2023.
[86] H. Ehya, A. Nysveen, and T. N. Skreien, "Performance evaluation of signal processing tools used for fault detection of hydrogenerators operating in noisy environments," *IEEE Trans. Ind. Appl.*, vol. 57, no. 4, pp. 3654–3665, Jul. 2021.
[87] F. I. Ihab and C. Xu, *Data Cleaning*. New York, NY, USA: ACM, 2019.
[88] I. F. Ilyas and T. Rekatsinas, "Machine learning and data cleaning: Which serves the other?" *J. Data Inf. Qual.*, vol. 14, no. 3, pp. 1–11, Sep. 2022.
[89] W. Lang, Y. Hu, C. Gong, X. Zhang, H. Xu, and J. Deng, "Artificial intelligence-based technique for fault detection and diagnosis of EV motors: A review," *IEEE Trans. Transport. Electrific.*, vol. 8, no. 1, pp. 384–406, Mar. 2022.
[90] L. Fischer, L. Ehrlinger, V. Geist, R. Ramler, F. Sobiezky, W. Zellinger, D. Brunner, M. Kumar, and B. Moser, "AI system engineering—Key challenges and lessons learned," *Mach. Learn. Knowl. Extraction*, vol. 3, no. 1, pp. 56–83, Dec. 2020.
[91] I. Remadna, L. S. Terrissa, Z. Al Masry, and N. Zerhouni, "RUL prediction using a fusion of attention-based convolutional variational AutoEncoder and ensemble learning classifier," *IEEE Trans. Rel.*, vol. 72, no. 1, pp. 106–124, Mar. 2023.
[92] A. Saxena, J. Celaya, E. Balaban, K. Goebel, B. Saha, S. Saha, and M. Schwabacher, "Metrics for evaluating performance of prognostic techniques," in *Proc. Int. Conf. Prognostics Health Manage.*, Oct. 2008, pp. 1–17.
[93] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *Int. J. Data Mining Knowl. Manage. Process*, vol. 5, no. 2, pp. 1–11, Mar. 2015.
[94] S. Khan, S. Tsutsumi, T. Yairi, and S. Nakasuka, "Robustness of AI-based prognostic and systems health management," *Annu. Rev. Control*, vol. 51, pp. 130–152, Jan. 2021.
[95] K. C. A. Khanzode, "Advantages and disadvantages of artificial intelligence and machine learning: A literature review," *Int. J. Library Inf. Sci.*, vol. 9, no. 1, p. 3, 2020.
[96] Z. Yang, P. Baraldi, and E. Zio, "A multi-branch deep neural network model for failure prognostics based on multimodal data," *J. Manuf. Syst.*, vol. 59, pp. 42–50, Apr. 2021.
[97] M. A. Awadallah and M. M. Morcos, "Application of AI tools in fault diagnosis of electrical machines and drives–An overview," *IEEE Trans. Energy Convers.*, vol. 18, no. 2, pp. 245–251, Jun. 2003.
[98] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mech. Syst. Signal Process.*, vol. 138, Apr. 2020, Art. no. 106587.
[99] L. C. Brito, G. A. Susto, J. N. Brito, and M. A. V. Duarte, "Fault diagnosis using eXplainable AI: A transfer learning-based approach for rotating machinery exploiting augmented synthetic data," *Expert Syst. Appl.*, vol. 232, Dec. 2023, Art. no. 120860.
[100] P.-Y. Chen and P. Das, "AI maintenance: A robustness perspective," *IEEE Comput. Mag.*, vol. 56, no. 2, pp. 48–56, Feb. 2023.
[101] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, "Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review," *IEEE Access*, vol. 7, pp. 122644–122662, 2019.
[102] M. Mostafaei and J. Faiz, "An overview of various faults detection methods in synchronous generators," *IET Electr. Power Appl.*, vol. 15, no. 4, pp. 391–404, Apr. 2021.
[103] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, Jun. 2015.
[104] R. B. de Santis, T. S. Gontijo, and M. A. Costa, "Condition-based maintenance in hydroelectric plants: A systematic literature review," *Proc. Inst. Mech. Eng., O, J. Risk Rel.*, vol. 236, no. 5, pp. 631–646, Oct. 2022.
[105] A. Silva and G. F. M. Souza, "Prognosis smart system AI-based applied to equipment health monitoring in 4.0 industry scenario," in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, May 2021, pp. 1–6.
[106] O. Das, D. Bagci Das, and D. Birant, "Machine learning for fault analysis in rotating machinery: A comprehensive review," *Heliyon*, vol. 9, no. 6, Jun. 2023, Art. no. e17584.
[107] X. T. Liu, F. Z. Feng, and A. W. Si, "Condition based monitoring, diagnosis and maintenance on operating equipments of a hydraulic generator unit," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 15, no. 4, Nov. 2012, Art. no. 042014.
[108] C. Fu, L. Ye, Y. Liu, R. Yu, B. Iung, Y. Cheng, and Y. Zeng, "Predictive maintenance in intelligent-control-maintenance-management system for hydroelectric generating unit," *IEEE Trans. Energy Convers.*, vol. 19, no. 1, pp. 179–186, Mar. 2004.
[109] V. Rousopoulou, A. Nizamis, T. Vafeiadis, D. Ioannidis, and D. Tzovaras, "Predictive maintenance for injection molding machines enabled by cognitive analytics for industry 4.0," *Frontiers Artif. Intell.*, vol. 3, Nov. 2020, Art. no. 578152.

HELENE BECHARA was born in Al-Hazmieh, Lebanon, in 1994. She received the B.Sc. and M.Sc. degrees in electrical engineering from Saint-Joseph University, Lebanon, in 2017 and 2019, respectively, and the Ph.D. degree from École de technologie supérieure (ÉTS), Montreal, Canada, in 2024. Her research interests include monitoring and diagnostic tools for complex electromagnetic systems and particularly hydroelectric generators.

RONY IBRAHIM (Member, IEEE) was born in Lebanon, in 1996. He received the B.Sc. degree in electrical engineering and the M.Sc. degree in renewable energies from Saint-Joseph University, Beirut, Lebanon, in 2019 and 2020, respectively, and the Ph.D. degree from École de technologie supérieure (ÉTS), Montreal, Canada, in 2024. His research interests include monitoring, diagnostic, and prognostic tools for complex electromechanical systems.

RYAD ZEMOURI was born in Algiers, Algeria, in 1974. He received the Engineering degree in electronics and control science from the Université Mouloud Mammeri de Tizi-Ouzou, Algeria, in 1998, and the Master's Diploma and Ph.D. degrees from the Franche-Comt University of Besanon, France, in 2000 and 2003, respectively. Since 2003, he has been an Associate Professor with the Engineering Department of Automatic Control, Electronics, and Computer Science, Conservatoire National des Arts et Métiers, Paris, France. In 2022, he joined the Research Center of Hydro Quebec (CRHQ) where he is currently a Researcher. His research interests include machine learning for safety-critical prognosis, diagnosis, and health management.

BACHIR KEDJAR received the B.Sc. degree in electrotechnical from École Nationale Polytechnique, Algiers, Algeria, in 1984, the M.Sc. degree in electrical engineering from Universite Pierre et Marie Curie, Paris, France, in 1985, and the Ph.D. degree in power electronics from Université des Sciences et Techniques de Lille Flandres-Artois, Lille, France, in 1988. From September 1989 to July 2001, he was a Lecturer with Université de Bejaia, Algeria. He is currently with the Electrical Engineering Department, Ecole de Technologie Supérieure, Montréal, QC, Canada, as a Research Assistant and a Lecturer. His research interests include unity power factor ac-dc converters, power quality, electric machines, drives and interfacing renewable energy sources to the grid in terms of modeling, and simulation and control.



AREZKI MERKHOUF (Senior Member, IEEE) received the B.Sc.Eng. degree from the University of Algeria, in 1990, the M.Sc. degree from the University of Pierre and Marie Curie, France, in 1992, and the Ph.D. degree in electrical engineering from the University of Sherbrooke, Québec, Canada, in 1999. In June 1999, he joined the GE Energy Power Systems, where he was a member of the Research and Development Team, until 2008. In 2008, he joined the Hydro Québec Research Institute (IREQ), where he is currently a Senior Researcher and as a Project Leader on the diagnosis of large electrical machines. He is the author of more than 65 scientific articles. His research interests include generator and motor design and optimization, generator acoustics and vibrations, analytical and numerical simulation of large rotating electrical machines, numerical, and analytical computation of electromagnetic fields.

ANTOINE TAHAN was born in Aleppo, Syria, in 1966. He received the B.Sc., M.Sc., and Ph.D. degrees from the University of Laval, Canada, in 1991, 1992, and 1998, respectively. Since 2004, he has been a Full Professor with the Department of Mechanical Engineering, École de technologie supérieure (ÉTS), Montreal, Québec, Canada, and an Active Member of the Engineering Laboratory for Products, Processes, and Systems (LIPPS). He specializes in dimensional metrology and geometrical characterization, tolerance management, and propagation of uncertainties. His research focuses on the aerospace and energy industries. He is working on the propagation of uncertainties in engineering models, diagnostic, and prognostic tools for complex electromechanical systems. He contributes to the development of probabilistic approaches to detect the state of degradation and damage to improve reliability and design.

KAMAL AL-HADDAD (Life Fellow, IEEE) received the B.Sc. and M.Sc. degrees from the University of Québec ã Trois-Riviéres, Canada, in 1982 and 1984, respectively, and the Ph.D. degree from the Institut National Polytechnique, Toulouse, France, in 1988. Since June 1990, he has been a Professor with the Electrical Engineering Department, École de technologie supérieure (ÉTS), Montreal, Québec, Canada, where he has been the holder of the Senior Canada Research Chair of the Electric Energy Conversion and Power Electronics, since 2002. He is a Consultant and has established a very solid link with many Canadian and international industries working in the fields of power electronics, electric transportation, aeronautics, and telecommunications. His research interests include highly efficient static power converters, harmonics and reactive power control using hybrid filters, voltage level multiplier, resonant and multilevel converters, including the modeling, control, and development of prototypes for various industrial applications in electric traction, renewable energy, power supplies for drives, and telecommunication. He is a member of the Academy of Sciences, a fellow of the Royal Society of Canada, and a fellow member of the Canadian Academy of Engineering. He was a recipient of the 2014 IEEE IES Dr.-Ing. Eugene Mittelmann Achievement Award. He has been elected the IEEE 2022 Division VI Director Elect. He is the IEEE IES President 2016–2017, an Associate Editor of the Transactions on Industrial Informatics, and an IES Distinguished Lecturer.