IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

Analyzing Transformer Insulation Paper Prognostics and Health Management: A Modeling Framework Perspective

ANDREW ADEWUNMI ADEKUNLE[1], ISSOUF FOFANA[1], (Senior Member, IEEE), PATRICK PICHER[2], (Senior Member, IEEE), ESPERANZA MARIELA RODRIGUEZ-CELIS[2], OSCAR H. ARROYO-FERNANDEZ[2*]

[1]Canada Research Chair tier 1, in Aging of Oil-Filled Equipment on High Voltage Lines (ViAHT), University of Quebec at Chicoutimi, Chicoutimi, QC G7H 2B1, Canada
[2]Hydro Quebec's Research Institute, Varennes, QC (Canada)

*Corresponding author: ArroyoFernandez.Oscar@hydroquebec.com

This work is co-sponsored by the Canada Research Chair tier 1, in Aging of Oil-Filled Equipment on High Voltage Lines (ViAHT) under grant number CRC-2021-00453.

ABSTRACT In the era of Industry 4.0, digital transformation has spurred the swift advancement of technologies, including intelligent predictive maintenance scheduling, prognostics and health management. The accurate prediction of remaining useful life plays a crucial role in these technologies as it extends power equipment's safe operational duration and decreases the maintenance costs associated with unforeseen shutdowns. Also, the increased accessibility of data for monitoring system conditions has paved the way for the more immense adoption of machine learning models in prognostics and health management for power transformers. At the moment, with the ever-increasing demand for electricity, there is a corresponding increase in the degradation processes of power transformers. Transformers insulation system and more importantly, the paper insulation happens to be the principal part where the degradation is prominent. Therefore, an accurate prediction of the insulating paper condition through its degree of polymerization is required to guarantee the reliability of power transformers. In this regard, the predictions, reliability, and health monitoring of this power equipment can be actualized by modeling the degradation of transformer insulation paper through several machine learning frameworks. In this view, this review paper has been drafted not only to serve as a guide for researchers interested in the fields of transformer insulation system fault prognosis but also to offer insights into potential research directions as existing literature in modeling and evaluating transformer paper insulation is presented.

INDEX TERMS Insulating paper, ageing, degree of polymerization, prognostics and health management, machine learning model.

I. INTRODUCTION

Over time, power transformers experience degradation in their insulation system which to a large extent causes tragic failures which limit the proper operations of power networks [1], [2]. The degradation of solid insulation among other transformer-insulating components remains the primary cause behind the ultimate failure of power transformers. As a result, assessing the state of paper insulation has become an established measure of the power transformer's health condition [3]. Also, according to IEEE, the lifetime models of solid insulation can be used for reliability assessment and health monitoring for power transformers [4]. The transformer solid insulation which is known as the paper insulation consists of extended chains of glucose rings forming the cellulose polymer molecule. The collective length of these chains is computed as the degree of polymerization [5]. Throughout the power transformer's operational life, the degree of polymerization diminishes due to some ageing mechanisms like oxidation, hydrolysis, and pyrolysis [6]. Measuring the degree of polymerization directly from an insulating paper sample is a practicable approach. However, when it comes to a power transformer in active operation, adopting this practice involves undesirable disconnection and invasive handling of the unit [7]. Hence, several techniques discussed later in this study have been employed as an indirect measure due to the established correlations between degree of polymerization (DP) and these parameters.

The effective operation, performance, and reliability of the power equipment can be enhanced by diagnostics and prognostics approaches as they give detailed information on the maintenance and replacement of any component of the power transformer to prevent any critical deprecatory state of the system. Furthermore, corrective and preventive maintenance are the two predominant maintenance schemes used since the 1990s [8]. Corrective maintenance is only performed after the equipment has experienced a breakdown, which results in prolonged repair durations and other penalties and expenses related to the equipment breakdown. This type of maintenance is often referred to as failure-driven maintenance. Preventive maintenance is performed on a routine basis as experts make these maintenance decisions based on their past knowledge of the equipment manufacturers and past breakdown and failure data. Nevertheless, making an accurate maintenance schedule in advance proves to be demanding. This makes preventive maintenance to be inadequate and gradually becoming outdated. In recent times, predictive maintenance has made use of predictive tools to ascertain the optimal timing when maintenance actions are required. This maintenance approach also referred to as condition-based maintenance can mitigate unexpected downtime, reduce maintenance expenses, and increase the equipment's lifetime [9], [10]. Furthermore, predictive maintenance utilizes non-invasive testing methods like acoustics dissolved gas analysis (DGA), electrical tests, infrared analysis, thermodynamics, and vibration analysis to observe and estimate the trend in the performance of the equipment. Data gathering, fault identification, diagnostics, and prognostics are the basic steps to execute predictive maintenance. According to equipment degradation data, classification, regression, and survival models are three major modeling schemes for predictive maintenance. The classification model aims to forecast if a failure will take place within a specified time frame. The regression model aims to forecast when the equipment is expected to fail by modeling the trajectory of the deterioration path. Finally, elemental concept of survival model is to address how the risk of failure evolves over time [9]. Figure 1 shows the classification of equipment failure [11] and the fault tree analysis of the transformer that connects the system-level transformer fault with low-level transformer failure is given in Figure 2 [12].

Now, the emergence of machine learning (ML) presents a promising technique for monitoring and assessing the condition of transformer insulation degradation through its capacity to understand the role of predictive maintenance [13]. This learning has helped to mitigate the errors in communication and sensor malfunction by addressing the missing data obtained from the sensor, which ensures a reliable predictive model [14]. Furthermore, ML models provide an effective technique for detecting, diagnosing and predicting transformer insulation status than the traditional approach [15]. Therefore, utilizing ML techniques for predicting the DP correlated parameters will offer a more cost-effective and readily accessible method for monitoring the health condition of transformers' paper insulation.

In this regard, this review seeks to provide researchers and engineers with insights into the vast potential, applications, and challenges associated with transformer insulation paper prognosis approaches. Furthermore, our goal is to aid researchers in obtaining a comprehensive understanding of the current and future applications, as well as the challenges, in the field of fault prognosis from a machine perspective. We anticipate that the application of machine learning in fault prognosis will significantly transform maintenance practices of power transformer insulation systems in the coming years.

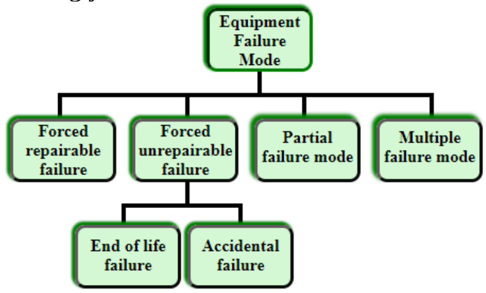


FIGURE 1. **Equipment failure classification [11].**

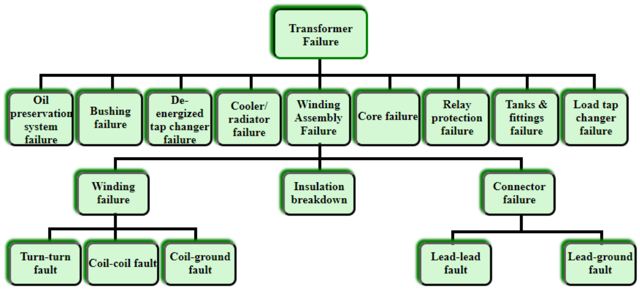


FIGURE 2. **Fault tree analysis for transformers [12].**

II. PAPER INSULATION

Cellulose pulp served in the early days of electrical engineering as an insulating paper and according to [16], [17], cellulose insulation paper represents approximately 10% of the weight of the transformer. It plays an important role as an insulator used for wrapping around copper windings and as a mechanical/structural support in oil-immersed transformers [18]. Also, spacers or pressplates, and pressrings, which are components of power transformers make use of pressboard or paper board [19]. Cellulose material is still utilized in power transformers because of its availability, low price outstanding insulation performance when dried, and acceptable mechanical behaviour at high temperatures relative to synthetic material, making the electrical power industry worldwide process several million tons of cellulose pulp into insulating materials. However, the paper's chemical process complexity when deteriorating has been a major setback limiting its technical utilization [18]. The common raw material for insulating paper is electrical grade softwood Kraft pulp containing about 85 % cellulose, 5 % lignin, and 10 %

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3391823

hemicelluloses. Bleaching of the paper is circumvented to increase the content of lignin in the paper, which helps improve its thermal strength and integrity [20]. Figure 3 represents the molecular structure of cellulose [7], [21]. Typically, solid insulation comprises two types of cellulose, which are thermally upgraded paper (TUP) and Kraft paper [22]. TUP, which is either chemically modified or incorporated with additives, has been recently utilized in power transformers due to their excellent improvement of thermal stability of insulating paper. However, chemically modified TUP requires extensive quantities of toxic and hazardous reactants, needs a distinct industrial process for the pulping stage, and there is a decrease in the mechanical performance of the insulating paper due to the replacement of the hydroxyl groups. Furthermore, in addition to excellent thermal stability, additives incorporated in TUP are cheap and have seamless integration into current industrial processes. However, it has the potential of discharging corrosive ammonia [18].

As per the GB/T29305-2012 standard [23] for measuring insulating cellulosic materials average DP for new and aged paper, the degree of polymerization of insulating paper is assessed through viscosity measurement. This involves utilizing positive hexane and copper ethylenediamine as the extractant and solvent respectively. The paper is then fragmented into small particles, dissolved in hexane, and completely defatted in a Sow extraction device. After drying in a vacuum oven, the mass of the sample is measured. Subsequently, the dried sample is introduced into the copper ethylenediamine solvent, and stirred until fully dissolved using a magnetic stirrer. The degree of polymerization is then determined by measuring the outflow time of the prepared solution with a Ubbelohde viscometer at a temperature of 20°C [7], [24]. The DP of new insulation paper is about 1000 – 1200 where its lifespan is approximately 180 000 h while the insulation paper DP has its least value (200) at the hottest spot of the transformer, which is considered its end-of-useful life [4], [25]. At this stage, the insulation paper becomes weak and brittle, which diminishes its insulation integrity and a loss of its capacity to sustain additional stresses [26]. Paper with 200 DP is also considered to have lost 70% of its tensile strength [27]. However, a fundamental prerequisite for insulating paper pulp is to assume a concentration of extremely low ions of transition and primary group metals enabling them to possess low conductivity and ensure reliable performance when exposed to very high temperatures [18].

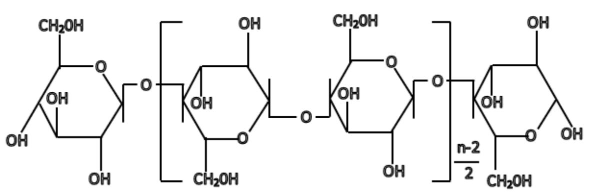


FIGURE 3. **Cellulose molecular structure [7, 21].**

A. MATHEMATICAL MODEL CONCEPT

Cellulose degradation kinetics have been developed for indirect evaluation of paper insulation, as the paper insulation cannot be easily accessed like the insulating liquid. The DP of insulation paper in power transformers for a given time can be estimated by employing the Arrhenius equation proposed by Emsley and Stephen model [28], which is also known as the zero-order kinetic model [29].

$$\frac{1}{DP_t} - \frac{1}{DP_0} = Ae^{-\frac{E_a}{RT}}t \quad (1)$$

Where DP_0 and DP_t are the original DP for the new transformer and DP value at time *t*, respectively, *R* is the gas constant in J/mol/K, E_a is the activation energy in Jmol^{-1}, T is the hot-spot temperature in Kelvin, A in $\mathrm{hr}^{-1}$is the pre-exponential factor that depends on the chemical environment, and t is the time in hours [27], [30]. *A* and E_a can be known from the plot of the logarithms of *K* against the reciprocal of *T*. where *K* is the rate of reaction given as [28], [31] given as:

$$K = Ae^{-\frac{E_a}{RT}} \quad (2)$$

Also, a recursive form of equation (1) given as equation (3) was employed in [25] to improve the accuracy of paper deterioration estimation.

$$\frac{1}{DP_n} - \frac{1}{DP_{(n-1)}} = A_{(n-1)}e^{-\frac{E_{a(n-1)}}{RT_{(n-1)}}}.\left[t_n - t_{(n-1)}\right] \quad (3)$$

Where n is the iteration stage.

The author in [27] utilized the pseudo-zero-order kinetic equation in equation (1) to model the DP as given in equation (4) considering the temperature, moisture content and oxygen level of the insulating paper. The DP_n is DP after the ageing period t_n, DP_{n-1} is the paper DP at the end of the last interval, E_a in J/mol is the activation energy, *R* in $\mathrm{Jmol}^{-1}\mathrm{K}^{-1}$ is the ideal gas constant, t_n is time-period, and *T* is the temperature in Kelvin. The value for *A* is obtained from the oxygen level of the insulating liquid and the moisture content of the insulating paper as seen in [7].

$$DP_n = \frac{1}{A \times t_n \times e^{\left(\frac{-E_a}{RT}\right)} + \frac{1}{DP_{n-1}}} \quad (4)$$

Furthermore, the Arrhenius equation in equation (1) was disintegrated in [32] to account for processes that involve oxidation, pyrolysis, and hydrolysis.

$$\frac{1}{DP(t)} - \frac{1}{DP(t_o)} = \sum_{t_0}^{t} k(t), \quad (5)$$

Where,

$$k(t) = A_{oxi}(t)e^{-\frac{E_{a,oxi}}{(\theta_{HS}(t)+273)R}} + A_{pyr}(t)e^{-\frac{E_{a,pyr}}{(\theta_{HS}(t)+273)R}} + A_{hyd}(t)e^{-\frac{E_{a,hyd}}{(\theta_{HS}(t)+273)R}} \quad (6)$$

The subscripts oxi, pyr, and hyd represent oxidation, pyrolysis, and hydrolysis respectively. $k(t)$ is the rate of degradation. Also, an expression in the form of a quadratic equation that relates the value of A_{hyd}, range of oxygen level and insulating paper moisture for both thermally upgraded and non-thermally upgraded paper was proposed in [33], [34]. In [35], the authors estimated the degradation of insulating paper by utilizing moisture content and temperature as the ageing parameter. The influence of oxygen was not considered as a larger number of oil-immersed transformers concurrently manufactured are known to be a system devoid of air (air free).

TUP (Nomex paper) for dry insulation also employs the first-order reaction law as seen in equation (7) to obtain its degree of polymerization [36], [37].

$$\frac{1}{DP_t} - \frac{1}{DP_0} = \frac{k_1}{k_2}[1 - e^{-k_2 t}] \quad (7)$$

Where DP_0 and DP_t are the Nomex paper DP at initial time and time t respectively, t is the ageing time, k_1 and k_2 are the rate of reaction constants. Furthermore, an improved second-order kinetic model is proposed in [38] to investigate the insulating paper ageing state under an axial temperature gradient. In [30], the authors estimate the effect of moisture, oxygen, and hot-spot temperature on the lifespan of the transformer with TUP and Kraft paper. The result of the study considering hot-spot temperatures ranging from 50 ℃ to 140 ℃, and varying moisture content ranging from 0.5% to 5% shows that at low oxygen conditions, the paper lifespan is longer than at medium and high oxygen levels. Also, both papers behave similarly with an increase in the oxygen level. However, the result indicates that lifespan TUP doubles that of Kraft paper.

In addition, Calvini proposed the kinetics of cellulose degradation as given in equation (8) based on the inherent structure of cellulose. This proposed equation suggests that cellulose degradation primarily involves the accumulation of bond scission of weak, amorphous, and crystalline links.

$$\frac{1}{DP_t} - \frac{1}{DP_0} = n_{w0}(1 - e^{-k_w t}) + n_{a0}(1 - e^{-k_a t}) + n_{c0}(1 - e^{-k_c t}) \quad (8)$$

Where n_t is the number of bond scission at t, n_0 is the initial number of links available for degradation, a, c, and w represent amorphous, crystalline, and weak links respectively, and all other parameters have their usual meaning as earlier used [39, 40].

B. EXPERIMENTAL CONCEPT

When estimating the DP of an insulating paper, researchers have incorporated new experimental techniques beyond assessing furanic compounds, oxides of carbon, ethanol, and methanol levels, as discussed in our previous review paper in [7].

The authors in [41] employed dispersion staining colours (DSC) as a marker to determine the ageing of insulating paper under accelerated thermal ageing. The cellulose fibres DSC was observed using dispersion staining techniques. It was observed that the transition of DSCs from blue to purple and then to red or orange signifies an increase in the refractive index of cellulose fibres obtained from the insulating paper. Also, the ageing of insulating paper increases with an increase in the refractive index of the cellulose fibres. The proposed marker for the ageing of insulating paper shows an advantage over the carbon oxides and 2-fulfural (2FAL) as a comparison was done with other works in the study. The sugar concentration was utilized as a marker for the degradation of cellulose-insulating paper in [42]. The study unveils a linear correlation between the DP of the paper and the logarithm of the total sugar concentration. This correlation becomes particularly pronounced during the initial stages of ageing as the total sugar concentration experiences a significant increase. A study in [43] reveals that the degree of polymerization of insulation paper can be examined using an optical method. Their experimental work and results show a significant enhancement over the ASTM D4243 viscometric method as measurement is said to be completed in a short duration. Furthermore, the ageing condition of insulation paper was identified in [44] by utilizing the texture eigenvalues of the captured paper images. This was simply achieved by collecting and pre-processing the images of the paper with different thermal ageing states and subsequently estimating the gray scale co-occurrence matrix for these images. The resulting texture eigenvalues then serve as input for the input layer of the backpropagation neural network (BPNN) used. In [45], the authors estimated the ageing condition of insulating paper at the hot-spot of a transformer utilizing a modified dielectric response model and reinforcement learning-based optimized with genetic algorithms. This estimation is achieved by collecting the polarization and depolarization currents (PDC) associated with the ageing state of transformer insulation. Subsequently, the modified dielectric response model is introduced to express the PDC characteristics of the insulating paper. Then, the reinforcement learning-based genetic algorithm (RLGA) is used to explore optimal model parameters specified in the modified dielectric model, aiming to effectively represent the insulating paper ageing condition at the hot-spot.

C. EOL ASSESSMENT

Models for estimating insulating paper EOL in given in Table 1. where $\%LL$ is the percent of life lost, F_{EQA} is the equivalent ageing factor, L is the loss of life, V is the paper rate of ageing, t_n is the time between measurements, and LL_n is the life lost after the time-period. The values of F_{EQA}, V and A can be computed as reported in [7].

III. PROGNOSTICS AND HEALTH MANAGEMENT (PHM)

PHM is a recent development that enables the tracking of the health condition of a system as a result of actual knowledge, data, and information gathered from the systems and their elemental units to identify the onset of anomalies, isolate and diagnose ongoing failures, and forecast the future health state of the system thereby estimating its remaining useful life (RUL), which enables dynamic support for maintenance decisions [51], [52]. PHM aids in implementing routine maintenance as it enhances the system's reliability and availability as well as reducing the cost of maintenance [53]. However, the contemporary impediments to PHM are how to construct an absolute PHM system for a particular sector and how to convert raw signals into knowledge and information to aid maintenance decision activity [54], [55]. The place of prognostics in PHM is shown in Figure. 4 [56].

Furthermore, PHM provides solutions for reducing operation and maintenance costs as it aids the extension of equipment lifetime by early identification of abnormality, diagnostics of equipment health state, RUL prediction, and solutions to condition-based maintenance [57], [58]. Its capacity to utilize the historical, current, and future data of equipment aids in estimating degradations, fault identification, and predictions to effectively manage equipment failures. Also, it relies on sensor technologies to gather precise, long-term, in-situ data about health indicators. This information is essential for assessing degradation levels and predicting RUL. The effectiveness of PHM deployment in enhancing system safety and minimizing maintenance costs is directly correlated with the accuracy of the collected information [59], [60]. As described in Figure 5, PHM generally consists of seven layers, where the layers are further divided into three major stages (observe, analyse, and act) [61], [62]. Also, according to [63], PHM can take place in five processes, which are presented in Table 2.

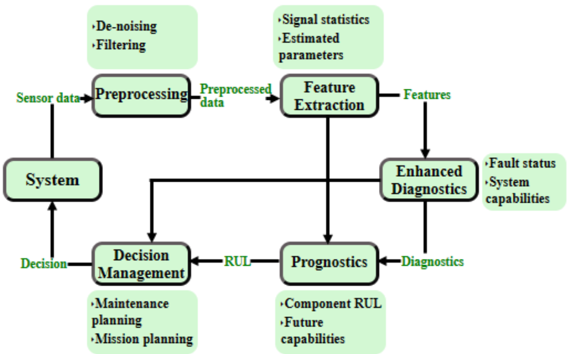


FIGURE 4. Prognostics in PHM [56, 64].

TABLE 1. End-of-useful life model for insulating paper.

Authors	EOL formulation	Considerations	Ref.
IEEE C57.091-2011	$\%LL = \frac{F_{EQA} \times t \times 100}{Normal\ insulation\ life}$	- Oxygen level and paper moisture content are not considered in this model as temperature is taken to be the only degradation parameter for paper insulation. - It does not apply to old transformer design, which could result in an overestimation of RUL	[46], [47]
IEC 60076-7:2005	$L = \int_{t_1}^{t_2} V dt \text{ or } L \approx \sum_{n=1}^{N} V_n \times t_n$	- Moisture content and oxygen level are not considered in this model as temperature is taken to be the only degradation parameter for paper insulation - Not applicable to old transformer design, which could result in overestimation of RUL	[48]
IEC 60076-7:2017	Utilizes the same formulation as IEC 60076-7:2005	- Considered the effect of water and oxygen as the paper ageing rate is given as: $V = \frac{k}{k_r} = \frac{A}{A_r} e^{\frac{1}{R}\times\left(\frac{E_r}{\theta_{h,r}+273} - \frac{E}{\theta_h+273}\right)}$	[49]
Martin et. al.	$LL_n = \frac{t_n}{\frac{\left(\frac{1}{200}-\frac{1}{1000}\right)}{A} \times e^{\frac{E_a}{RT}}}$	- IEEE and IEC models are modified by introducing water content and oxygen level to the degradation parameter parameters. - Applicable to old transformer design	[27]
Pradhan et al.	$L = \frac{41}{2} \ln\left(\frac{1100}{DP}\right)$	- Only considered paper DP value	[50]

TABLE 2. Processes for PHM.

PHM core operational processes	Functional block	Illustration
Act	Health management	The function utilizes insights gathered during the advisory generation phase, employing them to execute actions that bring the system back to an optimal and healthy state.
Advise	Advisory generation	This function delivers actionable information to operational staff or external systems, enabling them to take effective actions based on the provided insights.
	Prognostic assessment,	This function offers prognostic information regarding future health, remaining performance life, or indicators of useful life for the system

Analyse	Health assessment	This function offers information necessary to assess the present health status of the system.
	State detection	This function assesses the state conditions of equipment by comparing them to normal operating profiles, generating indicators for both normal and abnormal conditions.
Acquire	Data manipulation	This function handles the processing and transformation of sensor data and health state information collected by the data acquisition system.
	Data acquisition	This function is responsible for obtaining and recording sensor data and health state information from the internal monitors of the system, as well as from the system's data bus or data recorder.
Sense	Sensors	This encompasses both physical sensors and any soft system performance variables that are accessible within the system.

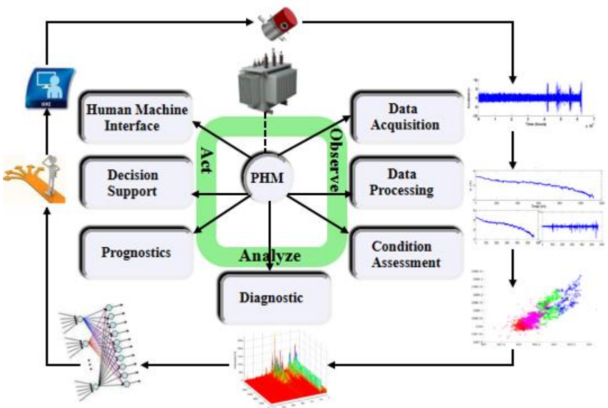


FIGURE 5. PHM chart.

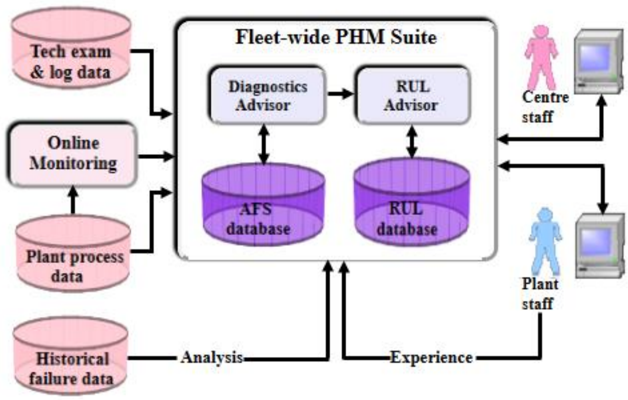


FIGURE 6. FW-PHM data flow by EPRI [66].

A. FW-PHM SUITE SOFTWARE

Fleet wide- prognostics and health management (FW-PHM) suite software developed by the Electric Power Research Institute (EPRI) uniquely for the fossil fuel and nuclear power industry. FW-PHM suite is a collection of online diagnostic and prognostic databases and tools to set up an integrated architecture for equipment health monitoring ranging from individual components to the whole power unit. It entails four major modules explained based on functionality in Table 3 while Figure 6 shows the shows data flow for the FW-PHM suite [65].

TABLE 3. Modules of FW-PHM

Modules	Functionality
AFS	ASF helps to classify the fault signatures obtained from several member utilities
DA	DA helps to recognize potential faults by correlating ASFs with operating data
RULA	RULA helps to estimate the RUL for equipment according to DA diagnostic information as well as the model parameters, model type, and input process parameters
RULD	RULD helps to classify equipment RUL model/signatures obtained from several industries

B. RUL PREDICTION

The RUL of a power transformer is defined as the duration extending from the present moment to the end of its operational viability [67], [68]. The time distance from the current prediction time t_p, to the failure time indicates the RUL of the equipment as given in equation (9).

$$RUL_i = EOL_i - t_{pi} \,|EOL > t_p \qquad (9)$$

Where EOL is the failure time (end of useful life).

Uncertainty modeling is key to predicting RUL accurately as the remaining time after t_p till EOL is random [69]. The concept of RUL prediction is given in Figure 7, where the data samples collected till the prediction point t_p is denoted by $Y = [y_1, \ldots, y_n]$.

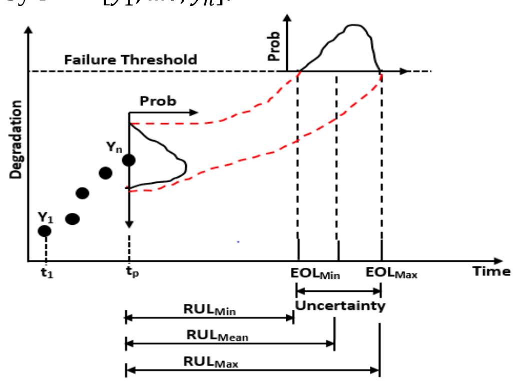


FIGURE 7. Prediction of RUL [70, 71].

IV. PROGNOSTICS APPROACH

In the industrial sphere where reliability, safety, and reduction in cost are highly prioritized, prognostics is a crucial activity of condition-based maintenance. In this regard, the key prognostic objective is to provide the RUL of degrading equipment, which is to forecast the time the equipment will no longer be competent to meet its operational demands and functions [72]. The prognostic approach tends to examine the degree of deviation and deterioration of the equipment from its anticipated normal operational state. The state of health of any degrading component is proportional to its operating time and usage. Comprehensive degradation data of equipment as well as selecting an appropriate modeling method is required to provide an accurate prognostic [9]. The Prognostics approach as other advancing technology is faced with some difficulties, which are addressed by the prognostics centre for excellence. These challenges are highlighted in Table 4 and the questions for each of the challenges have been justified in [56]. Furthermore, the prognostics approach is generally classified into three major classes, which are the physics-based, data-driven, and hybrid approaches [73]. Generally, prognostic approaches can be classified into three classes and the choice for selecting any of the approaches is shown in Figure 8 [56], [74].

TABLE 4. Challenges faced by the prognostics approach.

Challenges	Question
Uncertainty management	How can we effectively capture and process information from various sources of uncertainty?
Autonomic control configuration	How can local prognostic information be converted into changes at the controller level to ensure long-term satisfaction of the controller objectives?
Integration	How can we appropriately merge and process information from various interacting subsystems?
Validation and verification of prognostics	How can we validate the accurate functioning of prognostic algorithms, especially when applied to new systems?
Post-prognostic reasoning	How can data from a prognostic reasoner be translated into actionable steps, considering additional factors like logistics, mission details, and fleet management?

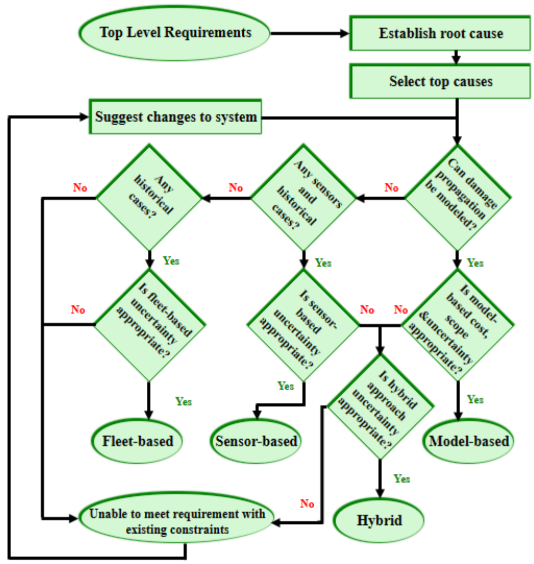


FIGURE 8. Selection decision for prognostics approach.

A. PHYSICS-BASED MODEL APPROACH

This approach sets up a mathematical model that directly links the physical processes that govern the equipment's health. It could be semi-empirical and mechanistic representing the grey box model and white box model respectively [75]. Dynamic systems like differential equations, nonlinear equations, and state space models are solved appropriately to characterize the model. However, building physics-based models for complex systems is difficult as essential knowledge of the material characteristics, failure mechanisms, and operation conditions, and degradation of physical phenomena is usually unknown [68], [76], [77]. In addition, physics-based approaches cannot be generalized as they are specific to certain applications [72], [78]. The physics-based approach refers to high fidelity simulation, differential equations, and finite element models [64].

B. DATA-DRIVEN APPROACH

Data-driven models also known as black box models are derived from condition monitoring data collected regularly, which aids in the learning of system behaviours rather than developing extensive human expertise and system physics models [75], [79], [80]. In data-driven approaches, a predictive model is created by training data and the validation of the model is achieved by testing data. This approach is effortlessly executed as it builds on historical accounts to predict output with regard to condition-monitored data [81]. Also, it employs interpolation, extrapolation, and ML for prediction [81]. However, this approach offers less accurate outputs compared to the physics-based model [72]. Also, the potential for this approach to account for uncertainties is limited [51], [60]. The classification and methodology for a data-driven model are illustrated in Figures 9 and 10 respectively [9], [56], [58] and Table 5 highlights some of the pros and cons of the different classes of data-driven models.

Some pros and cons of the data-driven models utilized in transformer insulation prediction are presented in Table 6.

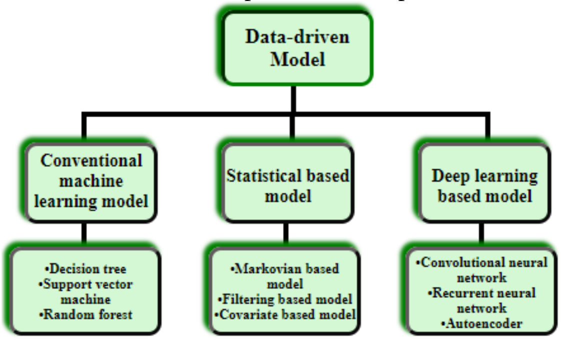


FIGURE 9. Classification of data-driven model.

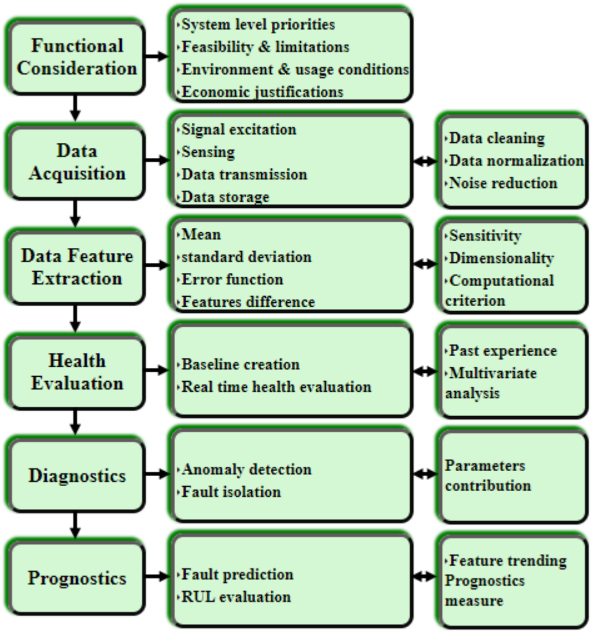


FIGURE 10. Methodology for data-driven model.

C. HYBRID APPROACH

The hybrid approach integrates physics-based models and data-driven approaches. The former offers offline validation of the physical model while the latter is used to increase accuracy by updating the parameters of the model [72]. This approach provides understandable results as it utilizes the advantages of the two approaches to accomplish finely tuned prognostics models that possess an excellent ability to handle uncertainties from many sources leading to accurate evaluation of RUL [134]. This is easily achievable as data is employed to compensate for the absence of knowledge [64]. Furthermore, this approach can be modeled in three ways, which are the series approach, parallel approach and parallel-series approach [61], [135]. Chao et al. [136] utilized hybrid approaches, leveraging the advantages of both physics-based and data-driven models for accurate prognostics. Physics-based models, which entail deducing unobservable model parameters associated with the health of a system's components through the resolution of a calibration problem amalgamated with sensor readings generated by commercial modular aero-propulsion system simulation (CMAPSS), which then serve as inputs for the deep neural network (feedforward neural network (FNN), convolutional neural network (CNN), long short-term memory (LSTM)) used. The performance of the proposed hybrid method was assessed by comparing it with an alternative data-driven method where sensor data solely serves as input to three distinct types of deep neural networks (multilayer perceptron-feedforward neural network (MLP-FNN), recurrent neural network (RNN), CNN). It was then observed that the prediction horizon increased by approximately 127%. Furthermore, the proposed hybrid method surpasses the data-driven approaches with an outstanding RUL estimation amidst highly variable operating conditions and an incomplete representation of the training dataset. It was concluded that the hybrid framework necessitates less training data relative to purely data-driven algorithms. Following this, Table 7 highlights some of the pros and cons of the different types of prognostic approaches.

TABLE 5. Pros and cons of a data-driven approach.

Data-driven model	Pros	Cons	Ref.
Conventional machine learning model	- Benefit from feature engineering - Is inherently interpretable - Requires less computational resources - Performs excellently well with smaller datasets	- Low computational speed - Labour intensive and expertise dependent as feature engineering is required - Inadequate to deal with long-time series analysis leading to errors in prediction results	[84]-[87]
Statistical based model	- Inherently consider uncertainty in data - Gives interpretable and clear representation of the relationships between datasets	- Assumptions not suitable for real-world industrial applications - Displays large error when complete data is unavailable	[61], [72], [88]
Deep learning-based model	- Has a high level of accuracy - Capability to autonomously extract essential features and representations of high-level - Does not require feature engineering	- Difficult to explain and interpret the results obtained - Unknown relationship between inputs and output - Requires remarkable computational power	[72], [86], [87], [89]

	- Appropriate where a physics-based model is not conducive to replicating system behaviour	- Requires more time for training and inference	

TABLE 6. Pros and cons of some data-driven models.

Table 6. Pros and cons of some data-driven models.

Model	Pros	Cons	Applicable scenario	Ref.
ANN	- Display high classification accuracy. - Has strong parallel distributed processing capacity - Do not generate uncertainty information. - Can be trained by backpropagation, which enables the model to learn from its errors and adjust its internal parameters (weights and biases) - Excellent capacity to extrapolate and interpolate	- Lack capacity to address misclassified data samples due to their deterministic diagnostics. - Poor generalization capacity - Computational speed is low - Vulnerable to overfitting - Difficulty in parameter adjustments - As a black box model, it has limited capacity to interpret and clarify their results	- Where large data is available. - In a complex pattern recognition. - In a non-linear relationship model - Where feature extraction is crucial - In real-time decision-making systems. - Adaptability to changing environments and evolving datasets.	[15], [90]- [96]
PNN	- Easy to train - Arbitrary nonlinear approximation - Has a fast convergence speed - Erroneous data are acceptable - Capable of integrating the benefit of both non-linear and linear algorithms	- Can not accurately predict intricate hierarchical structures - Sensitive to feature scaling - Computational intensive with large data - Prone to overfitting	- In pattern classification with uncertainty. - In non-linear classification. - Where noisy data is available. - In small to medium-sized datasets. - In sequential data analysis.	[95], [97], [98]
RNN	- Appropriate for time series signals and prediction with strong stability and adaptability - Demonstrates robustness in handling random length sequence data - Capable of retaining short-time information in dynamic processes - Capable of capturing the temporal correlation in sequential data - Uses its cells for capturing model uncertainty.	- Training and implementation are difficult tasks and cannot be computed in parallel - Vanishing or exploding gradients problem when addressing long sequence data, which leads to non-convergence - Avoids vital information right from the initial input level. - Constrained by recurrent mode, which inherently restricts their computational speed.	- In sequential data processing. - Where modeling of temporal relationship is required. - In memory-based tasks. - In transfer learning. - In online learning and real-time prediction. - Where the length of the input or out can be varied.	[9], [70], [84], [85], [94], [99]- [104]
CNN	- Excellent performance in local feature extraction as it preserves all the localized cues even as feature maps - Ability to address learning problems that involve multi-dimensional input data with intricate spatial structures. - Enhance convergence speed by preventing over-fitting - Can capture strong temporal and spatial correlation in signals	- Insufficient to capture and learn long-term dependencies in data - Efficiency and training speed not sufficient - Has poor global modeling capability. - Requires a fixed input size - Unable to capture remote features when processing temporal features	- In image recognition and classification. - In feature extraction from images. - In spatial hierarchical representation learning. - In transfer learning anf fine tuning. - In time-series data analysis.	[100], [104]- [112]
	- Appropriate for time series	- Networks face challenges in achieving optimal performance	- In sequential data modeling.	

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3391823

LSTM	analysis and long-sequence prediction - Vanishing or exploding gradient problem is partially solved through memory cells and a gating mechanism - Exceptionally capacity to learn long-sequence dependencies - Easy to identify and capture significant features over a long-distance - Has better performance than RNN as its hidden layer neuron is being substituted with memory cells	during parallel processing due to their intrinsic sequential nature - Neglect meaningful signals and give preference to recent data in the handling of extremely long time-series signals - Still struggle to capture long-term dependencies as the gradient becomes smaller	- In natural language processing. - Time-series anomaly detection. - In gesture and speech recognition.	[9], [84], [86], [87], [94], [99], [100], [113]-[116]
AE	- Used for dimensionality reduction - Does not require data labels - Solve the issue of noise interference	- Poor convergence - Difficulty in capturing the correlation of features - Learning requires an abundant amount of data	- In dimensional reduction. - In anomaly detection task. - In data denoising. - In feature learning.	[104], [107]
SVM	- Display good classification accuracy - Solve nonlinearity and high-dimension problem by utilizing kernel function - Better generalization ability with slight over-fitting	- Lack capacity to address misclassified data samples due to their deterministic diagnostics. - Greatly affected by the input vectors quality - Inability to classify nonlinear samples	- In a binary classification task - In a non-linear classification task. - In regression analysis - In the outlier detection task. - In text and image classification.	[90], [91], [117]-[119]
kNN	- A non-parametric model - Suitable prediction model for a small quantity of data - Can handle missing data	- Sensitive to outlying data points, especially when the k value is small - Less efficient for real-time application and less datasets - depends on the optimal value of k	- In classification and regression tasks. - In the anomaly detection task - Where imputing missing value is required. - In clustering task.	[14], [120], [121]
GBN	- A statistical and white box model that incorporated expert knowledge either by first-principle model or as a causal model - Deduce uncertainty information as it captures causality or probabilistic dependencies among random variables - has the potential to handle continuous data	- Cannot handle categorical data successfully - Difficult to capture intricate nonlinear relationships present in data - depends on assumptions that may not hold in real-world applications	- In text classification task.	[90], [91], [122], [123]
XGBoost	- Create an accurate learner through the combination of numerous regression trees. - Minimize the training loss while preventing overfitting through the incorporation of regularization terms. - Has scalability ability in all situations	- Only give more accurate results when features exhibit coherent relationships and are well-defined - Sensitive to noisy data and outliers - Computational intensive when training deep trees - Is a complex ensemble model with many hyperparameters that need to be tuned	- In classification and regression tasks. - Where feature importance analysis is required. - In time-series forecasting. - in ensemble learning	[124]-[128]

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3391823

	- Increase the number and depth of weak classifiers to optimal classification - Capability to address class imbalances by weight adjustment for unevenly distributed target classes - Able to identify features with the most significant impact on prediction to understand the factor influencing the model decisions - Has robust performance even without extensive hyperparameter tuning			
LSSVR	- Acceptable computational capacity - Good generalization capacity	- Poor accuracy when there is loss or incomplete historical data - Demands a substantial amount of prediction time	- In function approximation tasks. - In multi-output regression tasks. - In noise reduction tasks.	[93]-[129]
RVM	- Capacity to effectively handle high-dimensional data and provide probabilistic outputs as it has a Bayesian foundation. - Good generalization capacity - Acceptable computational capacity	- Requires additional optimization algorithms to find the kernels optimized sparse weights distribution as it is needed to reduce the prediction time	- Where a sparse representation of the model is desired - For feature ranking and selection tasks. - Classification with imbalanced data tasks.	[93]-[130]
Random forest	- Ability to train faster, robust and good balance of errors - Offers minimized risk of overfitting when extrapolating unobserved data - Reduced set of hyperparameters to configured	- Limited when extrapolating data beyond previously unobserved ranges - Challenges in understanding the decision-making process for individual tree - It is computationally expensive	- In classification and regression tasks. - Where feature importance scores are desired. - For outlier detection tasks.	[94]-[131]
GM	- Desirable predictions with small-scale data. - Most appropriate for scenarios where observed variables change monotonously over time - Need a limited amount of historical data	- Occurrence of inherent error persists - sensitive to initial conditions - Assumes linearity in the essential data trend - Difficult to capture nonlinear and complex data	- For small sample size tasks. - For missing or incomplete data tasks - For emerging or new trends cases.	[132]-[133]

TABLE 7. Pros and cons of different prognostic approaches.

Approach	Pros	Cons	Requirement	Ref.
Physics-based model	- RUL is evaluated at the early stages - Fewer failure data is required for RUL assessment - Seamlessly implemented and executed in real-time (online). - Offers insight into the internal physical parameters of the system - Has good generalization capacity. - Suitable for component-level	- Not frequently employed in practical applications as specific failure mechanisms which are challenging to gather are required - Difficult for complex systems. - Wide scope of the system is limited - Prediction accuracy relies on the understanding of the underlying system's physical behaviour. - Complex systems domain knowledge is often either unavailable or excessively costly.	- Working model not needed - Failure history needed - No requirement for past status - No requirement for current status - Identifying types of fault not needed - No requirement for maintenance history - Sensors and models not needed	[61], [72], [75], [82], [104], [137], [138]

Data-driven	- Assumptions about the model is not needed as knowledge about the insight parameter of the internal system is not required - Appropriate for the analysis of complex systems to complement the rarely available degradation knowledge about the system. - Deployment is achieved faster at low cost - Human experts or prior knowledge is not required - Wide scope of the system is accessible - Effective than a physics-based model for complex systems.	- Suffers inadequate representativeness of training data. - Heavy computational load is required - Give less accurate results compared to the physics-based model as a large quantity of exhaustive training data is required from systems of the same kind and maker - Depend on the assumption that the statistical characteristics of system data remain relatively constant unless a malfunction occurs - Limitation in observing and correlating the variation in the internal state parameters within the system - poor generalization and extrapolation capability	- No requirement for a working model - Failure history not needed - Past status not needed - Current status needed - Identifying types of fault is needed - Maintenance history not needed - Sensors needed, and models not needed	[61], [71], [72], [75], [104], [136], [138]
Hybrid	- Employs physics-based model updated by data-driven approach to increase prediction accuracy	- Suffers limitation of model incompleteness - Rarely utilized as it is computationally expensive and physical knowledge is required	- Working model needed - No requirement for failure history - Past status needed - Current status needed - Identifying types of fault is needed - No requirement for maintenance history - Both sensors and models are needed	[72], [104], [136], [138]

V. PROGNOSTICS DATA ANALYSIS

Prognostic degradation models usually need a precise and extensive historical degradation signal database. These signals help to identify degradation trends essential for predicting lifespan. However, for practical applications in real-world scenarios, these degradation signals often have missing observations, making it challenging to identify a suitable degradation model when there is a significant amount of missing data. Also, the data collected from several sensors may have different scale and numerical ranges which could remarkably influence the convergence speed and prediction accuracy [86]. Therefore, studies have been performed in this line to manage these issues by developing some approaches to mitigate these issues. Some ways of pre-processing data obtained from sensors are presented in Table 8.

A. ADDRESSING IMBALANCE DATA

In [139], the authors employed semi-parametric techniques to create a prognostic degradation model for data that are insufficient and fragmented. In their techniques, functional principal component analysis (FPCA), which is a non-parametric functional data analysis method that helps to identify crucial patterns of variation in functional data was utilised to recognise the principal features of the insufficient signal. FPCA can offer a concise and low-dimensional representation of each curve by condensing it into a set of scores known as functional principal components scores (FPC-scores), which are signal features. However, this technique cannot be applied for multiple failure modes, as it is limited to a single failure mode application. Also, the technique can only utilize observation from a single sensor, which can limit the practical application of the techniques as multiple sensors are employed in system monitoring. The authors in [140] extract valuable and irrelevant information from sensor data utilizing the principal component analysis (PCA), which is then linked with an auto-associative neural network (AANN) for classification. Furthermore, PCA requires linear or non-linear dependency among the observed variables. This makes it effective when there is the existence of either linear or non-linear correlation among sets of measurement vectors, unlike Pearson correlation coefficient (PCC) which only has linear dependency among the observed variables [141].

TABLE 8. **Concept of sensors data pre-processing.**

Term	Method	Function	Ref.
Feature selection	Pearson correlation coefficient (PCA)	- Identifies and selects the most informative features from the sensor as some sensor measurements remain constant and do not offer valuable degradation information. - Reduce overfitting	[87], [110], [142], [143]
Dimension ality reduction	Principal correlation analysis (PCA)	- Reduces the number of features in a dataset to retain essential information - Prevent overfitting and minimize computational stress - Discover the primary axes of maximum variance within a high-dimensional data space and project it onto a new subspace, maintaining equal or fewer dimensions than the original	[110]
Feature scaling	Min-max normalization method Z-score normalization (standardization)	- Helps in data normalization or standardization - Ensures values of features are on a similar scale to prevent certain features from dominating the learning process - Helps to speed up the training process - Utilized after feature selection and dimensionality reduction - Helps to prevent convergence difficulty	[86], [87], [108], [110], [144]

B. MODEL EVALUATION METRICS

Several researchers have proposed and developed different algorithms and have utilized different assessment metrics for their proposed models. Among others, the most popular performance assessment metrics used are mean square error (MSE), root mean squared error (RMSE), mean absolute error (MAE), scoring function (SF), mean absolute percentage error (MAPE), relative absolute error (RAE), symmetric mean absolute error (SMAPE), root mean squared logarithmic error (RMSLE), and coefficient of determination (R^2), which are expressed as follows.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{11}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{12}$$

$$RAE = 100 \times \frac{\sum_{i=1}^{n}|\hat{y}_i - y_i|}{\sum_{i=1}^{n}|\hat{y}_i - \bar{y}|} \tag{13}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{y}_i - y_i|}{|y_i|} \tag{14}$$

$$SMAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{2|\hat{y}_i - y_i|}{|\hat{y}_i| + |y_i|} \tag{15}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2} \tag{16}$$

$$SF = \begin{cases} \sum_{i=1}^{N}\left(e^{+\frac{\hat{y}_i - y_i}{10}} - 1\right), & for\ \hat{y}_i - y_i \geq 0 \\ \sum_{i=1}^{N}\left(e^{-\frac{\hat{y}_i - y_i}{13}} - 1\right), & for\ \hat{y}_i - y_i < 0 \end{cases} \tag{17}$$

$$RMSLE = \sqrt{\frac{\sum_{i=1}^{N}[log(\hat{y}_i + 1) - log(y_i + 1)]^2}{N}} \tag{18}$$

Where N represents the number of samples or units used as the input. y_i is the real or actual value, $\hat{y}_i$ is the predicted value, and $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ is the average value of y_i [9], [86], [103], [107], [127], [145]- [148] A smaller value of RMSE and MAPE signifies more accurate predictions with reduced errors. Furthermore, R^2 is employed to assess the interpretability of the prognostic model. A higher R^2 value signifies good prediction, and an R^2 value of 1 signifies a perfect fit for the prediction [149]. Also, The scoring function is introduced to severely penalized late prediction [150].

VI. PAPER DEGRADATION PARAMETERS

Ageing of the paper insulation is an irreversible rupturing of covalent and hydrogen bonds within and between the cellulose polymer chains. This degradation results from the combined influences of pyrolysis, oxidation and hydrolysis reactions, which are predominantly governed by temperature, reactive oxygen species and moisture respectively.

A. TEMPERATURE AND AGEING

Paper insulation's thermal stability has drawn numerous researchers' attention as power transformers operate at high temperatures while a lifetime of about 40 years is expected. Paper with a higher content of lignin exhibits thermal stability as it can produce highly condensed structures during thermal decomposition [18]. Pyrolytic degradation arises solely from elevated temperatures. Its activation energy is about 1.4 to 2 times greater than that of hydrolysis. Above 130 °C, pyrolysis becomes the predominant degradation process and it transited into a self-accelerated reaction at a temperature of 140 ℃ due to the generation of water and oxygen [47]. Temperature increase can lead to increasing overheating and conductivity, which causes localized carbonization of the insulating paper [151]. Furthermore, the insulating paper extracted from the uppermost part of the winding has the lowest DP values due to temperature difference [152]. The primary factor governing the RUL of the insulation paper is the hot-spot temperature as it influences the acceleration production of furans, moisture, CO, and CO_2 that could lead to direful breakdown during high loading conditions [47], [153]. However, underestimation of the hot-spot temperature has been a common practice as the estimation of the hot-spot temperature is complex, which could cause the power equipment to operate at a reduced cooling system [154]. Following IEEE C57.91 guidelines, the evaluation of insulating paper loss-of-life is based on the hot-spot temperature experienced during a specific period of power transformer operation. This methodology is predicated on the understanding that heat predominantly drives the ageing process and presumes that the levels of acidity and humidity within the insulation system remain unchanged [155].

The Arrhenius equation is broadly recognized as a standard model for describing the ageing process of insulating paper. The model given in equation (19) employs temperature as the only parameter influencing insulating paper degradation.

$$\frac{1}{DP(t)} - \frac{1}{DP(t_o)} = A\Delta t e^{-\frac{E_a}{R(\theta_{HS}(t)+273)}} \qquad (19)$$

Where E_a in Jmol^{-1} is the ageing reaction activation energy, R in $\mathrm{Jmol}^{-1}\mathrm{K}^{-1}$ is the gas constant, A is the pre-exponential factor whose outcome is dependent upon the chemical environment, Δt is the ageing period, θ_{HS} in ℃ is the insulating paper temperature (hot-spot temperature). $DP(t)$ and $DP(t_o)$ are the DP at the end t and start t_o respectively [47]. A post-mortem analysis was carried out in [156] which validated that the parameters E and A in equation (19) lack the required precision for evaluating insulating paper DP as they do not report for the influences of oxygen, moisture, and acidity. In [157], several values A in hour^{-1} are presume in equation (19) to be either hydrolysis or oxidation pre-exponential factor values that depend on θ_{HS} value. Also, in this work, a steady state equation is proposed for hot-spot temperature and it is expressed as:

$$\theta_{HS} = \theta_A + \theta_{TO} + HSF \times \theta_{OW} \qquad (20)$$

Where θ_{HS} is the hot-spot temperature, θ_A is the ambient temperature, θ_{TO} is the temperature of the top oil, and θ_{OW} is the oil-to-winding temperature gradient under rated load. The HSF is the hot-spot factor. The transformer HSF value characterizes the additional temperature increase in the windings above the top winding temperature, primarily due to the increased eddy current losses at the top of the winding. The HSF-specific value depends on a variety of design and manufacturing factors and can be determined accurately by directly measuring the winding θ_{HS} using fibre optic sensors. However, this technique is typically feasible only in new transformers. Consequently, when no calculations or test data are available, an HSF value of 1.3 for power transformers is suggested by IEC loading guidelines [158]. The increase in the hot-spot temperature over the temperature of the top-oil after a step load change has been documented as a time-dependent function influenced by the duration and the load of the transformer, which is referred to as an overshoot time-dependent function. The modeling of this overshoot phenomenon led to the introduction of a mathematical representation based on observed data in the IEC 600076-7 loading guide [49], and a graphical illustration comparing this overshoot based on the IEC 600076-7 model with the real transformer with forced and natural cooling can be seen in [159]. Also, this overshoot is reported to be unobservable with the IEEE Clause 7 loading guide model but observed when employing the IEEE Annex G model as it utilizes the physics-based modeling that incorporates the impact of viscosity and DC variation, stray and eddy losses with temperature. Generally, the IEEE C57:91 Clause 7, IEC 60354, and IEC 60076-7 loading guide for dynamic thermal models are accurately presented in [160] and are commonly utilized in several applications because their models necessitate only fundamental input parameters that can be obtained from a standard transformer heat-run test report.

A dynamic model was proposed in [161] for evaluating the hot-spot temperature θ_{HS} as the temperature distribution within a power transformer is not uniform. This model relies on solving the following pair of differential equations:

$$\frac{(1+RK_L^2)}{(1+R)}\mu_{pu}^n \Delta\theta_{TO,R} = \left(\mu_{pu}^n \tau_{TO,R}\frac{d\theta_{oil}}{dt}\right) + \frac{(\theta_{TO} - \theta_{amb})^{n+1}}{\Delta\theta_{TO,R}^n} \qquad (21)$$

$$K^2 P_{W,pu}(\theta_{HS})\mu_{pu}^n \Delta\theta_{HS,R} = \left(\mu_{pu}^n \tau_{W,R}\frac{d\theta_{HS}}{dt}\right) + \frac{(\theta_{HS} - \theta_{TO})^{m+1}}{\Delta\theta_{HS,R}^m} \qquad (22)$$

Where θ_{HS} measured in ℃ is situated at the top portion of the winding, θ_{HS} depends on factors such as the ambient temperature, load and the structural characteristics of the unit. θ_{TO} is the temperature of the top oil, K_L is the load factor, R is the ratio of load to no-load losses, μ_{pu} in pu is the viscosity of the oil, θ_{amb} is the ambient temperature,

$\Delta\theta_{HS,R}$ is the rated θ_{HS} above the θ_{TO}, $\Delta\theta_{TO,R}$ is the rated θ_{TO} rise above the θ_{amb}, $\tau_{W,R}$ is the winding time constant, $\tau_{TO,R}$ is the oil time constant, and n and m are the oil and winding exponents respectively [155]. More insight into the constant values of m and n as related to variation in the viscosity of the insulating oil is discussed in [161], [162]. $\theta_{HS,i}$ can be gotten by solving equations (20) and (21) simultaneously for each interval i, which can be utilized in equation (5) and (6) to estimate the degree of polymerization [47].

Furthermore, a recurrence equation from IEEE C57.91 which determines the ageing acceleration factor of transformer paper insulation was given in [26] to estimate the RUL. The RUL at a given time t, can be transformed into a Markovian recurrence relation form. In this form, the health state of the insulation paper is contingent solely on its preceding state and current conditions [154], [163].

$$RUL(t) = RUL(t-1) - e^{\left(\frac{15000}{383} - \frac{15000}{273 + \theta_H(t)}\right)} \tag{23}$$

Where $\theta_H(t)$ in °C is the temperature of the hottest-spot of the transformer winding. This parameter can be evaluated as:

$$\theta_H(t) = \theta_{TO}(t) + \Delta\theta_{TO,H}(t) = \theta_A(t) + \Delta\theta_{A,TO}(t) + \Delta\theta_{TO,H}(t) \tag{24}$$

Where θ_{TO}, θ_A, $\Delta\theta_{TO,H}$ and $\Delta\theta_{A,TO}$ are the temperature of the top oil, ambient temperature, temperature rise of the hottest spot over the temperature of the top-oil, and temperature rise of top oil over ambient temperature at time t respectively. $\Delta\theta_{TO,H}(t)$ and $\Delta\theta_{A,TO}(t)$ can be estimated by:

$$\Delta\theta_{TO,H}(t) = \left[\Delta\theta_{TO,H_u}(t) - \Delta\theta_{TO,H_i}(t)\right]\left(1 - e^{-\frac{\Delta t}{\tau_H}}\right) + \Delta\theta_{TO,H_i}(t) \tag{25}$$

$$\Delta\theta_{A,TO}(t) = \left[\Delta\theta_{A,TO_u}(t) - \Delta\theta_{A,TO_i}(t)\right]\left(1 - e^{-\frac{\Delta t}{\tau_{TO}}}\right) + \Delta\theta_{A,TO_i}(t) \tag{26}$$

Where $\Delta\theta_{TO,H_i}(t)$ is the hot-spot temperature rise over ambient temperature, $\Delta\theta_{A,TO_i}(t)$ is the top-oil temperature rise over ambient temperature, τ_H and τ_{TO} are the windings and oil time constant respectively, and Δt is the interval of the loading time. In a steady state, $\Delta\theta_{TO,H_u}(t)$ and $\Delta\theta_{A,TO_u}(t)$, estimated below are the hot-spot temperature and top-oil temperature rise over ambient temperature and top-oil temperature respectively.

$$\Delta\theta_{TO,H}(t) = \Delta\theta_{H,R}.\left(\frac{i(t)}{i_r}\right)^{2m} \tag{27}$$

$$\Delta\theta_{A,TO}(t) = \Delta\theta_{TO,R}.\left(\frac{\left(\frac{i(t)}{i_r}\right)^2.(\gamma+1)}{\gamma+1}\right)^n \tag{28}$$

Where m and n are parameters of a transformer obtained from a lookup table based on the transformer cooling system, $i_r, i(t), \gamma, \Delta\theta_{H,R}, \Delta\theta_{TO,R}$ are the rated load, transformer load, the temperature rise of the hottest spot at the rated load, and the temperature rise of the top oil at the rated load [4], [154]. However, in [154] values for uncertainty sources were considered, this makes equations (26) and (26) become:

$$\theta_H(t) = [\theta_{TO}(t) + \varphi_{TO}] + \Delta\theta_{H,R}.\left[\frac{i(t) + \varphi_i}{i_r}\right]^{2m} \tag{29}$$

Also, equation (22) requires incorporation of the uncertainty information correlating to the paper degradation process as it is not a deterministic process. Therefore, the equation is then given by:

$$RUL(t)\ RUL(t-1) + \omega RUL_{t-1} - e^{(15000 + \omega_t)\left[\frac{1}{383} - \frac{1}{\theta_H(t)+273}\right]} \tag{30}$$

Where φ_i and φ_{TO} measurement error for load and top oil respectively, ωRUL_{t-1} is the uncertainty of the lifetime evaluation at $(t-1)$, ω_t is the uncertainty of the degradation process [91].

Furthermore, the hottest-spot temperature of a transformer according to IEC 60076-7 [164] is given as:

$$\theta_H(t) = \theta_{TO}(t) + \Delta\theta_H(t) \tag{31}$$

Where $\theta_H(t)$ is the hottest-spot temperature at a given instant t in (°C), $\theta_{TO}(t)$ is the temperature of the top-oil in (°C), and $\Delta\theta_H(t)$ is the hottest-spot temperature rise over the temperature of the top-oil in (°C).

The differential equation can be estimated to different equations for a small Δt. Therefore, the top-oil temperature is evaluated as follows:

$$\theta_{TO}(t) = D\theta_{TO}(t) + \theta_{TO}(t-1) \tag{32}$$

$$D\theta_{TO}(t) = \frac{\Delta t}{k_{11}\tau_{TO}}\left(\Delta\theta_{H,R}\left(\frac{1 + K(t)^2 R}{1 + R}\right)^x + \theta_A(t) - \theta_{TO}(t-1)\right) \tag{33}$$

Where R in W is the ratio of load losses to no load losses, τ_{TO} in minutes is the time constant of the oil, x is the exponent constant of the oil that models the total losses exponential power with respect to the top oil temperature heating, $\Delta\theta_{H,R}$ in °C is the hottest-spot temperature rise at the rated load, $\theta_A(t)$ in °C is the ambient temperature, and k_{11} is the experimental thermal constant obtained.

Also,

$$K(t) = \frac{i(t)}{i_r} \tag{34}$$

Where $i(t)$ in per unit is the load at an instant t, and i_r in A is the rated load.

In addition, the hottest-spot temperature can be evaluated as:

$$\Delta\theta_{H_i}(t) = D\Delta\theta_{H_i}(t) + \Delta\theta_{H_i}(t-1) \tag{35}$$

For $i = [1, 2]$, where

$$D\Delta\theta_{H_1}(t) = \frac{\Delta t}{k_{22}\tau_w}\left(k_{21}\Delta\theta_{H,R}K(t)^y - \Delta\theta_{H_1}(t-1)\right) \tag{36}$$

$$D\Delta\theta_{H_2}(t) = \frac{k_{22}\Delta t}{\tau_{TO}}\left((k_{21}-1)\Delta\theta_{H,R}K(t)^y - \Delta\theta_{H_2}(t-1)\right) \tag{37}$$

Where, $\Delta\theta_{H,R}$ in °C is the hottest-spot temperature rise at rated load, y is the exponential constant of the windings that models the exponential power of the loading with the heating of the windings, τ_w in minutes is the time constant of the winding, $k_{21,}$ and k_{22} are the thermal constants of the transformer [57].

In [165], the authors utilized the physics-based model in equation (22) by the IEEE standard C57.91 to predict the ageing of insulation paper by incorporating the statistical filtering techniques known as particle filter. The particle filter was used to improve the available information by taking into consideration the uncertainties associated with both the exact degradation and the measurement. This was implemented as the IEEE standard C57.91 model for the ageing of paper only offers estimates of the influence of temperature and load, which does not account for uncertainty estimates. Also, in addition to the effect of thermal ageing on insulation paper, the authors in [166] considered the effects of winding vibration on the mechanical-thermal ageing of paper insulation. According to their study, it was reported that standalone mechanical ageing has minimal impact on the ageing of insulation paper compared to combined mechanical-thermal ageing, which has a significant effect on the paper. Furthermore, the ageing process in combined mechanical-thermal ageing aligns with thermal ageing where the insulation paper functional group remain constant as ageing is initiated with the breakage of the paper glycosidic bond. However, based on the degradation process of cellulose under various ageing conditions, it is evident that thermal stress remains the principal ageing factor, with electrical and mechanical stresses acting as accelerating factors for ageing [39].

B. MOISTURE AND AGEING

The presence of moisture in oil-immersed transformers is majorly a result of the ageing of insulation paper and due to its high dielectric strength, it represents a crucial factor affecting insulation reliability [167]. Therefore, paper insulation is vulnerable when the moisture content exceeds 3 % and failure is anticipated if it exceeds 4% [151]. Studies indicate that the rate of degradation at standard service temperatures, with 4% moisture in the paper, is 20 times higher than at 0.5 % [168]. Also, the mechanical strength of an insulating paper has been reported to diminish by halve when the level of moisture in the paper is doubled [118]. Generally, the process of degradation of insulating paper occurs in the presence of moisture at 70 ℃ – 130 ℃ [32]. To maintain dielectric strength and prolong ageing, it is essential to keep the moisture content below 0.5 % and 20 ppm in the paper and oil respectively [16]. The potential cellulose hydrolysis degradation, reaction types, reactants, and products are presented in Table 9. The conventional techniques utilized in detecting and measuring water content are the Karl Fisher titration, spectroscopic frequency-domain technique, spectroscopic time-domain techniques, electrical techniques and paper adsorption isotherms for insulating paper [7], [169], [170]. The mentioned techniques often lead to the overestimation of paper water content. In addition, the adsorption and desorption of moisture in the insulating paper oscillate around an equilibrium point or within a specific range. However, [33] explores an improved approach for assessing the moisture content in transformer paper, which is based on cellulose isotherms. This technique lies in its ability to circumvent the requirement for thermodynamic equilibrium between transformer oil and cellulosic insulation. Instead, it employs cyclic temperature variations to ascertain shifts in the direction of water migration whether into or out of the insulating liquid. These alterations in water migration direction offer a means to determine the vapour pressure of water absorbed by cellulose. Therefore, water activity probes can be employed to measure the temperature and water activity of the insulating liquid around its tip points, which reduces the uncertainties due to the condition of the liquid and non-equilibrium conditions [171]. To estimate the water concentration on insulating paper, the hot-spot temperature of the winding insulation should be considered as the temperature of the winding insulating paper is usually higher than the temperature of the insulating liquid. Furthermore, the use of a water activity probe can overestimate the water concentration on the paper as paper releases moisture into the insulating liquid when it is being heated up. As such, the difference in the temperature of the hot-spot position and the water activity probe should be noted. IEC 60076/7 thermal model [172], and the installation of an optical sensor close to the winding hot-spot can be used to determine this temperature difference. Also, the position of the water activity probes should be considered as the water activity estimation is affected by the speed of the circulating liquid. In this regard, the water activity probe should be placed at the region where the liquid has the highest speed [173]. Also, the dielectric response techniques typically provide data that aligns with online probe data but have the drawback of requiring complete isolation of the transformer from the grid [171]. Furthermore, a capacitive sensor, which is more sensitive to water and has no response to other insulating liquid molecules has been currently employed. However, for optimum data collection, the location to place this sensor remains a problem. Therefore, it is imperative that the sensor is placed to avoid areas with stagnant oil as it exclusively sensitive to the insulating liquid at the surface [22]. Furthermore, the authors in [170] proposed a system utilizing high-frequency sensors to access the deterioration of transformer insulation by measuring water concentration, where an artificial neural network (ANN) model is used to optimize the traditional coaxial probe technique used.

In [47], an empirical expression that establishes the relationship between paper moisture content, hot-spot temperature, and insulating oil moisture content is given in equation (38) under the assumption of equilibrium conditions. This expression is known as ABB's equation.

$$H_p = 2.06915e^{-0.0297\theta_{HS}} \times (H_0)^{0.4089\theta_{HS}^{0.09733}} \qquad (38)$$

Where H_0 in ppm is the humidity of the insulating oil, and H_p in % is the insulating paper moisture content. In [174], a novel and easily accessible technique was used to estimate the oil-immersed insulating paper moisture content. This technique involves the creation of an frequency domain spectroscopy (FDS) curve database that is being facilitated by employing an exponential decay model. The moisture content of insulating liquid-impregnated paper is measured in [175] by establishing a nonlinear coefficient α (feature parameter) of the U-I per-unit curve, which has a directly proportional relationship with moisture content and an inversely proportional relationship with the excitation frequency. In [117], the moisture content in an oil-immersed polymer paper insulation was predicted by employing support vector machine (SVM) with genetic algorithm (GA) to optimize SVM key parameters. The result shows that the model is a potential and robust tool for an oil-immersed polymer insulation moisture content prediction. BPNN enhanced with AdaBoost algorithms was proposed in [1] to address the unreliable evaluation of moisture content. AdaBoost was used for optimization due to its ability to

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3391823

cascade a weak classifier and thoroughly consider the classification outcomes of each individual weak classifier, which helps to avoid the problem of overfitting and gives a higher degree of generalization. Also, BPNN was considered as it excellently helps in the fitting of nonlinear problems. The result of the proposed model gives a more accurate result with lower MAE when compared with some other algorithms like kNN, SVM, BPNN, and GA-BPNN.

TABLE 9. Cellulose degradation reaction types, reactants, and products.

Chemical reaction	Reactant	Product	Ref.
Enzymatic hydrolysis	Cellulose + cellulase	Low molecular weight products	[171, 172]
Acid hydrolysis	Cellulose + H_30^+	D-glucose	[173, 174]
Alkaline hydrolysis	Cellulose + base	Low molecular weight products	[175]
Thermal hydrolysis	Cellulose + H_20	D-glucose	[176]

C. OXYGEN AND AGEING

The chemical reaction between oxygen and the hydrocarbons of any insulating liquid causes deterioration in terms of oxidation and ageing of the liquid which leads to the formation of compounds like acids, sludge, water, and gases as presented in Table 10 with their corresponding effects [182], [183]. Also, the DGA results involving about 1.5 million samples based on IEEE C57.104-2019 standard revealed that the O_2/N_2 ratio is a prominent parameter that exerts a remarkable influence on the typical gas levels detected irrespective of the insulating liquid volume, rating, and voltage class. Therefore, the ratio of O_2/N_2 was proposed for assessment as an indicator to differentiate free-breathing from sealed units. Figure 11 shows transformer classification based on O_2/N_2 ratio [184]. Oxygen, being one of the factors contributing to ageing, demands careful consideration as it decreases the ageing by a factor of 16 when its content in the insulating liquid is reduced from 30000 ppm to 300 ppm [185]. Therefore, the nitrogen-sealed type, diaphragm-sealed type conservators, and the use of oxidation inhibitors have been utilized by researchers to provide a solution as the removal of dissolved oxygen from transformer-insulating liquid is not optional but imperative.

Oxidative degradation of cellulose paper is primarily initiated by ions of transition metal rather than primary metal group ions. In the presence of water and oxygen, they catalyze the formation of hydroxyl anions, hydroxyl radicals, and several other reactive oxygen species. Despite electrical engineers' preeminent endeavours to eliminate metal ions, oxygen, and water within the system, these parameters remain in trace quantities and are ready to instigate a detrimental cycle of paper degradation [8], [186]. Furthermore, the presence of oxygen accelerates the process of paper degradation by increasing its rate by a factor of 2.5 [185]. This was confirmed as an experiment on accelerated ageing at 130 ℃ with and without oxygen was reported in [182], which reveals the ageing of paper insulation was decelerated by inhibiting access to oxygen. Oxidation of paper insulation results in an increase in acidity as a reaction takes place between oxygenated insulating liquid and the paper. This process is most notable at temperatures below 60 ℃ – 75 ℃, indicating that in an unloaded power transformer, acidity is the primary degradation of insulating paper. As such, thermally upgraded paper doesn't offer any benefit over non-thermally upgraded paper when considering oxidation as a degradation factor [47]. Oxygen as one of the parameters for insulating paper degradation has been integrated and captured in other parameters such as moisture and temperature. Nowadays, most of the new and widely used transformers produced are air-free called the hermetically sealed unit, making it a less-prominent factor to be considered by researchers among others.

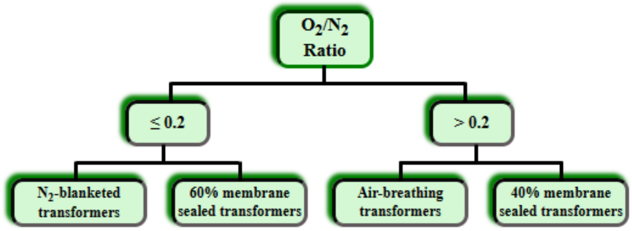


FIGURE 11. Classification based on O_2/N_2 ratio [184].

TABLE 10. Compounds formed and their effects during the chemical reaction between oxygen and insulating liquid.

Compound	Effect
Sludge	- Diminishes dielectric strength
	- Degrades the cellulose paper
	- Interferes with the core and coil cooling
	- Catalyzes further oxidation
Acids	- Induces and promotes corrosion process
	- Diminishes the tensile strength of insulating paper
	- Catalyzes further oxidation
Water	- Degrades the cellulose paper
	- Diminishes dielectric strength
	- Catalyzes further oxidation
Gases	- Produces several faults related gases such as CO_2, CO, H_2, C_4H_6, CH_4, C_2H_4, C_6H_8, etc.

D. INSULATING LIQUID CONDITION AND AGEING

DGA technology stands out as a crucial and efficient tool for detecting early-stage faults in oil-immersed power transformers. This method is instrumental in pinpointing the deterioration of both insulating paper and oil. Carefully examining the CO_2/CO ratio, ethanol content, methanol content, and furan compound analysis can validate the confirmation of paper insulation involvement and potential carbonization through dissolved gas analysis [187]. This technology based on conventional diagnostics techniques such as the Rogers ratio, IEC three-ratio technique, and modified three-ratio technique suffers some limitations like incomplete ratio coding and limited usage conditions [188], [189]. The relationship between some ageing parameters and DP is presented in Figures 12 and 13 (a) to (g). [190]. Also, some comparison models between 2FAL and DP and some transformer ageing models have been highlighted in [191]. Furthermore, a careful study of several researchers' work results has shown different values of DP for the same ageing parameter and insulating liquid as seen in Figure 14. This may be due to the inaccuracy and incapacity of the mathematical models to correlate the degree of polymerization with each individual dissolved gas. Furthermore, IEEE C57.104-2019 [184] and IEC 60599-2015 [192] have proposed some DGA interpretation techniques, which have been generally accepted by electrical industries. However, accurate results from the analysis can only be obtained if the observed gases relate closely to the condition of the equipment as gas data always show a non-linearity characteristic. Furthermore, DGA suffers some limitations such as incorrect collection of samples, uncommon causes of gas generation, several phenomena happening concurrently, accelerations and rate of gas generation which serves as the most reliable basis requires multiple measurements to be conducted over time, and the gas formation pattern and rate characterization is not always adequate to determine the gas generation source [184]. Generally, both offline and offline techniques for DGA suffer the limitations of inability to classify some fault especially when the DGA results do not fall within the IEEE/IEC specifications, leading to ambiguity in the analysis and it is not appropriate for an air-cooled transformer [15], [193]. Therefore, to enhance the accuracy and efficiency of fault identification methods in oil-immersed transformers using DGA, researchers globally have explored the integration of various ML techniques, which have been investigated as a new solution for assessing fault conditions by amalgamating parameters like gas levels, gas rates, and DGA interpretations [194]. Also, incorporating ML approaches with the DGA chromatographic method will help to overcome the impediments in employing standalone DGA-based techniques [195]- [197].

A self-learning technique was presented in [190] to evaluate the DP of insulating paper according to multiple insulating liquid ageing parameters. The authors utilized the FCM-LR method to forecast the DP of insulating paper employing several ageing parameters. The report in this study confirms that DP value prediction using multiple parameters is more accurate than a single parameter. Also, the better performance exhibited by the fuzzy C mean-linear regression (FCM-LR) method positions it as a robust tool for accurately evaluating the DP value of insulation paper. The authors in [198] utilized some ageing parameters present in the insulating liquid of 108 power transformers to determine the expected life estimation and condition of insulating paper by building an ANFIS for the DP of insulating paper. The input parameters considered are acidity, interfacial tension, CO, CO_2, and colour. The result presented shows that the ANFIS model has 85.75% and 89.07% accuracy in testing and training respectively. Also, the proposed model can be utilized to evaluate the paper condition when the 2FAL data of the power transformer is not obtainable. In [102], a combined model based on kernel principal component analysis (KPCA) and a generalized regression neural network (GRNN) utilizing an enhanced fly fruit optimization algorithm (FFOA) to select the smoothing factor parameter in the GRNN network was used to predict the concentration of dissolved gases in an aged insulating liquid. The FFOA optimization techniques were employed to overcome the limitation associated with the neural network approach by aiding the adjustment of training parameters. The proposed model in this study demonstrates superior data fitting and more accurate prediction capabilities relative to the grey model (GM) and the SVM. The authors in [199] employ the SVM classifier model to assess the insulation condition of insulation paper impregnated in insulating liquid. 149 transformers were considered and some dielectric properties, furanic and dissolved gas analysis compounds were used as measurement data. Also, 19 transformers whose furan data were unreachable were considered. After feature extraction and selection were done, the model was observed to have an accuracy of 90.63%. This implies that the model was able to make 29 correct classifications out of 32 based on healthy, moderate, and extensive state categories. In [200], ANN, Gaussian process regression (GPR), SVM, and least-square support vector regression (LSSVM) algorithms were used to predict the dissolved gas employing grey relational analysis (GRA) to estimate the grey relational coefficients for dissolved gas feature selection, coupled with GPR for predicting the dissolved gas values. The test was carried out with eight datasets of dissolved gases. Their results after comparison show that the GRA technique is efficient in identifying and eliminating extraneous and redundant features from the original feature sets. The content of dissolved gases in insulating liquid was forecasted using support vector regression (SVR), which is an amplified edition of the SVM technique in [201]. SVR is considered in this study due to its accuracy, simple structure, and good generalization performance as it is founded on the principle of reducing structural risk rather than empirical risk. Also, kernel functions (polynomial kernel and Gaussian kernel) is integrated to SVR, which is formulated to be mixed kernel function-support vector regression (MKF-SVR) helps in the transformation of non-linear and inseparable problems into linearly divisible problem. This is important as the dissolved gas prediction of liquid immersed transformers involves addressing a non-linear time series problem. Furthermore, the application of GA is utilized to fine-tune the parameters of the SVR model integrated with kernel functions, thereby enhancing the overall forecasting performance. The result from the study after comparison shows that MKF-SVR has superior prediction accuracy and fitting capacity than GM, radial basis function neural network (RBFNN), and GRNN. Time series prediction based on the LSTM method for dissolved gas level in insulation was presented in [84]. In [202], the authors proposed a new asset management technique for power transformer insulating liquid by utilizing online DGA data. Their proposed model consists of two different submodules, which are fault diagnostics and life management modules, these two modules were made to undergo a training process using the CNN machine learning framework. The former utilized the six main types of fault that can be analysed using the DGA approach, while the latter utilized the RUL of the insulating paper as the life expectancy of the power transformer principally depends on paper insulation. LSTM was considered in this work based on its suitability to predict long-term non-linear sequence problems through gating mechanisms. Also, to enhance the prediction accuracy, Bayesian optimization algorithm (BOA) was employed to optimize the model hyperparameters. The proposed LSTM prediction result shows superior prediction accuracy as it was compared with the prediction of GM, SVM, and BPNN. Three data-driven models, which are ANN, SVM, and Gaussian Bayesian network (GBN) are used to diagnose the dissolved gases to monitor the transformer health state in [123]. The dataset classifiers were trained and tested through Monte Carlo cross-validation, which is to allow estimation for statistical classifiers. The result indicates that ANN model emerges as a more accurate diagnostic method than SVM and GBN, where GBN model is observed to have the lowest accuracy. However, information derived from the results of GBN models proves to be more valuable as it was able to generate the probability density function (PDF) with uncertain information. This helps to represent the initial and final health states, which delineate the uncertainty that is linked with these states. Furthermore, the authors suggested that GBN model trained with dissolved gas samples should offer confident predictions for known data and also should be able to express uncertainty when faced with unseen dissolved gas sample data. As such, the uncertainty formation can be employed to refine and enhance the accuracy of the model in the real-world applications. The authors in [203] integrated fuzzy logic and adaptive neuro-fuzzy inference system (ANFIS) to forecast the remnant life cycle of oil-immersed transformer by utilizing the concentration of 2FAL and the degree of polymerization of insulating cellulose paper. Three models (De Pablo, Chendong, and Vaurchex models) relating DP was employed and ANFIS was considered as it provides remarkable good learning ability. The ANFIS model on the testing data was reported to have to offer a 99.86% accuracy indicating the model prediction capability. The summary of several machine learning models use for ageing parameters prediction with the evaluation metrics used is presented in Table 11.

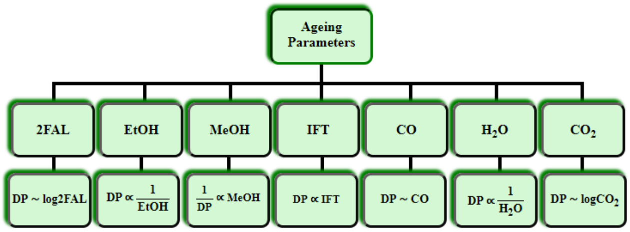
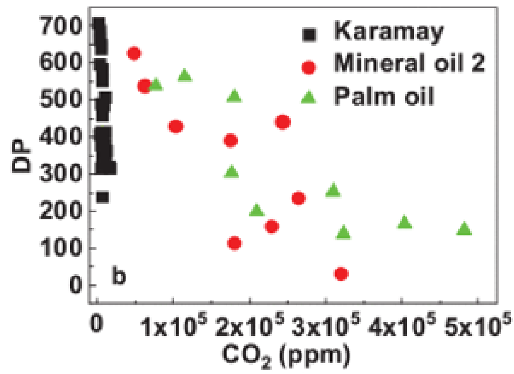
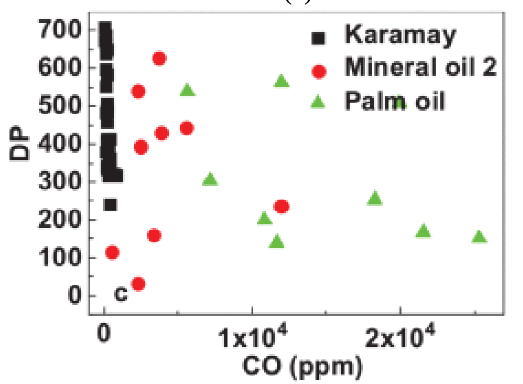


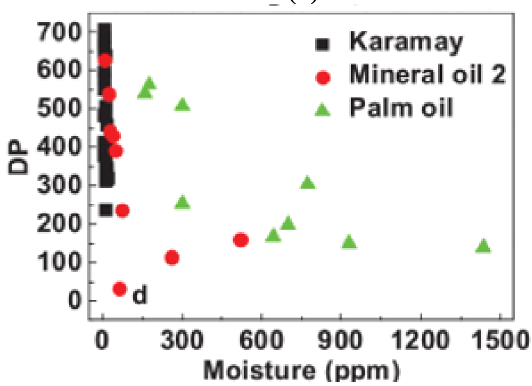
FIGURE 12. Ageing parameters in relationship with paper DP.



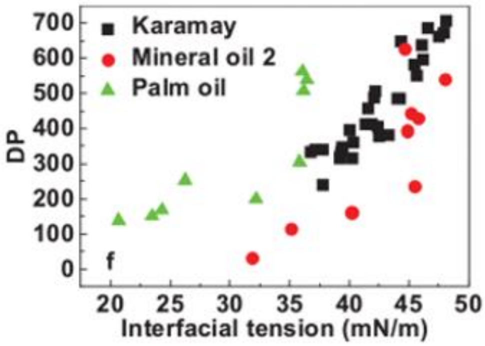
(a)



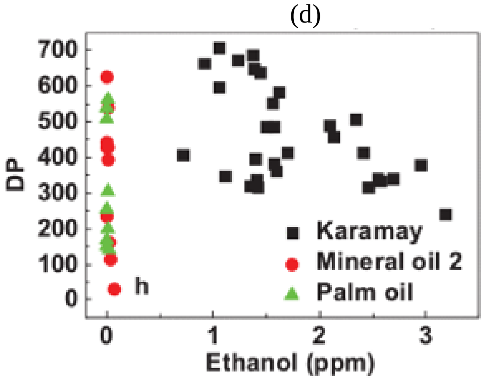
(b)



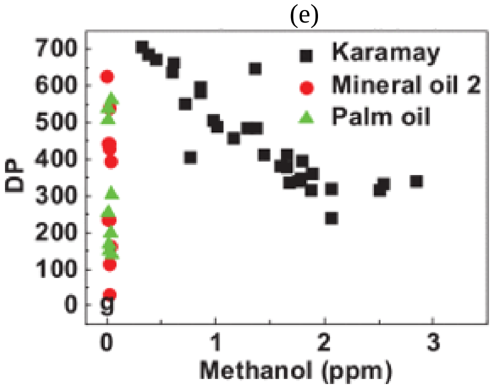
(c)



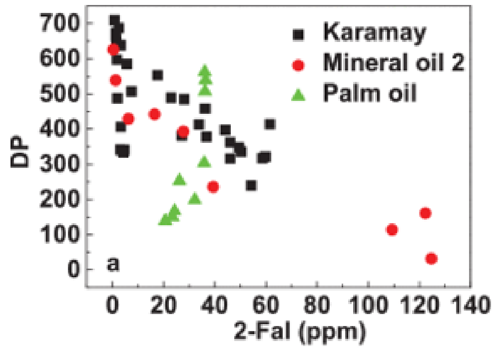
(d)



(e)



(f)



(g)

FIGURE 13 (a) – (g). Ageing parameters and DP [190].

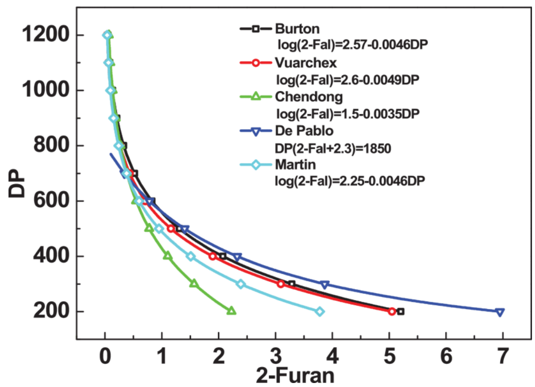


FIGURE 14. 2FAL and DP based on some models[190].

TABLE 11. Summary of models used for ageing parameters prediction with metrics used.

Authors	Year	Model	Optimization technique	Prediction	Evaluation metrics	Ref.
Shaban et al.	2016	kNN	Wrapper method	Furan level	Not stated	[204]
Kari et al.	2018	MKF-SVR	Not stated	Oil DGA	MAPE and R^2	[201]
Li et al.	2019	FCM-LR	Not stated	Paper DP	R^2, MSE	[190]
Hu et al.	2020	LSTM	BOA	Oil DGA	MAE and MAPE	[84]
Zhang et al	2020	SVM	GA	Moisture	Not stated	[117]
Liu et al.	2020	SVM	GA	Moisture	Not stated	[118]
Aciu et al.	2021	RBFN & FFNN	Not stated	DGA	MSE	[205]
Zhang et al.	2021	AHM	Not stated	Paper DP	MSE	[206]
Zhang et al.	2022	XGBoost	BOA	DGA	Not stated	[126]
Wu et al.	2022	SVM	GA	Ethanol, methanol	MSE	[207]
Liu et al.	2022	BPNN	AdaBoost	Moisture	MAE	[1]
Jin et al.	2023	BPNN & SVM	Not stated	DGA	Not stated	[188]

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3391823

Du et al.	2023	BPNN	Not stated	Paper DP	Not stated	[44]
Thango et al.	2023	ANFIS	Not stated	Cellulose DP	MAE, MAPE, RMSE	[203]
Zhong et al.	2023	HATT-RLSTM	Not stated	DGA	MAE, RMSE	[208]
Kunakorn et al.	2023	kNN	Not stated	DP & moisture	Not stated	[209]
Jiang et al.	2023	Reinforced learning	GA	Ageing state	ARE	[45]
Malik et al.	2023	FL	Not stated	DGA	Not stated	[210]

E. COMPARISON OF DEGRATION PARAMETERS

Based on our earlier findings, Table 12 provides a feature comparison of the impacts of various degradation parameters on the ageing of insulation paper.

TABLE 12. Some highlights of the effect of different degradation parameters.

Degradation factors	Effects on insulation paper	Observations	Ref.
Temperature	- Accelerates degradation, especially above 130 ℃ and pyrolysis becomes the predominant degradation process above 140 ℃ - Leads to increased overheating and conductivity, causing localized carbonization. - Governs RUL based on hot-spot temperature	- Activation energy for pyrolytic degradation is 1.4 to 2 times greater than hydrolysis. - Insulation paper from uppermost winding has the lowest DP due to temperature difference. - Hot-spot temperature estimation is complex, often leading to underestimation.	[47], [151]-[155]
Moisture	- Accelerates ageing with a degradation rate 20 times higher at 4% moisture than at 0.5%. - Promotes degradation at 70 – 130 ℃ in the presence of moisture. - Facilitates the migration of ions within the paper, leading to enhanced conductivity.	- Mechanical strength decreases by half when the moisture level doubles. - Moisture content below 0.5% and 20 ppm in insulating paper and liquid respectively is recommended.	[16] [32], [151], [167], [168]
Oxygen	- Accelerates degradation by a factor of 2.5. - Increases acidity, especially below 60 - 70 ℃, impacting on unloaded transformers. - Accelerates the formation of the carbonyl group within the cellulose fibres. - Leads to the breakdown of polymer chains, reducing the mechanical properties of insulation paper.	- Initiates oxidative degradation in the presence of water and catalytic metal ions. - Oxygen effects are integrated into parameters like moisture and temperature.	[47], [182], [185]
Insulating liquid condition	- The presence of sludge, acids, and gases in the insulating liquid accelerates the degradation of insulation paper.	- DGA is a crucial tool for detecting early-stage faults in transformers. - ML techniques such as ANN and SVM, provide superior diagnostics and prognostics capabilities when integrated with DGA.	[194]-[197]

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3391823

VII. DISCUSSION AND FUTURE DIRECTIONS

Among researchers and industries around the globe, the operation of ML technologies and their applications have become widespread as they provide solutions to numerous industrial difficulties. In this regard, they can substitute classical prognostic and diagnostic techniques leading to smart monitoring systems [211]. It is generally acknowledged that the effective selection of features significantly influences the performance of models as nearly all models invest a substantial amount of time and computational resources in the process of feature selection. Therefore, automatic feature extraction should be looked into as a scope for researchers. Furthermore, parameters, being the foundation of every ML algorithm should be considered as a penalty factor and the learning rate is the only parameter tuned to achieve better results. Following this, other imperative parameters should be explored for tuning to enhance the model's performance.

The sensor, which is a micro-electro-mechanical system operating within the same environment as the equipment, is anticipated to undergo degradation with time. Therefore, considering the critical role of precise information in predicting the health condition of power transformers insulating paper and making informed decisions thereafter. It is imperative to carefully account for the impact of sensor degradation and to actively manage and address the degradation of these sensors as differentiating sensor degradation from that of the main component will be necessary to precisely assess the health condition of the component at specific intervals.

An attention-based model like the transformer model allows the model to identify, prioritize, and concentrate on the most crucial features in the input by enabling them to assign importance to each input element while generating the output, thereby eliminating the constraint of gradient vanishing when handling long-term sequences. Therefore, this recent neural network architecture model can be employed in transformer insulation paper health prognosis as it is rarely used in this field. Furthermore, data obtained from power transformers encounters consequential background noise that is not connected to faults, this could pose some challenges in isolating fault-related features leading to inadequacy of accurate prediction. Therefore, priority should be placed on the identification of key fault signals by utilizing the attention mechanisms. Also, a variety of multiple kernel function types and employing different optimization algorithms can be explored to enhance prediction performance.

The intricate structure of power transformers may pose challenges in identifying the most suitable locations for sensors to accurately capture degradation signals for insulation systems. Therefore, a future study can focus on researching the optimal sensor placement methods, which could probably be achieved through simulation and some modeling methods. Also, the new approach can be automated and integrated into upcoming online condition monitoring sensors for power equipment, providing real-time insights into the status of paper insulation.

Probabilistic operations warranted for energy reliability evaluation and diagnostics (POWERED), which is a machine learning or hybrid artificial intelligence tool to diagnose and predict the RUL of power transformers have been presented in some studies. This system can be integrated with physics-based machine learning to aid accurate prediction of the health and status of a transformer insulation system. Also, the authors proposed the development of software that will serve as a valuable resource for the transformer insulation condition monitoring community, where a standardized platform will be created for engineers and researchers to simulate the behaviour and operation of the transformer insulation system. Furthermore, this will be able to provide researchers with the platform to compare and benchmark their algorithms through the provision of datasets for transformer insulation system health monitoring and prognostics.

VIII. CONCLUSION

Incorporating prognostics for power equipment holds the promise of substantially enhancing equipment management within the power industry as the volume of data obtained from power equipment increases with a reduced cost of sensors and storage. Online prognostics help to predict the future health state of a given equipment and present distinct advantages relative to total dependence on the judgement of an expert. However, the primary challenge in maintenance tasks for power transformers is conducting failure analysis for each component. Therefore, in this study, from the perspective of machine learning, the prediction of power transformer insulation paper is reviewed, challenged, and prospected. The methods of estimation of the degree of polymerization of insulating paper were introduced, and the prognostics approaches and how it is being done are summarized. Also, the influencing parameters are elaborately discussed based on mathematical models and machine learning models. Finally, potential directions for future research are prospected to further stimulate case studies, research, and industrial utilization.

APPENDIX A
LIST OF ABBREVIATIONS

Acronyms	Meaning	Acronyms	Meaning
2FAL	2-Fulfural	LR	Linear regression
AE	Auto-encoder	LSTM	Long short-term memory
AFS	Asset fault signature	LSSVR	Least-square support vector regression

AG	Advisory generation	MAE	Mean absolute error
ANFIS	Adaptive neuro-fuzzy inference system	MAPE	Mean absolute percentage error
ANN	Artificial neural network	MKF	Mixed kernel function
ARE	Average relative error	ML	Machine learning
BOA	Bayesian optimization algorithm	MLP	Multilayer perceptron
BPNN	Backpropagation neural network	MSE	Mean square error
CNN	Convolutional neural network	PCC	Pearson correlation coefficient
DA	Diagnostic advisor	PCA	Principal component analysis
DGA	Dissolved gas analysis	PDC	Polarization and depolarization currents
DP	Degree of polymerization	PDF	Probability density function
DSC	Dispersion staining colours	PHM	Prognostics and health management
EOL	End of useful life	PNN	Probabilistic neural network
EPRI	Electric power research institute	R^2	Coefficient of determination
FCM	Fuzzy C mean	RAE	Relative absolute error
FDS	Frequency domain spectroscopy	RLSTM	Recurrent long short time memory
FFOA	Fly fruit optimization algorithm	RMSE	Root mean squared error
FL	Fuzzy logic	RMSLE	Root mean squared logarithmic error
FPCA	Functional principal component analysis	RUL	Remaining useful life
FNN	Fleet wide	RULA	Remaining useful life advisor
GA	Genetic algorithm	RULD	Remaining useful life database
GM	Grey model	RNN	Recurrent neural network
GBN	Gaussian Bayesian network	RVM	Relevance vector machine
GPR	Gaussian process regression	SF	Scoring function
GRA	Grey relational analysis	SMAPE	Symmetric mean absolute error
GRNN	Generalized regression neural network	SVM	Support vector machine
HATT	Hierarchical attention network	SVR	Support vector regression
KPCA	Kernel principal component analysis	TUP	Thermally Upgraded paper
kNN	k nearest neighbour	XGBoost	Extreme gradient boosting

APPENDIX B
LIST OF ABBREVIATIONS

Symbols	Meaning	Symbols	Meaning
R	Gas constant	F_{EQA}	Equivalent ageing factor
A	Pre-exponential factor	L	Loss of life
E_a	Activation energy	V	Paper rate of ageing
K	Rate of reaction	N	Number of samples or units
t	Time	y_i	Real or actual value
T	Temperature	$\hat{y}_i$	Predicted value
$\%LL$	Percent of life lost	$\bar{y}$	Average value of y_i
θ_{HS}	Hot-spot temperature	m	Winding exponent
θ_{TO}	Temperature of the top oil	n	Oil exponent
$\Delta\theta_{H,R}$	Hottest-spot temperature rise at the rated load	$\Delta\theta_{TO,H}$	Temperature rise of the hottest spot over top-oil temperature
ω_t	Uncertainty of the degradation process	$\Delta\theta_{A,TO}$	Temperature rise of top oil over ambient temperature
θ_A	Ambient temperature	y	Exponential constant of the windings
$\tau_{W,R}$	Winding time constant	i_r	Rated load
θ_{OW}	Oil-to-winding temperature gradient	$i(t)$	Load at an instant t
$\tau_{TO,R}$	Oil time constant	k	thermal constants of the transformer
H_0	Humidity of the insulating oil	K_L	Load factor
φ_i	Measurement error for load	φ_{TO}	Measurement error for top oil
$\Delta\theta_{TO,H_i}$	Hot-spot temperature rise over ambient temperature	$\Delta\theta_{A,TO_i}$	Top-oil temperature rise over ambient temperature
n_t	Number of bond scission at t	n_0	Initial number of links available for degradation

REFERENCES

[1] Liu, J., Ding, Z., Fan, X., Geng, C., Song, B., Wang, Q., & Zhang, Y. (2022). A BPNN model-based AdaBoost algorithm for estimating inside moisture of oil–paper insulation of power transformer. *IEEE Transactions on Dielectrics and Electrical Insulation*, *29*(2), 614-622.

[2] A. Adekunle and S. Oparanti, "A Review on Physicochemical and Electrical Performance of Vegetable Oil-Based Nanofluids for High Voltage Equipment," *Electric Power Systems Research,* vol. 214, p. 108873, 2023.

[3] Chen, Q., Sun, W., Cheng, S., & Huang, G. (2023). A review on a novel method for aging evaluation of transformer insulating paper based on methanol. *IET Generation, Transmission & Distribution*, *17*(9), 1955-1971.

[4] I. Power and E. Society, "Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage Regulators," ed: IEEE Standards Association New York, NY, USA, 2012.

[5] C. Oria, C. Méndez, I. Carrascal, D. Ferreño, and A. Ortiz, "Degradation of the compression strength of spacers made of high-density pressboard used in power transformers under the influence of thermal ageing," *Cellulose,* pp. 1-20, 2023.

[6] A. Abu-Siada and S. Islam, "A new approach to identify power transformer criticality and asset management decision based on dissolved gas-in-oil analysis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 19, no. 3, pp. 1007-1012, 2012.

[7] A. A. Adekunle, S. O. Oparanti, and I. Fofana, "Performance Assessment of Cellulose Paper Impregnated in Nanofluid for Power Transformer Insulation Application: A Review," *Energies,* vol. 16, no. 4, p. 2002, 2023.

[8] Sakib, N., & Wuest, T. (2018). Challenges and opportunities of condition-based predictive maintenance: a review. *Procedia Cirp*, *78*, 267-272.

[9] Y. Wen, M. F. Rahman, H. Xu, and T.-L. B. Tseng, "Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective," *Measurement,* vol. 187, p. 110276, 2022.

[10] R. Ahmad and S. Kamaruddin, "An overview of time-based and condition-based maintenance in industrial application," *Computers & industrial engineering,* vol. 63, no. 1, pp. 135-149, 2012.

[11] H. Ghasemi, E. S. Farahani, M. Fotuhi-Firuzabad, P. Dehghanian, A. Ghasemi, and F. Wang, "Equipment failure rate in electric power distribution networks: An overview of concepts, estimation, and modeling methods," *Engineering Failure Analysis,* p. 107034, 2023.

[12] IEEE, " IEEE Guide for the Evaluation and Reconditioning of Liquid Immersed Power Transformer," *IEEE Std. C57.140,* 2006.

[13] Wang, Y. B., Chang, D. G., Fan, Y. H., Zhang, G. J., Zhan, J. Y., Shao, X. J., & He, W. L. (2017). Acoustic localization of partial discharge sources in power transformers using a particle-swarm-optimization-route-searching algorithm. *IEEE Transactions on Dielectrics and Electrical Insulation*, *24*(6), 3647-3656.

[14] P. M. Sekatane and P. Bokoro, "Partial Discharge Localization through k-NN and SVM," *Energies,* vol. 16, no. 21, p. 7430, 2023.

[15] S. Y. Wong, X. Ye, F. Guo, and H. H. Goh, "Computational intelligence for preventive maintenance of power transformers," *Applied Soft Computing,* vol. 114, p. 108129, 2022.

[16] C. Oria, A. Ortiz, D. Ferreño, I. Carrascal, and I. Fernández, "State-of-the-art review on the performance of cellulosic dielectric materials in power transformers: Mechanical response and ageing," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 26, no. 3, pp. 939-954, 2019.

[17] O. Girlanda, D. Tjahjanto, S. Östlund, K. Wei, and L. Schmidt, "Characterization and modelling of the mechanical properties of pressboard," in *2013 Annual Report Conference on Electrical Insulation and Dielectric Phenomena*, 2013: IEEE, pp. 563-566.

[18] P. Jusner, E. Schwaiger, A. Potthast, and T. Rosenau, "Thermal stability of cellulose insulation in electrical power transformers–A review," *Carbohydrate Polymers,* vol. 252, p. 117196, 2021.

[19] O. Girlanda, K. Wei, L. E. Schmidt, M. Evenbom, and J. Forslín, "Static and quasi-static behavior of dry and oil-impregnated pressboard," in *2016 IEEE Electrical Insulation Conference (EIC)*, 2016: IEEE, pp. 105-108.

[20] L. E. Lundgaard, W. Hansen, D. Linhjell, and T. J. Painter, "Aging of oil-impregnated paper in power transformers," *IEEE Transactions on power delivery,* vol. 19, no. 1, pp. 230-239, 2004.

[21] C. Tang, S. Zhang, J. Xie, and C. Lv, "Molecular simulation and experimental analysis of Al2O3-nanoparticle-modified insulation paper cellulose," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 24, no. 2, pp. 1018-1026, 2017.

[22] Arsad, S. R., Ker, P. J., Jamaludin, M. Z., Choong, P. Y., Lee, H. J., Thiviyanathan, V. A., & Yang Ghazali, Y. Z. (2023). Water content in transformer insulation system: A review on the detection and quantification methods. *Energies*, *16*(4), 1920.

[23] G. T. 29305-2012, "Measurement of the average viscometric degree of polymerization of new and aged cellulosic electrically insulating materials " 2012.

[24] M. Chi, R. Xia, Y. Guan, W. Sun, J. Liu, and Q. Chen, "Insulation degradation and fracture cracks of mechanical impacted thermal-ageing oil-impregnated papers," *High Voltage,* vol. 8, no. 1, pp. 21-30, 2023.

[25] M. Ariannik, A. A. Razi-Kazemi, and M. Lehtonen, "An approach on lifetime estimation of distribution transformers based on degree of polymerization," *Reliability Engineering & System Safety,* vol. 198, p. 106881, 2020.

[26] J. I. Aizpurua, B. G. Stewart, S. D. McArthur, B. Lambert, J. G. Cross, and V. M. Catterson, "Improved power transformer condition monitoring under uncertainty through soft computing and probabilistic health index," *Applied Soft Computing,* vol. 85, p. 105530, 2019.

[27] D. Martin, Y. Cui, C. Ekanayake, H. Ma, and T. Saha, "An updated model to determine the life remaining of transformer insulation," *IEEE Transactions on power delivery,* vol. 30, no. 1, pp. 395-402, 2014.

[28] A. Emsley and G. Stevens, "Review of chemical indicators of degradation of cellulosic electrical paper insulation in oil-filled transformers," *IEE Proceedings-Science, Measurement and Technology,* vol. 141, no. 5, pp. 324-334, 1994.

[29] Zhou, L., Zhang, J., Li, Z., Wen, R., Wang, D., Wu, J., & Tang, H. (2023). Degree of polymerisation estimation for insulation paper immersed in silicone oil using frequency domain dielectric spectroscopy. *High Voltage*, *8*(5), 1041-1054.

[30] M. Ariannik, A. A. Razi-Kazemi, and M. Lehtonen, "Effect of cumulative moisture content on degradation of transformer paper insulation," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 24, no. 1, pp. 611-618, 2017.

[31] J. V. Costa, D. F. d. Silva, and P. J. C. Branco, "Large-power transformers: Time now for addressing their monitoring and failure investigation techniques," *Energies,* vol. 15, no. 13, p. 4697, 2022.

[32] C. T. Force, "Ageing of cellulose in mineral-oil insulated transformers," in *CIGRE: Paris, France*, 2007.

[33] D. Martin, C. Perkasa, and N. Lelekakis, "Measuring paper water content of transformers: A new approach using cellulose isotherms in nonequilibrium conditions," *IEEE Transactions on Power Delivery,* vol. 28, no. 3, pp. 1433-1439, 2013.

[34] M. Martins, "Study of moisture equilibrium in oil-paper system with temperature variation," in *Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials (Cat. No. 03CH37417)*, 2003, vol. 3: IEEE, pp. 1124-1129.

[35] A. Sbravati, L. V. Cheim, and G. K. Frimpong, "Insulation Residual Life Estimation Based on Transformer Condition Assessment Data," in *2023 IEEE Electrical Insulation Conference (EIC)*, 2023: IEEE, pp. 1-4.

[36] S. Mishra, A. Baral, D. Mishra, and S. Chakravorti, "A novel method to predict severity of thermal aging and degree of polymerization for reliable diagnosis of dry-type insulation," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 29, no. 2, pp. 631-640, 2022.

[37] M. Wen, J. Song, Y. Song, Y. Liu, C. Li, and P. Wang, "Reliability assessment of insulation system for dry type transformers," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 20, no. 6, pp. 1998-2008, 2013.

[38] Liu, J., Jiang, K., Wang, Q., Zhang, H., Zhang, E., Fan, X., & Zhang, Y. (2023). An improved second-order kinetic model for degradation analysis of transformer paper insulation under non-uniform thermal field. *High Voltage*, *8*(1), 81-90.

[39] Li, S., Li, S., Yang, L., Ge, Z., Tan, B., Mao, G., & Guo, Q. (2021). Characteristics and kinetics of cellulose degradation in oil-

impregnated pressboard under simultaneous thermal-electrical-mechanical stress accelerated aging. *IEEE Transactions on Dielectrics and Electrical Insulation*, *28*(4), 1363-1371.

[40] P. Calvini, A. Gorassini, and A. L. Merlani, "On the kinetics of cellulose degradation: looking beyond the pseudo zero order rate equation," *Cellulose,* vol. 15, pp. 193-203, 2008

[41] J. Zhang, F. Wang, J. Li, X. Li, and Q. Zhou, "Ageing characterization of oil impregnated paper insulation based on dispersion staining colors," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 25, no. 5, pp. 1587-1597, 2018.

[42] C. Chen, S. Li, T. He, and Y. Liu, "Concentration of total sugar produced by insulating paper in oil-paper insulation system," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 26, no. 6, pp. 2036-2040, 2019.

[43] Seiffadini, N., Sekongo, B., Meghnefi, F., Lim, K. S., Weng, O. C., Udos, W., ... & Ouhrouche, M. (2023, June). Correlation between the Optical Properties and the Degree of Polymerisation of Transformer Insulation Paper. In *2023 IEEE Electrical Insulation Conference (EIC)* (pp. 1-4). IEEE.

[44] B. Du, Y. Hou, X. Kong, C. Zhang, F. Li, and J. Mu, "Identification of Aging Status of Insulation Paper Based on BP Neural Network," in *2023 IEEE 4th International Conference on Electrical Materials and Power Equipment (ICEMPE)*, 2023: IEEE, pp. 1-4.

[45] Z. Jiang, J. Liu, X. Fan, Q. Wang, Y. Zhang, and T. Wu, "Reinforcement Learning Based Genetic Algorithm for Aging State Analysis of Insulating Paper at Transformer Hotspot," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[46] IEEE, "Guide for Loading Mineral-Oil-Immersed Transformers and STP-Voltage Regulators, IEEE Standard C57.91-2011.

[47] Medina, R. D., Romero, A. A., Mombello, E. E., & Ratta, G. (2017). Assessing degradation of power transformer solid insulation considering thermal stress and moisture variation. *Electric Power Systems Research*, *151*, 1-11.

[48] IEC, "Power Transformers — Part 7: Loading Guide for Oil-ImmersedPower Transformers, IEC 60076-7, 2005.

[49] IEC, "Power Transformers — Part 7: Loading Guide for Mineral-Oil-Immersed Power Transformers, IEC 60076–7:2017," 2017.

[50] M. K. Pradhan and T. Ramu, "On the estimation of elapsed life of oil-immersed power transformers," *IEEE Transactions on Power Delivery,* vol. 20, no. 3, pp. 1962-1969, 2005.

[51] Y. Hu, X. Miao, Y. Si, E. Pan, and E. Zio, "Prognostics and health management: A review from the perspectives of design, development and decision," *Reliability Engineering & System Safety,* vol. 217, p. 108063, 2022.

[52] T. Xia, Y. Dong, L. Xiao, S. Du, E. Pan, and L. Xi, "Recent advances in prognostics and health management for advanced manufacturing paradigms," *Reliability Engineering & System Safety,* vol. 178, pp. 255-268, 2018.

[53] A. Chaturvedi, M. Sarma, S. K. Chaturvedi, and J. Bernstein, "Performance assessment and RUL prediction of power converters under the multiple components degradation," *Microelectronics Reliability,* vol. 144, p. 114958, 2023.

[54] N.-H. Kim, D. An, and J.-H. Choi, "Prognostics and health management of engineering systems," *Switzerland: Springer International Publishing,* 2017.

[55] M. Pecht, "Prognostics and health management of electronics," *Encyclopedia of structural health monitoring,* 2009.

[56] H. M. Elattar, H. K. Elminir, and A. Riad, "Prognostics: a literature review," *Complex & Intelligent Systems,* vol. 2, no. 2, pp. 125-154, 2016.

[57] Aizpurua, J. I., Peña-Alzola, R., Olano, J., Ramirez, I., Lasa, I., del Rio, L., & Dragicevic, T. (2023). Probabilistic machine learning aided transformer lifetime prediction framework for wind energy systems. *International Journal of Electrical Power & Energy Systems*, *153*, 109352.

[58] Y. Hua, Y. Sun, G. Xu, S. Sun, E. Wang, and Y. Pang, "A fault diagnostic method for oil-immersed transformer based on multiple probabilistic output algorithms and improved DS evidence theory," *International Journal of Electrical Power & Energy Systems,* vol. 137, p. 107828, 2022.

[59] K. Mukhopadhyay, B. Liu, T. Bedford, and M. Finkelstein, "Remaining lifetime of degrading systems continuously monitored by degrading sensors," *Reliability Engineering & System Safety,* vol. 231, p. 109022, 2023.

[60] E. Zio, "Prognostics and Health Management (PHM): Where are we and where do we (need to) go in theory and practice," *Reliability Engineering & System Safety,* vol. 218, p. 108119, 2022.

[61] K. Javed, "A robust & reliable Data-driven prognostics approach based on extreme learning machine and fuzzy clustering," Université de Franche-Comté, 2014.

[62] M. Lebold and M. Thurston, "Open standards for condition-based maintenance and prognostic systems," in *Maintenance and reliability conference (marcon)*, 2001, vol. 200.

[63] A. M. Elmashtoly and C.-K. Chang, "Prognostics health management system for power transformer with IEC61850 and internet of things," *Journal of Electrical Engineering & Technology,* vol. 15, no. 2, pp. 673-683, 2020.

[64] M. Payette and G. Abdul-Nour, "Asset Management, Reliability and Prognostics Modeling Techniques," *Sustainability,* vol. 15, no. 9, p. 7493, 2023.

[65] V. Agarwal, N. J. Lybeck, R. Bickford, and R. Rusaw, "Development of asset fault signatures for prognostic and health management in the nuclear industry," in *2014 International Conference on Prognostics and Health Management*, 2014: IEEE, pp. 1-7.

[66] V. Agarwal, N. J. Lybeck, B. Pham, R. Rusaw, and R. Bickford, "Implementation of remaining useful lifetime transformer models in the fleet-wide prognostic and health management suite," Idaho National Lab.(INL), Idaho Falls, ID (United States), 2015.

[67] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation–a review on the statistical data driven approaches," *European journal of operational research,* vol. 213, no. 1, pp. 1-14, 2011.

[68] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mechanical systems and signal processing,* vol. 104, pp. 799-834, 2018.

[69] S. Sankararaman, "Significance, interpretation, and quantification of uncertainty in prognostics and remaining useful life prediction," *Mechanical Systems and Signal Processing,* vol. 52, pp. 228-247, 2015

[70] Aizpurua, J. I., Stewart, B. G., McArthur, S. D., Penalba, M., Barrenetxea, M., Muxika, E., & Ringwood, J. V. (2022). Probabilistic forecasting informed failure prognostics framework for improved RUL prediction under uncertainty: A transformer case study. *Reliability Engineering & System Safety*, *226*, 108676.

[71] J. I. Aizpurua, V. M. Catterson, Y. Papadopoulos, F. Chiacchio, and D. D'Urso, "Supporting group maintenance through prognostics-enhanced dynamic dependability prediction," *Reliability Engineering & System Safety,* vol. 168, pp. 171-188, 2017.

[72] R. Khelif, S. Malinowski, B. Chebel-Morello, and N. Zerhouni, "RUL prediction based on a new similarity-instance based approach," in *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, 2014: IEEE, pp. 2463-2468.

[73] Y. Peng, M. Dong, and M. J. Zuo, "Current status of machine prognostics in condition-based maintenance: a review," *The International Journal of Advanced Manufacturing Technology,* vol. 50, pp. 297-313, 2010.

[74] K. Goebel, "Prognostics and health management," *Guest lecture, ENME 808A University of Maryland,* 2007.

[75] L. Vichard, N. Y. Steiner, N. Zerhouni, and D. Hissel, "Hybrid fuel cell system degradation modeling methods: A comprehensive review," *Journal of Power Sources,* vol. 506, p. 230071, 2021.

[76] F. K. Wang and T. Mamo, "Hybrid approach for remaining useful life prediction of ball bearings," *Quality and Reliability Engineering International,* vol. 35, no. 7, pp. 2494-2505, 2019.

[77] D. Kwon, M. R. Hodkiewicz, J. Fan, T. Shibutani, and M. G. Pecht, "IoT-based prognostics and systems health management for industrial applications," *IEEE access,* vol. 4, pp. 3659-3670, 2016.

[78] Y. Li, T. Kurfess, and S. Liang, "Stochastic prognostics for rolling element bearings," *Mechanical Systems and Signal Processing,* vol. 14, no. 5, pp. 747-762, 2000.

[79] A. Muneer, S. M. Taib, S. M. Fati, and H. Alhussian, "Deep-learning based prognosis approach for remaining useful life prediction of turbofan engine," *Symmetry,* vol. 13, no. 10, p. 1861, 2021.

[80] T. Sutharssan, S. Stoyanov, C. Bailey, and C. Yin, "Prognostic and health management for engineering systems: a review of the data-driven approach and algorithms," *The Journal of engineering,* vol. 2015, no. 7, pp. 215-222, 2015.
[81] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing,* vol. 107, pp. 241-265, 2018.
[82] Z. Li, D. Wu, C. Hu, and J. Terpenny, "An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction," *Reliability Engineering & System Safety,* vol. 184, pp. 110-122, 2019.
[83] M. Pecht and S. Kumar, "Data analysis approach for system reliability, diagnostics and prognostics," in *Pan Pacific Microelectronics Symposium,* 2008, vol. 795: Kauai Hawaii, USA, pp. 1-9.
[84] C. Hu, Y. Zhong, Y. Lu, X. Luo, and S. Wang, "A prediction model for time series of dissolved gas content in transformer oil based on LSTM," in *Journal of Physics: Conference Series*, 2020, vol. 1659, no. 1: IOP Publishing, p. 012030.
[85] N. Khan, I. U. Haq, S. U. Khan, S. Rho, M. Y. Lee, and S. W. Baik, "DB-Net: A novel dilated CNN based multi-step forecasting model for power consumption in integrated local energy systems," *International Journal of Electrical Power & Energy Systems,* vol. 133, p. 107023, 2021.
[86] X. Liu and Y. He, "SD2SDFNet: A novel deep fusion network based on statistical denoising and dual-dimension self-attention model for power transformer hybrid-space fault prognosis," *Digital Signal Processing,* vol. 142, p. 104197, 2023.
[87] Q. Hu, Y. Zhao, and L. Ren, "Novel Transformer-based Fusion Models for Aero-engine Remaining Useful Life Estimation," *IEEE Access,* 2023.
[88] A. L. Smith, T. Zheng, and A. Gelman, "Prediction scoring of data-driven discoveries for reproducible research," *Statistics and Computing,* vol. 33, no. 1, p. 11, 2023.
[89] L. Capogrosso, F. Cunico, D. S. Cheng, F. Fummi, and M. Cristani, "A Machine Learning-oriented Survey on Tiny Machine Learning," *IEEE Access,* 2024.
[90] J. I. Aizpurua, V. M. Catterson, B. G. Stewart, S. D. McArthur, B. Lambert, and J. G. Cross, "Uncertainty-aware fusion of probabilistic classifiers for improved transformer diagnostics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 51, no. 1, pp. 621-633, 2018.
[91] Aizpurua, J. I., Stewart, B. G., McArthur, S. D., Garro, U., Muxika, E., Mendicute, M., ... & del Rio, L. (2019, October). Prognostics & health management methods & tools for transformer condition monitoring in smart grids. In *IEEE 6th International Advanced Research Workshop on Transformers (ARWtr2019)*.
[92] Z.-H. Zhou and J. Feng, "Deep forest," *National science review,* vol. 6, no. 1, pp. 74-86, 2019.
[93] T. Wang, Y. He, T. Shi, J. Tong, and B. Li, "Transformer health management based on self-powered RFID sensor and multiple kernel RVM," *IEEE Transactions on Instrumentation and Measurement,* vol. 68, no. 3, pp. 818-828, 2018.
[94] J. Leiva, J. A. Aguado, Á. Paredes, and P. Arboleya, "Data-driven flexibility prediction in low voltage power networks," *International Journal of Electrical Power & Energy Systems,* vol. 123, p. 106242, 2020.
[95] W. Zhang, X. Yang, Y. Deng, and A. Li, "An inspired machine-learning algorithm with a hybrid whale optimization for power transformer PHM," *Energies,* vol. 13, no. 12, p. 3143, 2020.
[96] C. L. Gan, "Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things: John Wiley & Sons Ltd.(2018). pp. 731, ISBN: 9781119515326 (Print), 9781119515326 (Online)," *Life Cycle Reliability and Safety Engineering,* vol. 9, no. 2, pp. 225-226, 2020.
[97] L. Wang, H. Wang, Y. Di, L. Dong, and G. Jin, "Predicting Sliding Angles on Random Pit-Distributed Textures Using Probabilistic Neural Networks," *Langmuir,* vol. 39, no. 18, pp. 6406-6412, 2023.
[98] M. Pereira and B. Glisic, "Detection and quantification of temperature sensor drift using probabilistic neural networks," *Expert Systems with Applications,* vol. 213, p. 118884, 2023.
[99] G. Xue, Y. Pan, T. Lin, J. Song, C. Qi, and Z. Wang, "District heating load prediction algorithm based on feature fusion LSTM model," *Energies,* vol. 12, no. 11, p. 2122, 2019.
[100] H.-K. Wang, Y. Cheng, and K. Song, "Remaining useful life estimation of aircraft engines using a joint deep learning model based on TCNN and transformer," *Computational Intelligence and Neuroscience,* vol. 2021, 2021.
[101] O. Ogunfowora and H. Najjaran, "A Transformer-based Framework For Multi-variate Time Series: A Remaining Useful Life Prediction Use Case," *arXiv preprint arXiv:2308.09884,* 2023.
[102] J. Lin, G. Sheng, Y. Yan, J. Dai, and X. Jiang, "Prediction of dissolved gas concentrations in transformer oil based on the KPCA-FFOA-GRNN model," *Energies,* vol. 11, no. 1, p. 225, 2018.
[103] L. Hongwei, Q. Binxin, H. Zhicheng, L. Junnan, Y. Yue, and L. Guolong, "An interpretable data-driven method for degradation prediction of proton exchange membrane fuel cells based on temporal fusion transformer and covariates," *International Journal of Hydrogen Energy,* 2023.
[104] K. Zhong, J. Wang, S. Xu, C. Cheng, and H. Chen, "Overview of fault prognosis for traction systems in high-speed trains: A deep learning perspective," *Engineering Applications of Artificial Intelligence,* vol. 126, p. 106845, 2023.
[105] A. Muneer, S. M. Taib, S. Naseer, R. F. Ali, and I. A. Aziz, "Data-driven deep learning-based attention mechanism for remaining useful life prediction: Case study application to turbofan engine analysis," *Electronics,* vol. 10, no. 20, p. 2453, 2021.
[106] K. Hong, M. Jin, and H. Huang, "Transformer winding fault diagnosis using vibration image and deep learning," *IEEE Transactions on Power Delivery,* vol. 36, no. 2, pp. 676-685, 2020.
[107] Y. Ding and M. Jia, "Convolutional transformer: An enhanced attention mechanism architecture for remaining useful life estimation of bearings," *IEEE Transactions on Instrumentation and Measurement,* vol. 71, pp. 1-10, 2022.
[108] Z. Zhang, W. Song, and Q. Li, "Dual-aspect self-attention based on transformer for remaining useful life prediction," *IEEE Transactions on Instrumentation and Measurement,* vol. 71, pp. 1-11, 2022.
[109] Y. Wang, Y. Zhao, and S. Addepalli, "Remaining useful life prediction using deep learning approaches: A review," *Procedia manufacturing,* vol. 49, pp. 81-88, 2020.
[110] Gu, X., See, K. W., Li, P., Shan, K., Wang, Y., Zhao, L., ... & Zhang, N. (2023). A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. *Energy*, *262*, 125501.
[111] Sharma, T., Chehri, A., Fofana, I., Jadhav, S., Khare, S., Debaque, B., ... & Arya, D. (2024). Deep Learning-Based Object Detection and Classification for Autonomous Vehicles in Different Weather Scenarios of Quebec, Canada. *IEEE Access*.
[112] A. H. Alshalawi and F. S. Al-Ismail, "Partial Discharge Detection Based on Ultrasound Using Optimized Deep Learning Approach," *IEEE Access*, 2024.
[113] J. Dai, H. Song, G. Sheng, and X. Jiang, "LSTM networks for the trend prediction of gases dissolved in power transformer insulation oil," in *2018 12th International Conference on the Properties and Applications of Dielectric Materials (ICPADM)*, 2018: IEEE, pp. 666-669.
[114] Z. Zhang, Y. Jeong, J. Jang, and C.-G. Lee, "A pattern-driven stochastic degradation model for the prediction of remaining useful life of rechargeable batteries," *IEEE Transactions on Industrial Informatics,* vol. 18, no. 12, pp. 8586-8594, 2022.
[115] K. T. Nguyen, K. Medjaher, and D. T. Tran, "A review of artificial intelligence methods for engineering prognostics and health management with implementation guidelines," *Artificial Intelligence Review,* vol. 56, no. 4, pp. 3659-3709, 2023.
[116] Y. Hu, L. Zhang, Y. Jiang, K. Peng, and Z. Jin, "A Hybrid Method for Performance Degradation Probability Prediction of Proton Exchange Membrane Fuel Cell," *Membranes,* vol. 13, no. 4, p. 426, 2023.
[117] Y. Zhang, J. Li, X. Fan, J. Liu, and H. Zhang, "Moisture prediction of transformer oil-immersed polymer insulation by applying a support vector machine combined with a genetic algorithm," *Polymers,* vol. 12, no. 7, p. 1579, 2020.
[118] Liu, J., Fan, X., Zhang, C., Lai, C. S., Zhang, Y., Zheng, H., ... & Zhang, E. (2020). Moisture diagnosis of transformer oil-immersed

insulation with intelligent technique and frequency-domain spectroscopy. *IEEE Transactions on Industrial Informatics*, *17*(7), 4624-4634.

[119] S. Zhang, Y. Bai, G. Wu, and Q. Yao, "The forecasting model for time series of transformer DGA data based on WNN-GNN-SVM combined algorithm," in *2017 1st International Conference on Electrical Materials and Power Equipment (ICEMPE)*, 2017: IEEE, pp. 292-295..

[120] M. A. Faraji, A. Shooshtari, and A. El-Hag, "Stacked Ensemble Regression Model for Prediction of Furan," *Energies*, vol. 16, no. 22, p. 7656, 2023.

[121] U. G. Inyang, F. F. Ijebu, F. B. Osang, A. A. Afolorunso, S. S. Udoh, and I. J. Eyoh, "A Dataset-Driven Parameter Tuning Approach for Enhanced K-Nearest Neighbour Algorithm Performance," *International Journal on Advanced Science, Engineering & Information Technology*, vol. 13, no. 1, 2023.

[122] Aizpurua, J. I., Catterson, V. M., Stewart, B. G., McArthur, S. D., Lambert, B., Ampofo, B., ... & Cross, J. G. (2018). Power transformer dissolved gas analysis through Bayesian networks and hypothesis testing. *IEEE Transactions on Dielectrics and Electrical Insulation*, *25*(2), 494-506.

[123] J. I. Aizpurua, B. G. Stewart, and S. D. McArthur, "Prognostics and health management oriented data analytics suite for transformer health monitoring," *Transformers Magazine*, vol. 9, no. SE2, pp. 66-75, 2022.

[124] Aizpurua, J. I., Garro, U., Muxika, E., Mendicute, M., Gilbert, I. P., Stewart, B. G., ... & Lambert, B. (2019, October). Probabilistic power transformer condition monitoring in smart grids. In *2019 6th International Advanced Research Workshop on Transformers (ARWtr)* (pp. 42-47). IEEE.

[125] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785-794.

[126] Zhang, D., Li, C., Shahidehpour, M., Wu, Q., Zhou, B., Zhang, C., & Huang, W. (2022). A bi-level machine learning method for fault diagnosis of oil-immersed transformers with feature explainability. *International Journal of Electrical Power & Energy Systems*, *134*, 107356.

[127] H. Oukhouya and K. El Himdi, "Comparing machine learning methods—svr, xgboost, lstm, and mlp—for forecasting the moroccan stock market," in *Computer Sciences & Mathematics Forum*, 2023, vol. 7, no. 1: MDPI, p. 39. [30] M. Ariannik, A. A. Razi-Kazemi, and M. Lehtonen, "Effect of cumulative moisture content on degradation of transformer paper insulation," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 1, pp. 611-618, 2017.

[128] Sudha, M., & Mehertaj, S. Exploring the Gradient Boosting and LSTM for Power Distribution-based Time Series Analysis.

[129] T. Wang, Y. He, Q. Luo, F. Deng, and C. Zhang, "Self-powered RFID sensor tag for fault diagnosis and prognosis of transformer winding," *IEEE Sensors Journal*, vol. 17, no. 19, pp. 6418-6430, 2017.

[130] S. Raj and K. C. Ray, "ECG signal analysis using DCT-based DOST and PSO optimized SVM," *IEEE Transactions on instrumentation and measurement*, vol. 66, no. 3, pp. 470-478, 2017.

[131] S. Meisenbacher and P. Norlander, "Transforming Unstructured Text into Data with Context Rule Assisted Machine Learning (CRAML)," *arXiv preprint arXiv:2301.08549*, 2023.

[132] M. Wang, "Grey-extension method for incipient fault forecasting of oil-immersed power transformer," *Electric Power Components and Systems*, vol. 32, no. 10, pp. 959-975, 2004.

[133] S. Liu, Y. Yang, and J. Forrest, "Grey data analysis," *Springer Singapore, Singapore, Doi*, vol. 10, no. 1007, pp. 978-981, 2017.

[134] F. A. Viana and A. K. Subramaniyan, "A survey of Bayesian calibration and physics-informed neural networks in scientific modeling," *Archives of Computational Methods in Engineering*, vol. 28, pp. 3801-3830, 2021.

[135] Z. Hajirahimi and M. Khashei, "Hybrid structures in time series modeling and forecasting: A review," *Engineering Applications of Artificial Intelligence*, vol. 86, pp. 83-106, 2019.

[136] M. A. Chao, C. Kulkarni, K. Goebel, and O. Fink, "Fusing physics-based and deep learning models for prognostics," *Reliability Engineering & System Safety*, vol. 217, p. 107961, 2022.

[137] L. Liao and F. Köttig, "Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction," *IEEE Transactions on Reliability*, vol. 63, no. 1, pp. 191-207, 2014.

[138] X. Su, H. Liu, L. Tao, C. Lu, and M. Suo, "An end-to-end framework for remaining useful life prediction of rolling bearing based on feature pre-extraction mechanism and deep adaptive transformer model," *Computers & Industrial Engineering*, vol. 161, p. 107531, 2021.

[139] X. Fang, R. Zhou, and N. Gebraeel, "An adaptive functional regression-based prognostic model for applications with missing data," *Reliability Engineering & System Safety*, vol. 133, pp. 266-274, 2015.

[140] K. Kim, G. Parthasarathy, O. Uluyol, W. Foslien, S. Sheng, and P. Fleming, "Use of SCADA data for failure detection in wind turbines," in *Energy sustainability*, 2011, vol. 54686, pp. 2071-2079.

[141] M. A. Djeziri, S. Benmoussa, and R. Sanchez, "Hybrid method for remaining useful life prediction in wind turbine systems," *Renewable Energy*, vol. 116, pp. 173-187, 2018.

[142] J.-z. Kong, F. Yang, X. Zhang, E. Pan, Z. Peng, and D. Wang, "Voltage-temperature health feature extraction to improve prognostics and health management of lithium-ion batteries," *Energy*, vol. 223, p. 120114, 2021.

[143] I. Jebli, F.-Z. Belouadha, M. I. Kabbaj, and A. Tilioua, "Prediction of solar energy guided by pearson correlation using machine learning," *Energy*, vol. 224, p. 120109, 2021.

[144] D.-N. Le, G. N. Nguyen, V. Bhateja, and S. C. Satapathy, "Optimizing feature selection in video-based recognition using Max–Min Ant System for the online video contextual advertisement user-oriented system," *Journal of computational science*, vol. 21, pp. 361-370, 2017.

[145] Y. Liu, J. Chen, Y. Chang, S. He, and Z. Zhou, "A novel integration framework for degradation-state prediction via transformer model with autonomous optimizing mechanism," *Journal of Manufacturing Systems*, vol. 64, pp. 288-302, 2022.

[146] D. Xu, X. Xiao, J. Liu, and S. Sui, "Spatio-temporal degradation modeling and remaining useful life prediction under multiple operating conditions based on attention mechanism and deep learning," *Reliability Engineering & System Safety*, vol. 229, p. 108886, 2023.

[147] S. Das, A. Paramane, U. M. Rao, and P. Rozga, "A Hybrid Regression Model to Estimate Remaining Useful Life of Transformer Liquid," *IEEE Transactions on Dielectrics and Electrical Insulation*, 2023.

[148] Sayadi, A., Mahi, D., Fofana, I., Bessissa, L., Bessedik, S. A., Arroyo-Fernandez, O. H., & Jalbert, J. (2023). Modeling and predicting the mechanical behavior of standard insulating kraft paper used in power transformers under thermal aging. *Energies*, *16*(18), 6455.

[149] J. Lv, Z. Yu, H. Zhang, G. Sun, P. Muhl, and J. Liu, "Transformer Based Long-Term Prognostics for Dynamic Operating PEM Fuel Cells," *IEEE Transactions on Transportation Electrification*, 2023.

[150] Y. Zhang, K. Feng, J. Ji, K. Yu, Z. Ren, and Z. Liu, "Dynamic model-assisted bearing remaining useful life prediction using the cross-domain transformer network," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 2, pp. 1070-1080, 2022.

[151] M. Muhr and C. Sumereder, "Transformer condition determination methods," in *2008 International Symposium on Electrical Insulating Materials (ISEIM 2008)*, 2008: IEEE, pp. 479-482.

[152] Vasovic, V., Lukic, J., Mihajlovic, D., Pejovic, B., Milovanovic, M., Radoman, U., & Radakovic, Z. (2019). Aging of transformer insulation of experimental transformers and laboratory models with different moisture contents: Part II—Moisture distribution and aging kinetics. *IEEE Transactions on Dielectrics and Electrical Insulation*, *26*(6), 1847-1852.

[153] Y. Cui, H. Ma, T. Saha, C. Ekanayake, and D. Martin, "Moisture-dependent thermal modelling of power transformer," *IEEE Transactions on Power Delivery*, vol. 31, no. 5, pp. 2140-2150, 2016.

[154] J. I. Aizpurua, S. D. McArthur, B. G. Stewart, B. Lambert, J. G. Cross, and V. M. Catterson, "Adaptive power transformer lifetime predictions through machine learning and uncertainty modeling in nuclear power plants," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 6, pp. 4726-4737, 2018.

[155] I. S. Association, "IEEE guide for loading mineral-oil-immersed transformers and step-voltage regulators," *URL(http://ieeexplore. ieee. org/stamp/stamp. jsp*, 2012..

[156] N. Lelekakis, W. Guo, D. Martin, J. Wijaya, and D. Susa, "A field study of aging in paper-oil insulation systems," *IEEE Electrical Insulation Magazine*, vol. 28, no. 1, pp. 12-19, 2012.

[157] D. Feng, Z. Wang, and P. Jarman, "Modeling thermal life expectancy of the UK transmission power transformers," in *2012 International Conference on High Voltage Engineering and Application*, 2012: IEEE, pp. 540-543.

[158] P. Transformers—Part, "7: loading guide for oil-immersed power transformers," *IEC standard*, vol. 60076, no. 7, 2005.

[159] S. P. Patrick Picher, Clemence Tissot, Antoine Mailhot, "Dynamic thermal modeling and overload testing for optimization of transformer loadability in the context of energy transition," presented at the 2022 CIGRE Canada Conference & Expo, 2022.

[160] CIGRE WG A2.38, "Transformer thermal modelling " presented at the CIGRE, Technical Brochure 659, 2016..

[161] D. Susa, M. Lehtonen, and H. Nordman, "Dynamic thermal modelling of power transformers," *IEEE transactions on Power Delivery*, vol. 20, no. 1, pp. 197-204, 2005.

[162] R. D. Medina, A. A. Romero, E. E. Mombello, and G. Ratta, "Comparative study of two thermal aging estimating methods for power transformers," *IEEE Latin America Transactions*, vol. 13, no. 10, pp. 3287-3293, 2015.

[163] S. Cvijic, N. Gupta, and S. Lux, "Need for AI in Transformer Diagnostics and Prognostics," in *2023 Annual Reliability and Maintainability Symposium (RAMS)*, 2023: IEEE, pp. 1-7.

[164] I. IEC, "Power transformers–part 7: loading guide for oil-immersed power transformers," *IEC: Geneva, Switzerland*, 2018.

[165] V. Catterson, J. Melone, and M. S. Garcia, "Prognostics of transformer paper insulation using statistical particle filtering of on-line data," *IEEE Electrical Insulation Magazine*, vol. 32, no. 1, pp. 28-33, 2015.

[166] J. Wang, R. Liao, L. Cheng, and Z. Cheng, "Effects of winding vibration on the mechanical-thermal aging properties of insulating paper," *IEEE Access*, vol. 8, pp. 67912-67920, 2020.

[167] O. Rahman, T. Islam, N. Khera, and S. A. Khan, "A novel application of the cross-capacitive sensor in real-time condition monitoring of transformer oil," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-12, 2021.

[168] M. Arshad and S. M. Islam, "Significance of cellulose power transformer condition assessment," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 18, no. 5, pp. 1591-1598, 2011.

[169] C. M. Banerjee, A. Baral, and S. Chakravorti, "Influence of de-trapped charge polarity in sensing health of power transformer insulation," *IEEE Sensors Journal*, vol. 22, no. 7, pp. 6706-6716, 2022.

[170] S. A. Wani, A. Murugan, R. Sarathi, and V. Subramanian, "Measurement of moisture in transformer insulation using the intelligent high frequency sensor system," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[171] D. Martin and T. Saha, "A review of the techniques used by utilities to measure the water content of transformer insulation paper," *IEEE Electrical Insulation Magazine*, vol. 33, no. 3, pp. 8-16, 2017.

[172] T. Gradnik, B. Čuček, and M. Končan-Gradnik, "Temperature and chemical impact on determination of water content in dielectric liquids by capacitive moisture sensors," in *2014 IEEE 18th International Conference on Dielectric Liquids (ICDL)*, 2014: IEEE, pp. 1-5.

[173] D. Martin, T. Saha, C. Perkasa, N. Lelekakis, and T. Gradnik, "Fundamental concepts of using water activity probes to assess transformer insulation water content," *IEEE Electrical Insulation Magazine*, vol. 32, no. 3, pp. 9-16, 2016.

[174] J. Liu, Z. Wang, X. Fan, Y. Zhang, and J. Wang, "A novel curve database for moisture evaluation of transformer oil-immersed cellulose insulation using FDS and exponential decay model," *IEEE Access*, vol. 8, pp. 180728-180737, 2020.

[175] W. Li, L. Yang, Y. Xia, C. Wu, and Z. Sha, "Assessment of the Moisture Content of Oil–Paper Insulation Based on the Nonlinear Coefficient of Low-Frequency–Per-Unit Curve," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 29, no. 6, pp. 2122-2130, 2022.

[176] L. Fan, Y. H. Lee, and D. H. Beardmore, "Mechanism of the enzymatic hydrolysis of cellulose: effects of major structural features of cellulose on enzymatic hydrolysis," *Biotechnology and Bioengineering*, vol. 22, no. 1, pp. 177-199, 1980.

[177] L.-t. Fan, M. M. Gharpuray, Y.-H. Lee, L.-t. Fan, M. M. Gharpuray, and Y.-H. Lee, "Enzymatic hydrolysis," *Cellulose hydrolysis*, pp. 21-119, 1987.

[178] L.-t. Fan, M. M. Gharpuray, Y.-H. Lee, L.-t. Fan, M. M. Gharpuray, and Y.-H. Lee, "Acid hydrolysis of cellulose," *Cellulose hydrolysis*, pp. 121-148, 1987.

[179] L.-t. Fan, M. M. Gharpuray, and Y.-H. Lee, *Cellulose hydrolysis*. Springer Science & Business Media, 2012..

[180] I. Pavasars, J. Hagberg, H. Borén, and B. Allard, "Alkaline degradation of cellulose: mechanisms and kinetics," *Journal of Polymers and the Environment*, vol. 11, pp. 39-47, 2003.

[181] P. Trnka, J. Hornak, P. Prosr, O. Michal, and F. Wang, "Various aging processes in a paper-natural ester insulation system in the presence of copper and moisture," *IEEE Access*, vol. 8, pp. 61989-61998, 2020.

[182] I. Fofana and J. Sabau, "The Service Reliability of Aging Power Transformers," in *Cigré Canada Conference*, 2013, pp. 9-11.

[183] C. M. Gutiérrez, A. O. Fernández, C. J. R. Estébanez, C. O. Salas, and R. Maina, "Understanding the Ageing Performance of Alternative Dielectric Fluids," *IEEE Access*, vol. 11, pp. 9656-9671, 2023.

[184] IEEE., "IEEE C57.104-2019—*Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers*;," *Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA*, 2019.

[185] J. Fabre and A. Pichon, "Deteriorating processes and products of paper in oil," *Application to transformers, CIGR'E*, vol. 137, 1960.

[186] E. Vänskä, T. Vihelä, M. S. Peresin, J. Vartiainen, M. Hummel, and T. Vuorinen, "Residual lignin inhibits thermal degradation of cellulosic fiber sheets," *Cellulose*, vol. 23, pp. 199-212, 2016.

[187] L. Cheim, M. Duval, and S. Haider, "Combined Duval pentagons: A simplified approach," *Energies*, vol. 13, no. 11, p. 2859, 2020.

[188] Y. Jin, H. Wu, J. Zheng, J. Zhang, and Z. Liu, "Power Transformer Fault Diagnosis Based on Improved BP Neural Network," *Electronics*, vol. 12, no. 16, p. 3526, 2023.

[189] I. B. Taha, A. Hoballah, and S. S. Ghoneim, "Optimal ratio limits of rogers' four-ratios and IEC 60599 code methods using particle swarm optimization fuzzy-logic approach," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 27, no. 1, pp. 222-230, 2020.

[190] S. Li, Z. Ge, A. Abu-Siada, L. Yang, S. Li, and K. Wakimoto, "A new technique to estimate the degree of polymerization of insulation paper using multiple aging parameters of transformer oil," *IEEE Access*, vol. 7, pp. 157471-157479, 2019.

[191] G. A. Sudrajad, N. Hariyanto, and B. A. S. Putro, "Comparison of the Degree of Polymerization Model of the Transformer and Lifetime," in *2023 4th International Conference on High Voltage Engineering and Power Systems (ICHVEPS)*, 2023: IEEE, pp. 733-738.

[192] IEC, "*IEC 60599—Mineral Oil-Impregnated Electrical Equipment in Service–Guide to the Interpretation of Dissolved and Free Gases Analysis*," *International Electrotechnical Commission: Geneva, Switzerland*, 2015.

[193] A. J. Patil, A. Singh, and P. Scholar, "A literature review: Traditional and advanced protection schemes of power transformer," *International Journal of Engineering Research and General Science*, vol. 7, no. 2, pp. 6-19, 2019.

[194] R. A. Prasojo, H. Gumilang, Suwarno, N. U. Maulidevi, and B. A. Soedjarno, "A fuzzy logic model for power transformer faults' severity determination based on gas level, gas rate, and dissolved gas analysis interpretation," *Energies*, vol. 13, no. 4, p. 1009, 2020.

[195] A. Abu-Siada, "Improved consistent interpretation approach of fault type within power transformers using dissolved gas analysis and gene expression programming," *Energies*, vol. 12, no. 4, p. 730, 2019.

[196] Pereira, F. H., Bezerra, F. E., Junior, S., Santos, J., Chabu, I., Souza, G. F. M. D., ... & Nabeta, S. I. (2018). Nonlinear autoregressive neural network models for prediction of transformer oil-dissolved gas concentrations. *Energies*, *11*(7), 1691.

[197] M. E. A. Senoussaoui, M. Brahami, and I. Fofana, "Combining and comparing various machine-learning algorithms to improve dissolved gas analysis interpretation," *IET Generation, Transmission & Distribution*, vol. 12, no. 15, pp. 3673-3679, 2018.

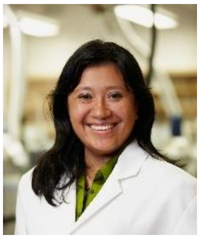
[198] R. A. Prasojo, K. Diwyacitta, Suwarno, and H. Gumilang, "Transformer paper expected life estimation using ANFIS based on oil characteristics and dissolved gases (Case study: Indonesian transformers)," *Energies*, vol. 10, no. 8, p. 1135, 2017. [186] E. Vänskä, T. Vihelä, M. S. Peresin, J. Vartiainen, M. Hummel, and T. Vuorinen, "Residual lignin inhibits thermal degradation of cellulosic fiber sheets," *Cellulose*, vol. 23, pp. 199-212, 2016.

[199] R. A. Prasojo, "Power transformer paper insulation assessment based on oil measurement data using SVM-classifier," *International Journal on Electrical Engineering and Informatics*, vol. 10, no. 4, pp. 661-673, 2018.
[200] S. X. Lu, G. Lin, H. Que, M. J. J. Li, C. H. Wei, and J. K. Wang, "Grey relational analysis using Gaussian process regression method for dissolved gas concentration prediction," *International Journal of Machine Learning and Cybernetics*, vol. 10, pp. 1313-1322, 2019.
[201] T. Kari, W. Gao, A. Tuluhong, Y. Yaermaimaiti, and Z. Zhang, "Mixed kernel function support vector regression with genetic algorithm for forecasting dissolved gas content in power transformers," *Energies*, vol. 11, no. 9, p. 2437, 2018.
[202] L. Jin, D. Kim, K. Y. Chan, and A. Abu-Siada, "Deep Machine Learning-Based Asset Management Approach for Oil-Immersed Power Transformers Using Dissolved Gas Analysis," *IEEE Access,* vol. 12, pp. 27794-27809, 2024.
[203] B. A. Thango and P. N. Bokoro, "A Technique for Transformer Remnant Cellulose Life Cycle Prediction Using Adaptive Neuro-Fuzzy Inference System," *Processes,* vol. 11, no. 2, p. 440, 2023
[204] K. B. Shaban, A. H. El-Hag, and K. Benhmed, "Prediction of transformer furan levels," *IEEE Transactions on Power Delivery,* vol. 31, no. 4, pp. 1778-1779, 2016.
[205] A.-M. Aciu, C.-I. Nicola, M. Nicola, and M.-C. Nițu, "Complementary analysis for DGA based on Duval methods and furan compounds using artificial neural networks," *Energies*, vol. 14, no. 3, p. 588, 2021.
[206] Zhang, E., Zheng, H., Zhang, C., Wang, J., Shi, K., Guo, J., ... & Zhang, C. (2021). Aging state assessment of transformer cellulosic paper insulation using multivariate chemical indicators. *Cellulose*, *28*, 2445-2460.
[207] Wu, S., Zhang, H., Wang, Y., Luo, Y., He, J., Yu, X., ... & Shuang, F. (2022). Concentration Prediction of Polymer Insulation Aging Indicator-Alcohols in Oil Based on Genetic Algorithm-Optimized Support Vector Machines. *Polymers*, *14*(7), 1449.
[208] Zhong, M., Cao, Y., He, G., Feng, L., Tan, Z., Mo, W., & Fan, J. (2023). Dissolved gas in transformer oil forecasting for transformer fault evaluation based on HATT-RLSTM. *Electric Power Systems Research*, *221*, 109431.
[209] A. Kunakorn, S. Pramualsingha, P. Yutthagowith, P. Nimsanong, and S. Kittiratsatcha, "Accurate assessment of moisture content and degree of polymerization in power transformers via dielectric response sensing
[210] I. M. Malik, A. Sharma, and R. Naayagi, "A Comprehensive and Practical Method for Transformer Fault Analysis with Historical Data Trend using Fuzzy Logic," *IEEE Transactions on Dielectrics and Electrical Insulation,* 2023.
[211] M. Bagheri, S. Nezhivenko, M. S. Naderi, and A. Zollanvari, "A new vibration analysis approach for transformer fault prognosis over cloud environment," *International Journal of Electrical Power & Energy Systems,* vol. 100, pp. 104-116, 2018.

ANDREW ADEWUNMI ADEKUNLE received a bachelor's degree in physics from the University of Ilorin, Ilorin, Nigeria, in 2016, and a postgraduate diploma in electrical engineering from Ahmadu Bello University, Zaria, Nigeria in 2021, where he had quality experience in synthesizing alternative insulating liquid for power transformer applications in the material physics laboratory. He is currently pursuing a master's degree in electrical engineering with the University of Quebec at Chicoutimi (UQAC), Québec, Canada. He is doing his master's research with the Canada Research Chair tier 1, in Aging of Oil-Filled Equipment on High Voltage Lines (ViAHT), UQAC. His main research interests include high voltage engineering, liquid dielectrics, power transformer diagnostics, and machine learning.

ISSOUF FOFANA (M'05–SM'09) obtained his PhD from École Centrale de Lyon, France, in 1996. He was a postdoctoral researcher in Lyon in 1997 and was at the Schering Institute of High-Voltage Engineering Techniques at the Leibniz University of Hanover, Germany from 1998 to 2000. He was a Fellow of the Alexander von Humboldt Stiftung from November 1997 to August 1999. He has been at Université du Québec à Chicoutimi (UQAC), Quebec, Canada since 2000, where he is presently a full Professor. He has held the Canada Research Chair, tier 2, of insulating liquids and mixed dielectrics for electrotechnology (ISOLIME) from 2005 to 2015. He is holding the Canada Research Chair tier 1 on the Aging of liquid filled power equipment installed on High Voltage networks (ViAHT), director of the MODELE laboratory and the International Research Centre for Nordic Engineering (CIIN). Prof. Fofana is an accredited professional engineer in the province of Québec, Fellow of the IET and the 2021 IEEE Chen Jidan Award. He is currently a member of few international conference scientific committees (ICDL, CEIDP, ICHVE, CATCON, ICEMPE and ISH). Since 2017, he is the chair of the IEEE DEIS Technical Committee on Liquid Dielectrics.

PATRICK PICHER (M'91–SM'09) received the B.Eng. degree in electrical engineering from the Université de Sherbrooke, Sherbrooke, Québec, Canada, in 1993, and the Ph.D. degree from the École Polytechnique de Montréal, Montréal, Québec, in 1997. He has been working as a Researcher and a Project Manager with the Hydro-Québec's Research Institute, IREQ since 1999. Since 2003, he was involved in several international CIGRE working groups related to transformer modeling, diagnosis, and monitoring. Mr. Picher is a registered professional engineer, a Senior Member of IEEE, and a Distinguished Member of CIGRE.



ESPERANZA MARIELA RODRIGUEZ-CELIS received the B.Sc. degree in chemistry from the Pontifical Catholic University of Peru, Lima, Peru, in 2002, and the M.Sc. degree in pharmacy and the Ph.D. degree in chemistry from the University of Florida, Gainesville, FL, USA, in 2007 and 2009, respectively.She joined the Hydro-Québec's Research Institute (IREQ), Varennes, QC, Canada, in 2011. Her field of research is chemical markers of cellulose degradation in power transformers.

OSCAR HENRY ARROYO-FERNANDEZ obtained his B.Sc. degree in mechanical engineering in 2003 from the Pontifical Catholic University of Peru. In 2008, he obtained his master's degree in chemical engineering at Ecole Polytechnique in Montreal. He has authored/coauthored technical papers in the field transformers diagnostics or monitoring. He obtained his PhD in engineering at the Université du Québec à Chicoutimi (UQAC) 2017, Canada. He works at Hydro-Québec's Research Institute (IREQ), Varennes, Canada, since 2020. His field of research is distribution/power transformers and degradation of utility poles.