IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 16, 2019, accepted October 18, 2019, date of publication October 25, 2019, date of current version November 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2949580

A New Technique to Estimate the Degree of Polymerization of Insulation Paper Using Multiple Aging Parameters of Transformer Oil

SHIJUN LI[1], ZHAO GE[1], AHMED ABU-SIADA[2], (Senior Member, IEEE), LIUQING YANG[1], SHENGTAO LI[1], (Senior Member, IEEE), AND KIYOSHI WAKIMOTO[3]

[1]State Key Laboratory of Electrical Insulation and Power Equipment, Xi'an Jiaotong University, Xi'an 710049, China
[2]Electrical and Computer Engineering Department, Curtin University, Perth, WA 6102, Australia
[3]Transformer Design & Development Department, Development Section, Meidensha Corporation, Numazu-shi 410-8588, Japan

Corresponding author: Shengtao Li (sli@mail.xjtu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB0902705, and in part by the Project of State Grid Corporation of China under Grant 522722160001.

ABSTRACT Transformer insulation paper is a key indicator for transformer remaining operational life. Paper decomposition is evaluated using the degree of polymerization (DP) which calls for samples of insulation paper from operating transformers. Since collecting such paper samples is extremely difficult, other indicators have been used to indirectly reveal the DP value of insulation paper. This includes dissolved gases in transformer oil such as carbon oxides and hydrocarbon gases, furan compounds, methanol, ethanol, and moisture along with some oil characteristics such as interfacial tension. However, for the same oil sample, these individual parameters lead to different DP values. This is attributed to the lack of accuracy of the established mathematical and artificial intelligence models correlating DP with each of the above mentioned individual parameters. This paper presents a self-learning method to estimate the DP value of transformer insulation paper based on multiple transformer oil aging parameters. The proposed method comprises data processing, fuzzy c-means and linear regression. Results reveal that estimating the DP value based on multiple aging parameters is more accurate than estimating it using one single parameter as per the current practice. The proposed method not only helps to understand the correlation between multiple oil aging parameters and the DP value of paper insulation, but also promotes the establishment of more accurate life assessment models for power transformers based on these oil aging parameters.

INDEX TERMS Degree of polymerization, oil aging parameters, paper insulation, power transformers.

I. INTRODUCTION

Oil impregnated paper (OIP) is widely used as a solid insulating material in power equipment, such as power transformers [1]. During the operation of the power equipment, OIP gradually degrades due to the multi-stresses it is subjected to. Since insulation condition is a key factor for reliable operation of power equipment, evaluation of insulation condition has received much attention by researchers and industry [2].

The useful remaining operational life of a power transformer is identified based on the condition of its solid insulation that is measured using the degree of polymerization (DP) [3]. According to [4], initial DP value of fresh paper is within the range 1000-1200 which drops to 200-300 when the insulation paper reaches end-of-life. While DP is the most accurate indicator to assess the decomposition of paper insulation, it is rarely used by industry due to the complexity of collecting paper samples, especially from hotspot locations. As such, much research efforts have been conducted to correlate the DP value with other measurable chemical indicators and oil characteristics [5]–[13]. These indicators include furan compound (2-Fal) [5], [6], carbon oxides (CO and CO_2) [7], [8], hydrocarbon gases (CH_4, C_2H_6, C_2H_4, C_2H_2) [9], methanol (MeOH) [10], ethanol (EtOH) [11], moisture in oil [12] and oil interfacial tension [13]. Paper insulation comprises approximately 90% of cellulose, 6-7% hemi-cellulose and 3-4% of lignin [14]. Cellulose is a linear polymer of glucose molecules connected through

The associate editor coordinating the review of this manuscript and approving it for publication was Dusmanta Kumar Kumar Mohanta.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

glycosidic bonds [15]. Due to high thermal stresses, hydrogen bonds tend to breakdown resulting in a shorter cellulose molecular chain and chemical by-products are formed and dissolve in the oil. Furanic compounds and moisture are the main hydrolysis by-products of cellulose degradation [5], [12]. Carbon oxides and hydrocarbon gases are used in the Chinese standard DLT 722 to diagnose the transformer health condition [9]. According to [11], methanol is proved to measure the rupture of 1, 4-β-glycosidic bonds of cellulose, and ethanol is regarded as a high temperature thermal decomposition by-product of insulating paper. Oil interfacial tension identifies the insulation aging activity correlated with the acidity of the oil [13]. Each of these 8 oil aging parameters has a strong correlation with the DP of insulating paper.

Estimating the DP based on one single indicator results in inconsistent DP values for the same oil-paper sample. This is attributed to the lack of accuracy of the developed mathematical and artificial intelligence models correlating these individual indicators with the DP value. Moreover, such correlations vary with the type of oil and paper, oil-paper ratio and transformer operating conditions. Table 1 shows the correlation between the DP value and various individual oil aging parameters as suggested in the literature [5]–[13].

TABLE 1. Correlation of DP value and various oil aging parameters.

Oil aging parameter	Correlation	Application limitation	Reference
2-Fal	$DP \sim \log(2-\text{Fal})$	Mineral oil and Ester oil	[5][6]
	$\frac{1}{DP} \sim \log(2-\text{Fal})$	Mineral oil, middle and late aging stages	[11]
CO_2	$DP \sim CO_2$	Valid for mineral oil, invalid for ester oil	[8]
	$DP \sim \log(CO_2)$	Mineral oil	[7]
CO	$DP \sim \log(CO)$	Valid for mineral oil, invalid for ester oil	[7]
Total hydrocarbon gases	Increases when DP decreases	-	[9]
Moisture	Increases when DP decreases	Depends on oil type	[12]
Methanol	$\frac{1}{DP} \sim \text{MeOH}$	$DP > 300$, or early aging stage	[10]
Ethanol	Increases when DP decreases	Depends on oil type	[10][11]
Interfacial tension	Decreases when DP decreases	-	[13]

These various correlations do not necessarily lead to the same DP value for the same oil-paper sample. Even for the same aging parameter, various mathematical models can be found. For instance the correlation between 2-Fal and DP value has been modeled using various mathematical equations as summarized in Figure 1. From this figure, it can be seen that the deviation of the DP values calculated using

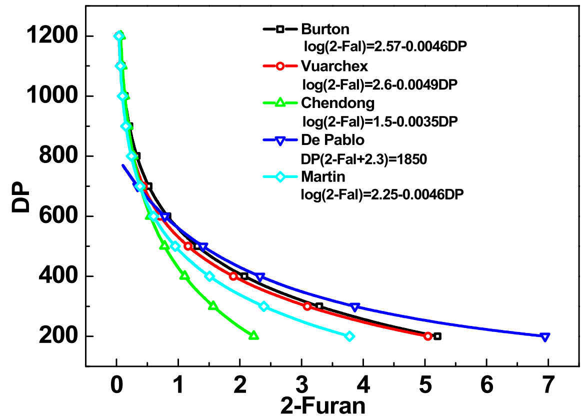

FIGURE 1. Comparison of different correlations of the DP and 2-Fal compound.

various mathematical models increases with the increase of 2-Fal concentration, especially when the DP value drops below 600. Such large deviation leads to inconsistent diagnosis for the paper condition when these mathematical models are employed for the same oil-paper sample. Therefore, an empirical equation with static coefficients using single aging parameter always fails to reveal accurate DP value of the insulation paper that is normally functioning under transformer harsh and dynamic operating conditions. Thus, the demand for a new self-learning model to estimate more accurate DP value based on multiple insulation aging parameters is essential.

The main contribution of this paper is the presentation of a new self-learning method to estimate the DP value of paper insulation based on multiple oil aging parameters that can be easily measured using transformer oil samples. In this context, series of thermal aging experiments on oil-paper samples are conducted and the proposed indicators are measured and randomly split into training and testing datasets. The proposed method comprises data processing, fuzzy-c means (FCM) and linear regression (LR). After data processing and determination of hyper parameters, the FCM-LR method is conducted on the training dataset to develop the model. This model is then examined using the reserved testing dataset. Furthermore, the proposed FCM-LR method is applied on other thermal aging experimental conditions and different types of oil and paper to further validate the robustness of the developed model when employed by different oil-paper materials.

II. THERMAL AGING EXPERIMENTS AND TESTING RESULTS

Series of thermal aging experiments were performed on oil-paper samples. In this regard, Weidmann T4 pressboard of 1 mm thickness and Karamay (KI 50X) transformer oil were chosen as the solid and liquid insulating materials, respectively. Several 2.5 mm $\times$ 4.7 mm paper samples were immersed in 500 mL of the insulating oil. All oil-paper samples were kept in glass bottles that were heated up in a thermally controlled oven for 240 days. During this period, the temperature within the oven was maintained at 90°C and aged oil-paper samples were extracted every 24 days to measure the above-mentioned aging parameters. This process was repeated under other oven temperatures; 110°C and 130°C. Thus, after the three thermal aging experiments, a total of 30 aged samples were obtained.

All tests on aged oil-paper samples were conducted in accordance to the standards listed in Table 2. The paper DP measurement was conducted following the IEC 60450 in which the specific viscosity of cellulose solution is measured to assess the DP value with the help of Mark Houwink equations [16]. DGA was conducted following the IEC 60567 standards using gas chromatography [17]. Interfacial tension of oil was measured following the ISO 6295-1983 guides [18]. 2-Fal in oil was determined following IEC 61198 standards using high performance liquid chromatography [19]. The moisture measurement was conducted following IEC 60814 guides using Karl Fisher titration [20]. Since specific international standard for measurement of methanol and ethanol in oil has not been set yet, test procedure is conducted by following the method in using headspace gas chromatography mass spectroscopy [21].

TABLE 2. Tests for aged samples with related standards.

Tests	Standard
paper DP	IEC 60450-2007[16]
DGA	IEC 60567-2011[17]
Interfacial tension	ISO 6295-1983[18]
2-Fal	IEC 61198-1993[19]
Moisture in oil	IEC 60814-1997[20]
Methanol and Ethanol of oil	Headspace-Gas Chromatography-Mass Spectroscopy [21]

In the proposed method, the target parameter is the DP value while the 8 oil aging parameters shown in the first column of Table 1 are the input parameters to the model. Therefore, the dataset used in this method consists of 30 individuals with 8 attributes and 1 label. The dataset is randomly split into training and testing datasets with a ratio of 4:1.

In order to examine the performance of the proposed method when applied to different oil-paper materials, other thermal aging experiments were performed on palm oil and another type of mineral oil (mineral oil 2), respectively. Since these two types of transformer oil are still in research stage, the commercial names are protected due to confidential policy. Various solid insulating materials as shown in Table 3 were used in this experiment. These solid and liquid materials were aged in a thermal aging chamber placed in thermal oven at 150°C for a duration of 200, 500, 1000 hours (for mineral oil samples) and 500, 1000, 2000 hours (for palm oil samples). The above mentioned parameters were measured in accordance to the standards in Table 2. This process was repeated at temperature of 165 °C for durations of 50, 100, 200 hours for mineral oil samples and 100, 200, 500 hours for palm oil samples. It was also repeated at a temperature of 180 °C for 20, 50, 100 hours for mineral oil samples and 50, 100, 200 hours for palm oil samples. Hence nine aged samples for each oil type were tested and therefore, 2 datasets consisting of 9 individuals with 8 attributes and 1 label were obtained.

TABLE 3. Quantity and size of solid and liquid materials placed in the thermal aging chamber.

Insulating materials	Quantity	Size (mm)
Insulating paper	20 pcs	100×100×0.13
Pressboard (A)	1 pc	100×100×1
Pressboard (B)	4 pcs	100×100×3
Paper insulating rectangular wire	11 pcs	160×10.9×2.4
Silicon steel sheet	1 pc	60×60×0.3
Transformer oil (mineral or palm oil)	2 L	-

The correlations of the paper DP value and the eight individual aging parameters for the above thermal aging experiments including the investigated three oil types are shown in Figure 2. As can be seen in these plots, each aging parameter exhibits inconsistent mathematical correlation with the paper DP value due to the different types of oil and paper as well as the thermal aging experimental conditions. Results show that some oil aging parameters such as interfacial tension and 2-Fal compound comprise similar correlation with

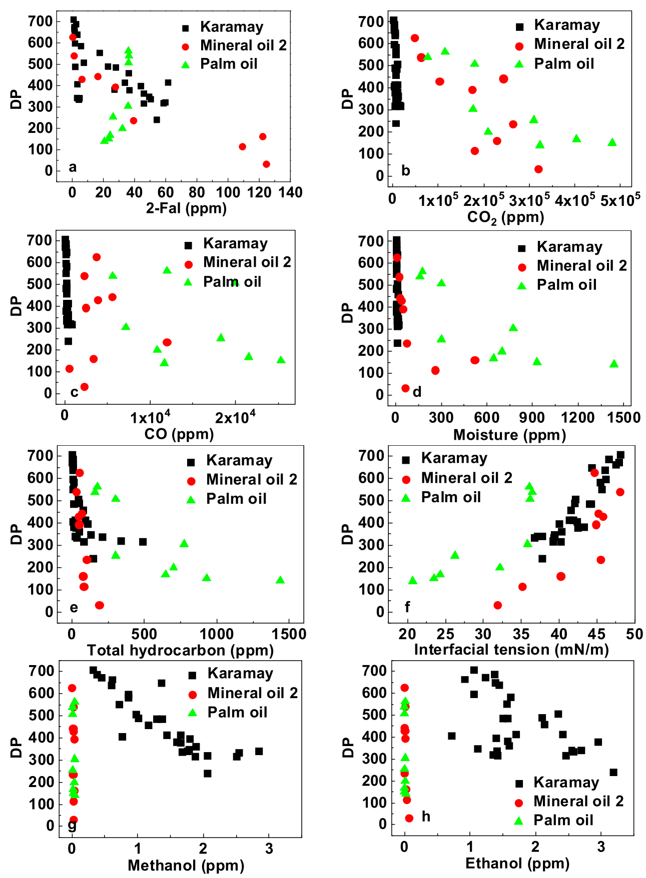

FIGURE 2. Correlation between DP and each single aging parameter.

the DP value for different thermal aging experimental results. However substantial difference on the range of DP value can be observed among different thermal aging experimental results. This validates the fact that a fixed equation correlating one oil aging parameter with the paper DP value is not very accurate which poses the demand for a more flexible and accurate method to reveal the DP correlation with multiple aging parameters.

III. ALGORITHM OF THE PROPOSED MODEL

The algorithm consists of several steps namely; data processing, fuzzy c-means (FCM), linear regression (LR) and evaluation. These steps are briefly elaborated below.

A. DATA PROCESSING BASED ON GREEDY ALGORITHM

As can be seen in the results of Figure 2, oil aging parameters exhibit various correlation patterns with the DP of paper insulation. For example, while this correlation is almost linear in case of interfacial tension, it is exponential in case of 2-Fal. Therefore, a kind of adaptive data processing method based on greedy algorithm as shown in Figure 3 is proposed in this paper to estimate this correlation. In this method, data processing is initially performed on one of the eight attributes to form a new combination $A_{i,s}$ with the original data of the other attributes. Comparison is then conducted between the new data combination ($A_{i,s}$) and the previous ones ($A_{i-1,s}$). Once the FCM-LR method on processed data reveal better performance, the new processed data replace the previous one and continue to the next attribute until all attributes are being considered. This optimization process is repeated 3 times in order to determine the ultimate proper combination of data processing method for each attribute.

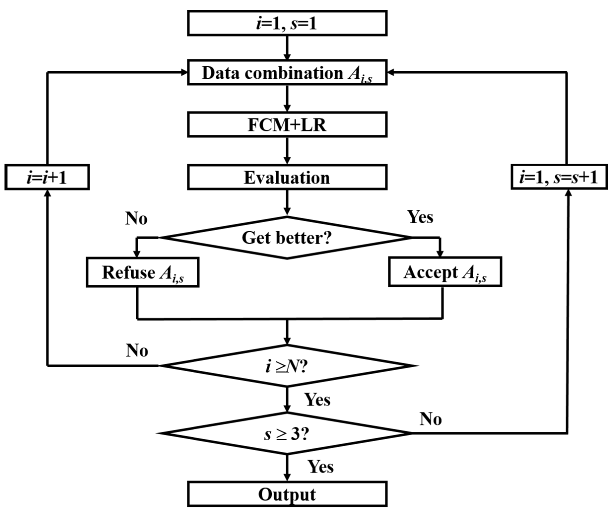

FIGURE 3. Data processing method based on greedy algorithm (*i* represents *i*th attribute, *s* represents the iteration step, *N* represents the total number of individuals).

After the data processing stage, normalization of input data as in (1) is conducted in order to limit the data into a standard scale. After normalization, the mean value of the normalized data vector x_i^* equals to 0, with a standard deviation equals to 1.

$$x_i^* = \frac{x_i - E(x_i)}{\sqrt{D(x_i)}} \tag{1}$$

where x_i^* is the normalized vector data of i^{th} attribute, $E(x_i)$ represents the mathematical expectation value of data vector x_i and $D(x_i)$ represents the variance of data x_i.

B. FUZZY C-MEANS (FCM)

Clustering is a typical unsupervised training method to solve unlabeled data analysis problem. FCM is a typical fuzzy clustering model, which assigns a degree of membership to every cluster center. FCM is built based on minimizing the below objective function J_m [22].

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{N_c} u_{ij}^m \left\| x_i - c_j \right\|^2, \quad 1 \le m < \infty \tag{2}$$

where N represents the total number of attributes, N_c represents the number of cluster centers, m is a real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, c_j is the j^{th} cluster center, and $||*||$ is the Euclidean norm that is revealing the similarity between any train input data and the cluster center.

In FCM method, N_c and m are two key hyper parameters which need to be determined before calculation. Fuzzy portioning is carried out using an iterative optimization of the objective function J_m, with the update of the degree of membership u_{ij} and the cluster centers c_j as below.

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \tag{3}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m} \tag{4}$$

This iteration process terminates when $\max_{ij}\{|u_{ij}^{(k+1)} - u_{ij}^{(k)}|\} < \delta$, where δ is a value between 0 and 1 and $k$is the iteration step.

This procedure converges to a local minimum or a saddle point of J_m. After iterative computation, the degree of membership values, u_{ij} are obtained to form the input matrix for the linear regression stage. For example, if FCM model of 10 cluster centers ($N_c = 10$) is trained and established on training dataset consisting of 24 individuals with 8 attributes, the calculation will generate a 10×8 cluster center matrix C as well as a 24×10 degree of membership matrix U.

C. LINEAR REGRESSION (LR)

LR serves as a linear approach to model the relationship between dependent and independent variables. In the proposed FCM-LR method, the dependent variable is the DP value of the paper insulation while the independent variables comprise the degree of memberships obtained from the FCM. The independent variable matrix Z given in (5) comprises degree of membership matrix U together with a column vector of ones.

$$Z = \begin{bmatrix} 1 & u_{11} & u_{12} & \cdots & u_{1j} \\ 1 & u_{21} & u_{22} & \cdots & u_{2j} \\ 1 & u_{31} & u_{32} & \cdots & u_{3j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_{i1} & u_{i2} & \cdots & u_{ij} \end{bmatrix} \tag{5}$$

LR algorithm can be presented in a matrix notation as in (6). The purpose of LR is to find a proper parameter vector β so as to minimize the error vector ε with the help of least squares estimation technique. For example, if FCM transfers a 24×10 degree of membership matrix U into LR, M will be a 24×11 independent variable matrix, and parameter β is an 11×1 vector.

$$DP = Z \cdot \beta + \varepsilon \tag{6}$$

D. EVALUATION

The coefficient of determination (R^2) is a typical evaluation indicator of regression performance, which is calculated as in (7). In the proposed FCM-LR method, R^2 is regarded as a single evaluation indicator of the FCM-LR model during the calculation.

$$R^2 = 1 - \frac{\sum (DP_{estimated} - DP)^2}{\sum \left(DP - \overline{DP}\right)^2} \tag{7}$$

where $DP_{estimated}$ represents the estimated DP value, which equals to $Z \cdot \beta$ obtained from (6), and $\overline{DP}$ represents the mean of the DP values.

Also, the coefficient of correlation (CC) and mean square error (MSE) given in (8) and (9) are applied to further evaluate the performance of the proposed method.

$$CC = \frac{Cov\,(DP, DP_{estimated})}{\sqrt{Var\,(DP)\;Var\,(DP_{estimated})}} \tag{8}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(DP_i - (DP_{estimated})_i)^2 \tag{9}$$

where $Cov(DP,\ DP_{estimated})$ represents the covariance between DP and $DP_{estimated}$, $Var(DP)$ represents the variance of DP, $Var(DP_{estimated})$ represents the variance of $DP_{estimated}$

IV. RESULTS AND DISCUSSION

A. INFLUENCE OF DATA PROCESSING

In order to show the influence of data processing, the value of R^2 and its bias are obtained as a function of the number of attributes using the row original and the processed data as shown in Figure 4. As can be seen in Figure 4, the performance of the processed data becomes more accurate than the original data. This accuracy is improved with the increase of the number of attributes. These results reveal that more attributes result in more accurate estimated value for the dependent parameter. Furthermore, after data processing, the performance of the model gets further progress compared with the model based on only original values, which indicates the necessity of data processing.

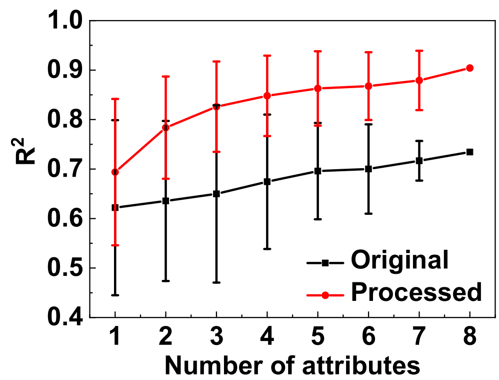

FIGURE 4. R^2 as a function of the number of attributes considered in the model based on original data and processed data ($N_c = 15$, $m = 3$).

B. DETERMINATION OF HYPER PARAMETERS

Hyper parameters in FCM model need to be determined before computation, including N_c and m parameters. The number of cluster centers N_c describes the number of groups which the input data belong to. This parameter influences the compactness of the fuzzy clustering model. Generally, more clustering centers improve the accuracy of the FCM, but increases the computation complexity from the other side, especially if it deals with big dataset. Besides, increase of the cluster number will probably face the risk of overfitting on training dataset. It is of importance to pick up an appropriate value to reduce the computation complexity and risk of overfitting in the case of high fitting performance. The other parameter m, which describes the fuzziness of the FCM, corresponds to the compactness and separation of FCM model. High value of m reduces the fuzziness of FCM, making it closer to crispy clustering method. In the proposed FCM model, these two hyper parameters are determined using grid search method. During the computation, the value of integrator N_c ranges from 2 to 15, while m ranges from 1 to 5 with a step of 0.1. The heat map of R^2 as a function of both N_c and m is shown in Figure 5(a) while Figure 5(b) shows the average of R^2 for the range of m at different N_c values and Figure 5(c) shows the average of R^2 for the range of N_c at different m values. This figure shows that despite of the same FCM method, the performance of each model depends on the values of the hyper parameters. The proper value for N_c falls within the range of 8 to 15 while m falls within the range of 1.1 to 2.5.

C. PERFORMANCE OF THE DEVELOPED MODEL ON TESTING DATASET

In order to examine the prediction accuracy of the developed FCM-LR model, proper data processing method of each attribute and position of cluster centers are obtained and then tested using the testing dataset after sufficient training.

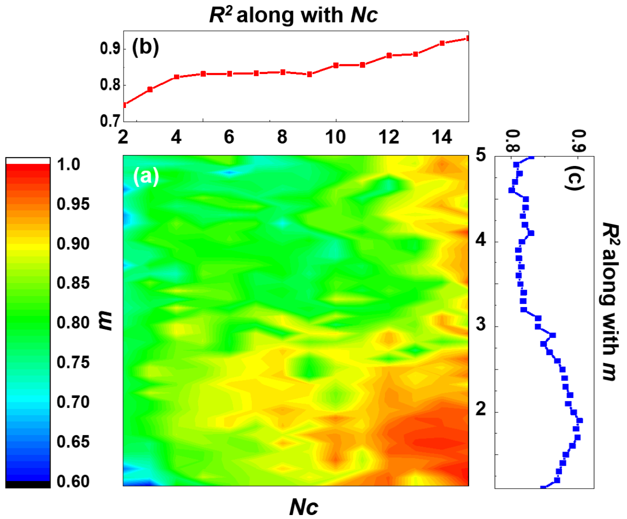

FIGURE 5. (a) Heat map of the R^2 of training dataset as a function of N_c and m, (b) R^2 as a function of N_c and (c) R^2 as a function of m.

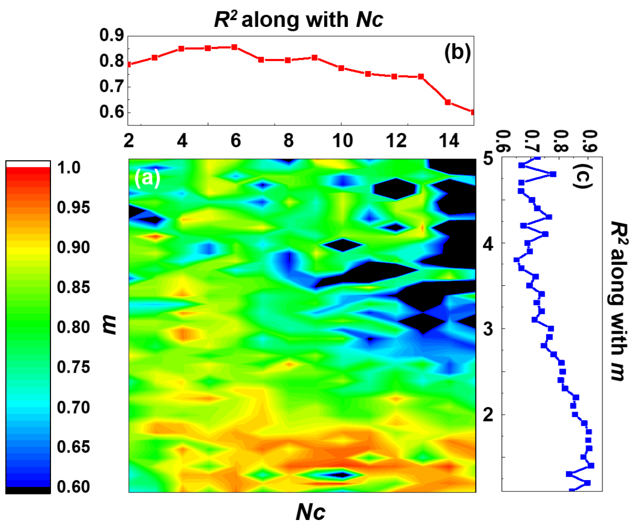

FIGURE 6. (a) Heat map of the R^2 of testing dataset as a function of N_c and m, (b) R^2 as a function of N_c and (c) R^2 as a function of m.

Similar to the training dataset performance shown in Figure 5, the heat map of R^2 of the FCM-LR model using the testing dataset as a function of N_C and m is shown in Figure 6. Generally, R^2 of FCM on testing dataset exhibits similar trend as the one resulted from the training dataset shown in Figure 5, especially for the plot of R^2 as a function of m (Figure 6 (c)). However, small difference appears on the plot of R^2 as a function of N_c (Figure 6(b)), which reveals that the proper value for N_c falls within the range 6 to 12 while it was in the range 8 to 15 from training dataset. This phenomenon reveals that the increase of cluster center numbers N_c may lead to the risk of overfitting on training dataset. Therefore, N_c should fall within the range 8 to 12 in order to obtain better computation results on both datasets.

Comparison of the performance of both training and testing datasets is shown in Figure 7, which indicates that the method

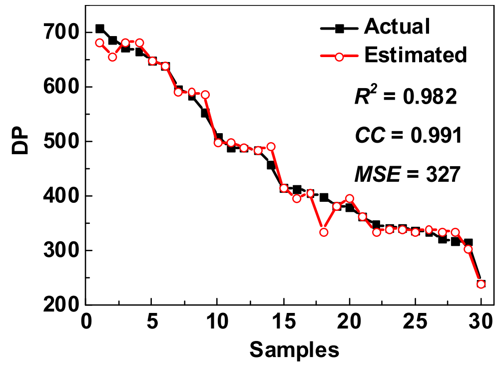

FIGURE 7. Performance of the proposed method ($N_c = 12$, $m = 1.2$).

works well on testing dataset. Table 4 compares the DP value calculated based on De Pablo's equation (DP vs 2-Fal) as in (10) [5] and Mandlik's equation (DP vs CO) as in (11) [7]. Comparison is made in terms of R^2, CC and MSE.

$$\mathrm{DP} = 1850 \cdot (2 - \mathrm{Fal} + 2.3) \tag{10}$$

$$\mathrm{DP} = -\frac{\ln\left(\frac{\mathrm{CO}}{555}\right)}{0.0016} \tag{11}$$

Results in Table 4 reveal that, the proposed method can estimate the DP value with higher accuracy than the current mathematical correlations as the R^2, CC and MSE are respectively found to be 0.982, 0.991 and 327 for the proposed method, while these parameters are respectively −3.70, 0.716, 8.33×10^4 based on De Pablo's equation and −5.07, 0.737, 1.07×10^5 based on Mandlik's equation.

TABLE 4. Evaluation indicators of the proposed method and other equations.

	FCM-LR ($N_c = 12$, $m = 1.2$)	De Pablo's equation	Mandlik's equation
R^2	0.982	-3.70	-5.07
CC	0.991	0.716	0.737
MSE	327	8.33×10^4	1.07×10^5

TABLE 5. Evaluation indicators of the proposed method and other equations for mineral oil-2 samples.

	FCM-LR ($N_c = 5$, $m = 1.8$)	PCA-LR	De Pablo's equation
R^2	0.993	0.969	-0.0414
CC	0.997	0.985	0.826
MSE	244	1.14×10^3	3.83×10^4

TABLE 6. Evaluation indicators of the proposed method and other equations for palm oil samples.

	FCM-LR ($N_c = 5$, m = 2.3)	PCA-LR	De Pablo's equation
R^2	0.995	0.944	-2.49
CC	0.998	0.972	-0.812
MSE	139	1.52×10^3	9.52×10^4

TABLE 7. Detailed information and DP values of training and testing datasets of the 30 aged OIP (KI50X) samples calculated by the proposed method and other equations.

	Sample	Aging Temperature (ºC)	Aging Duration (day)	Actual DP	FCM-LR (Nc = 12, m = 1.2)	Error (%)	De Pablo's equation	Error (%)	Mandlik's equation	Error (%)
Training set	1	90	20	709	682	3.808	578	18.477	1580	-122.849
	2	90	40	673	682	-1.337	498	26.003	1313	-95.097
	3	90	60	665	682	-2.556	495	25.564	972	-46.165
	4	90	100	648	648	0.000	479	26.080	645	0.463
	5	130	10	638	638	0.000	329	48.433	732	-14.734
	6	90	80	597	591	1.005	452	24.288	932	-56.114
	7	110	48	585	591	-1.026	232	60.342	634	-8.376
	8	130	20	507	498	1.775	189	62.722	679	-33.925
	9	130	30	489	498	-1.840	73	85.072	715	-46.217
	10	90	120	488	488	0.000	434	11.066	657	-34.631
	11	110	64	485	485	0.000	61	87.423	764	-57.526
	12	130	50	414	414	0.000	29	92.995	466	-12.560
	13	110	112	413	396	4.116	51	87.651	246	40.436
	14	90	140	406	406	0.000	330	18.719	635	-56.404
	15	110	80	381	381	0.000	63	83.465	1072	-181.365
	16	110	96	378	396	-4.762	47	87.566	650	-71.958
	17	110	128	363	363	0.000	38	89.532	275	24.242
	18	130	70	348	335	3.736	36	89.655	409	-17.529
	19	90	180	341	339	0.587	341	0.000	667	-95.601
	20	90	200	341	339	0.587	264	22.581	609	-78.592
	21	130	90	336	335	0.298	35	89.583	164	51.190
	22	90	240	334	339	-1.497	282	15.569	550	-64.671
	23	130	100	321	335	-4.361	30	90.654	-146	145.483
	24	110	192	240	240	0.000	33	86.250	223	7.083
Testing set	25	110	16	687	657	4.367	381	44.541	927	-34.934
	26	110	32	553	586	-5.967	93	83.183	888	-60.579
	27	130	40	458	490	-6.987	48	89.520	521	-13.755
	28	130	60	398	335	15.829	40	89.950	451	-13.317
	29	130	120	317	335	-5.678	30	90.536	-264	183.281
	30	110	144	316	304	3.797	38	87.975	407	-28.797

Detailed DP values and more information of training and testing sets are listed in Table 7 in the Appendix. The table shows that the proposed method in this paper is more accurate than the current methods used to estimate the paper DP value as evidenced from the least percentage error it results in for all training and testing data sets.

D. GENERALIZATION PERFORMANCE OF THE PROPOSED METHOD

The main feature of FCM-LR is the self-learning nature that facilitates its application for different thermal aging conditions and oil-paper types. The aging experimental results obtained using the other two oil samples (mineral oil-2 and palm oil) are used to assess this adaptive feature. The performance of the proposed FCM-LR model is compared with the performance of the principal component analysis linear regression (PCA-LR) proposed in [2]. Results of this comparison are listed in Tables 5 and 6 and are plotted in Figure 8. Evaluation indicators listed in Tables 5 and 6 attest the high accuracy of the proposed method in estimating the DP value based on multiple aging parameters when compared with the current methods.

Figure 8 shows the DP value predicted using PCA-LR and FCM-LR models for the 9 investigated samples of the two types of oil. As can been seen from Figure 8, the FCM-LR model has a better performance than the PCA-LR in estimating the DP values for various samples with a high degree of

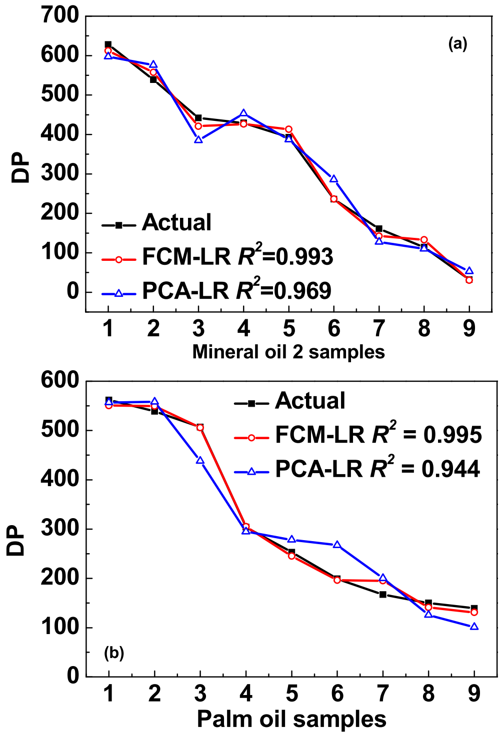

FIGURE 8. Actual and estimated DP values based on FCM-LR and PCA-LR method for (a) Mineral oil 2 and (b) Palm oil.

accuracy. Detailed DP values of the other aged OIP samples are listed in Tables 8 and 9 in the Appendix to compare the performance of the proposed method in this paper with the current methods in estimating the DP value based on one single parameter such as De Pablo's equation. Results in the tables reveal the high accuracy of the proposed method when compared with PCA-LR and De Pablo's equation. These results also reveal the suitability of the proposed method to be adopted by various oil-paper samples.

V. CONCLUSION

This paper proposes an FCM-LR method to predict the DP of transformer paper insulation using multiple aging parameters of the transformer oil. Results show that the proposed method is more accurate than calculating the DP value based on one single aging parameter such as 2-Fal as per the current practice. The proposed method features generalization performance that is examined on various oil-paper materials at different thermal aging conditions. The good performance of the proposed FCM-LR method makes it a powerful tool to estimate the DP value of transformer paper insulation with high degree of accuracy. The proposed technique can be automated and built within future power equipment online condition monitoring sensors to revel the condition of the solid insulation in real time.

APPENDIX

See Tables 7–9.

TABLE 8. Detailed DP values of aged OIP (mineral oil 2) calculated by proposed the method and other equations.

Actual DP	FCM-LR (N_c = 5, m = 1.8)	PCA-LR	De Pablo's equation
628	612	597	685
539	557	576	521
442	421	385	98
429	426	453	216
393	413	387	61
236	236	286	44
161	142	127	15
114	133	110	17
32	30	53	15

TABLE 9. Detailed DP values of aged OIP (palm oil) calculated by the proposed method and other equations.

Actual DP	FCM-LR (N_c = 5, m = 2.3)	PCA-LR	De Pablo's equation
562	550	557	48
539	550	558	48
507	506	438	48
304	305	295	49
253	245	278	65
199	196	267	54
167	195	200	70
150	141	126	72
139	131	101	80

REFERENCES

[1] Y. Zhu, S. Li, and D. Min, "Origin of dielectric processes in aged oil impregnated paper," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 3, pp. 1625–1635, Jun. 2017.
[2] S. Li, L. Yang, W. Yan, H. Cui, Z. Ge, and S. Li, "An effective principal component regression method for transformer life management based on indirect parameters," in *Proc. Condition Monit. Diagnosis*, Perth, WA, Australia, Sep. 2018, pp. 1–5.
[3] H.-Z. Ding and Z. D. Wang, "Author response to the comments by P. Calvini regarding the article 'On the degradation evolution equations of cellulose' by H.-Z. Ding and Z. D. Wang," *Cellulose*, vol. 15, pp. 229–237, Sep. 2007.
[4] 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, Sep. 2007.
[5] M. A. G. Martins, "Vegetable oils, an alternative to mineral oil for power transformers-experimental study of paper aging in vegetable oil versus mineral oil," *IEEE Elect. Insul. Mag.*, vol. 26, no. 6, pp. 7–13, Nov./Dec. 2010.
[6] L. Yang, Y. Lin, R. Liao, X. Zhao, W. Sun, and Y. Zhang, "Effects of temperature and aging on furfural partitioning in the oil-paper system of power transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 3, pp. 1393–1401, Jun. 2016.
[7] M. Mandlik and T. S. Ramu, "Moisture aided degradation of oil impregnated paper insulation in power transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 21, no. 1, pp. 186–193, Feb. 2014.
[8] L. J. Yang, "Study on aging characteristics of oilpaper in transformer and its lifetime estimation method," Ph.D. dissertation, Dept. Elect. Eng., Chongqing Univ., Chongqing, China, 2009.
[9] *Guide to the Analysis and the Diagnosis of Gases Dissolved in Transformer Oil*, Standard DLT 722, 2014.

[10] R. Gilbert, J. Jalbert, S. Duchesne, P. TÃľtreault, B. Morin, and Y. Denos, ''Kinetics of the production of chain-end groups and methanol from the depolymerization of cellulose during the ageing of paper/oil systems. Part 2: Thermally-upgraded insulating papers,'' *Cellulose*, vol. 17, pp. 253–269, Apr. 2009.

[11] S. Y. Matharage, Q. Liu, and Z. D. Wang, ''Aging assessment of kraft paper insulation through methanol in oil measurement,'' *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 3, pp. 1589–1596, Jun. 2016.

[12] L. E. Lundgaard, W. Hansen, D. Linhjell, and T. J. Painter, ''Aging of oil-impregnated paper in power transformers,'' *IEEE Trans. Power Del.*, vol. 19, no. 1, pp. 230–239, Jan. 2004.

[13] N. A. Baka, A. Abu-Siada, S. Islam, and M. F. El-Naggar, ''A new technique to measure interfacial tension of transformer oil using UV-Vis spectroscopy,'' *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 2, pp. 1275–1282, Apr. 2015.

[14] N. A. Bakar, A. Abu-Siada, and S. Islam, ''A review on chemical diagnosis techniques for transformer paper insulation degradation,'' in *Proc. Australas. Univ. Power Eng. Conf.*, Hobart, TAS, Australia, 2013, pp. 1–6.

[15] N. A. Bakar, A. Abu-Siada, and S. Islam, ''A review of dissolved gas analysis measurement and interpretation techniques,'' *IEEE Elect. Insul. Mag.*, vol. 30, no. 3, pp. 39–49, May/Jun. 2014.

[16] *Measurement of the Average Viscometric Degree of Polymerization of New and Aged Cellulosic Electrically Insulating Materials*, Standard IEC 60450:2004, 2004.

[17] *Oil-Filled Electrical Equipment—Sampling of Gases and Analysis of Free and Dissolved Gases—Guidance*, Standard IEC 60567:2011, 2011.

[18] *Petroleum Products—Mineral Oils—Determination of Interfacial Tension of Oil Against Water-Ring Method*, Standard ISO 6295:1983, 1983.

[19] *Mineral Insulating Oils—Methods for the Determination of 2-Furfural and Related Compounds*, Standard IEC 61198, 1993.

[20] *Insulating Liquids—Oil-Impregnated Paper and Pressboard—Determination of Water by Automatic Coulometric Karl Fischer Titration*, Standard IEC 60814, 1997.

[21] J. Jalbert, S. Duchesne, E. Rodriguez-Celis, P. Tétreault, and P. Collin, ''Robust and sensitive analysis of methanol and ethanol from cellulose degradation in mineral oils,'' *J. Chromatogr. A*, vol. 1256, pp. 240–245, Sep. 2012.

[22] J. C. Bezdek, W. Full, and R. Ehrlich, ''FCM: The fuzzy c-means clustering algorithm,'' *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, 1984.

SHIJUN LI was born in Shandong, China, in 1992. He received the B.Sc. degree from Xi'an Jiaotong University (XJTU), China, in 2015, the General Engineer degree from Ecole Centrale de Lyon, ECLyon, France, in 2016. He is currently pursuing the Ph.D. degree in domain of aging condition assessment of transformer with XJTU. He was with the Laboratory of TorreSupra, CEA Cadarache as a Researcher for three months in 2015 in the domain of superconducting in Tokamak.

ZHAO GE was born in Jiangsu, China, in 1994. He received the B.Sc. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2017, where he is currently pursuing the master's degree. His current research interests include thermal aging and life prediction of oil-paper insulation.

AHMED ABU-SIADA (M'07–SM'12) received the B.Sc. and M.Sc. degrees from Ain Shams University, Egypt, in 1998 and the Ph.D. degree from Curtin University, Australia, in 2004, all in electrical engineering. He is currently an Associate Professor and Lead of the Electrical and Electronic Engineering Discipline, Curtin University. His research interests include power system stability, condition monitoring, power electronics, and power quality. He has published over 230-research article, book and book chapters in these research areas. He is the Editor-in-Chief of the *International Journal Electrical and Electronic Engineering*, regular reviewer for various IEEE Transactions and a Vice-Chair of the IEEE Computation Intelligence Society, WA Chapter.

LIUQING YANG was born in Gansu, China, in 1986. She received the B.Sc. and Ph.D. degrees in chemistry from Lanzhou University, in 2008 and 2013, respectively. Since December 2017, she has been an Associate Professor with the School of Electrical Engineering, Xi'an Jiaotong University. Her research interests include energy storage, functional polyimide, and functional polymeric composites.

SHENGTAO LI (M'96–SM'11) received the Ph.D. degree in electrical engineering from Xi'an Jiaotong University (XJTU), China, in 1990. He was a Lecturer, an Associate Professor, and a Professor with XJTU, in 1990, 1993, and 1998, respectively. He was a Research Fellow with Waseda University, Japan, in 1996, and was also a Senior Visiting Scholar with the University of Southampton, U.K., in 2001. From 1993 to 2003, he was a Deputy Director of the State Key Laboratory of Electrical Insulating and Power Equipment (SKLEIPE), XJTU. Since 2003, he has been an Executive Deputy Director of SKLEIPE. His research interests include dielectrics and their application, insulating materials, and electrical insulation. He was awarded a Distinguished Young Scholar of China by the National Science Foundation, in 2006. He gave Liu Ziyu Memorial Lecture on the 11th International Conference on the Properties and Applications of Dielectric Materials (ICPADM), in 2015. He was the Chair of the 6th International Conference on Conditional Monitoring and Diagnosis (CMD), in 2016, and the IEEE Transactions on Dielectrics and Electrical Insulation Vol. 25, No. 4; August 2018 1575 1st International Conference on Electrical Materials and Power Equipment (ICEMPE), in 2017. He is serving on the Editorial Board of IEEE Access.

KIYOSHI WAKIMOTO was born in Hokkaido, Japan, in 1973. He received the B.Sc. and M.Sc. degrees from the Muroran Institute of Technology, Japan, in 1996 and 1998, respectively. He is currently with the Power Transformer Factory, Meidensha Corporation, and he is also the Manager of the Development Section. His research interests include insulation and cooling technology, analysis of electromagnetic transient phenomena, and fault diagnosis in power transformers. He is a Senior Member of IEE of Japan.