22nd International Conference on Renewable Energies and Power Quality (ICREPQ'24)
Bilbao (Spain), 26th to 28th June 2024
Renewable Energy and Power Quality Journal (RE&PQJ)
ISSN 2172-038 X, Volume No.22, September 2024

Use of a Decision Tree to Define Transformer Oil Contamination from On-Load Tap-Changer Gases to Ensure Power Quality in the Grid

S. Bustamante[1], J. L. M. Lastra[2], M. Manana[1], and A. Arroyo[1]

[1] Department of Electrical and Energy Engineering
Universidad de Cantabria
Avda. Los Castros, s/n, 39005 Santander, Spain.

[2] Faculty of Engineering and Natural Sciences
Tampere University
Korkeakoulunkatu 6, Tampere, 33720, Finland

Abstract. Power transformers are one of the most important and critical assets in the electricity distribution and transmission network. Power quality (PQ) can be disturbed when a power transformer is brought into or out of service, so it is very important to be sure of the reason for this action. Dissolved gas analysis (DGA) in oil can be used to diagnose the condition of transformer insulation. There may be situations where the DGA results indicate the presence of a serious fault which would lead to the transformer being taken out of service, when in fact the high gas concentrations are due to the leakage into the main oil tank of gases generated in the on-load tap-changer (OLTC) during normal operation. In previous work, using machine learning techniques and a distribution system operator's DGA database, a decision tree (DT) was developed to identify oil contamination from OLTC gases. In this work, the developed DT is applied to a new DGA database to identify contaminated transformers and test its accuracy. A total of 1161 DGA results from 95 transformers with OLTC were used, giving an initial DT accuracy of 83.13% when all samples were analysed and 85.26% when the last DGA result from each transformer was used.

Key words. Communicating OLTC, dissolved gas analysis, maintenance management, oil insulation, power transformer.

1. Introduction

One of the most important and critical assets in the electricity distribution and transmission network are the power transformers. As they are critical and very expensive machines, the maintenance work carried out on them is very important to ensure the correct operation of the network. Taking a power transformer in or out of service can lead to power quality (PQ) disturbances, so it is very important to know why this action is being taken. For example, taking a transformer out of service requires the load to be transferred to another part of the network or the back-up transformer to be energised, causing PQ disturbances in the network [1], [2], [3].
One of the most widely used tools for diagnosing the condition of transformers is dissolved gas analysis (DGA). It is possible to use DGA to diagnose the presence of insulation faults, which can lead to the need to take a power transformer out of service. The DGA measures the concentrations of gases in the transformer oil. These gases are hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4), ethane (C_2H_6), carbon monoxide (CO), and carbon dioxide (CO_2). According to IEC [4] and IEEE [5] guidelines, these gases are generated with a characteristic pattern that allows the type of fault to be identified. One of the existing fault identification methods, such as the Duval triangle (DTM) [6] and pentagon (DPM) [7] methods, can be used to diagnose the fault.

None of the methods for identifying transformer insulation faults takes into account the possibility of DGA results being contaminated by on-load tap-changer (OLTC) gases. In the in-tank OLTC design, gas leakage from the OLTC compartment to the main tank can occur [8]. This will cause the DGA result to be due to gases generated during OLTC operation and not due to an insulation fault.

In a previous work [9], a methodology for identifying oil contamination (OC) was defined using the DGA database and the knowledge of the distribution system operator's (DSO) maintenance engineers. This methodology was based on the application of the C_2H_2/H_2 ratio collected in the IEC [4] and IEEE [5] guides and the study of the DGA results of transformers with high C_2H_2 concentrations. By applying machine learning techniques to the results of the methodology, a decision tree (DT) was developed to automatically define the contamination [9].

In this work, the developed DT is applied to a new and different DGA database relating to power transformers with OLTC of a DSO. The results of the DT are compared with those defined by the maintenance engineers to verify the accuracy of the DT. Finally, the results obtained by applying the DT to the latest DGA results are studied in detail to identify new cases of contaminated transformers.

2. Dissolved gas analysis

The DGA is the most widely used tool for identifying and diagnosing insulation faults in power transformers. The gas concentrations generated in the transformer oil are related to the type of fault and the energy and temperature at the fault location [4], [5].

According to [10], C_2H_2 is formed at temperatures above 800 ºC, followed by rapid cooling. This occurs in high and low energy defects and thermal defects at temperatures above 700 ºC. However, the presence of high concentrations of C_2H_2 does not always indicate the presence of the above defects. During normal operation of OLTCs, high concentrations of C_2H_2 are generated during tap changes as a result of arcing between the fixed and moving parts and are dissolved in the OLTC oil compartment. In transformers with the in-tank OLTC design, the oil compartments may be in communication, causing the gases generated in the OLTC to leak into the main tank [8]. This results in altered DGA results which, when using any of the fault identification methods, will result in one of the faults discussed above, when in fact the gas concentrations are due to the correct operation of the OLTC. The typical C_2H_2 concentration ranges given in the IEC guide are different depending on whether OC is present or not, being 60–280 and 2–20 ppm respectively.

None of the methods for identifying transformer insulation faults take into account the possibility that DGA results may be contaminated by OLTC gases. The methodology presented in [9] is intended to assist in determining whether or not the transformer is contaminated by OLTC gases. This methodology is based on the application of the C_2H_2/H_2 ratio reported in [4], [5] and the study of DGA results of transformers with C_2H_2 concentrations above 10 ppm, referred to as expert knowledge (EK) in [9]. As a result of this methodology and the application of machine learning techniques, the DT in Fig. 1 was created with the aim of automatically detecting oil contamination. If the oil is not contaminated, the result of the DT is 0, and if it is contaminated, the result is 1 or 2, depending on whether the contamination is defined by the EK or by the application of the IEC ratio.

The accuracy ($\boldsymbol{A}$) of the DT was calculated using the following equation:

$$\boldsymbol{A}(\%) = \left(\frac{\boldsymbol{TP}+\boldsymbol{TN}}{\boldsymbol{TP}+\boldsymbol{TN}+\boldsymbol{FP}+\boldsymbol{FN}}\right).\mathbf{100} \qquad (1)$$

where TP and TN are the number of predicted true positives and true negatives, respectively, and FP and FN are the number of predicted false positives and false negatives, respectively.

3. Material and method

The initial data for this study were the 1274 results that make up the DGA database for 104 power transformers from a DSO.

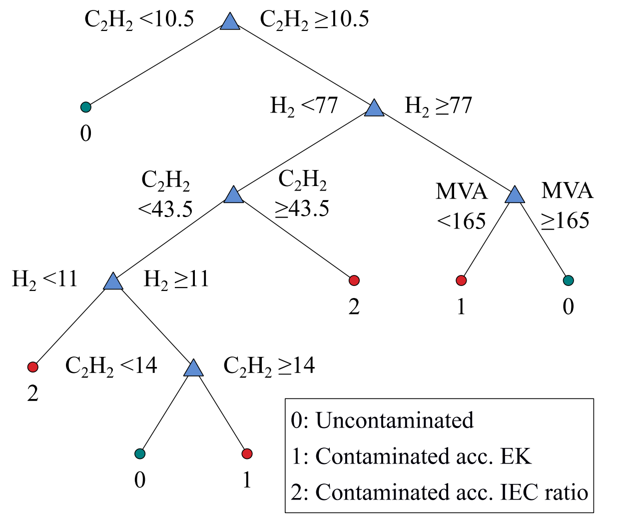

Fig. 1. Decision tree obtained in [9] (concentrations in ppm and power rating in MVA).

An initial screening of the database was carried out, removing DGA results from transformers without OLTC, as there can be no contamination from the OLTC gases. This first screening resulted in a working database of 1174 DGA results from 95 power transformers. Fig. 2 shows the distribution of the 95 transformers in terms of age, voltage class, and power rating. The number of transformers by age is broadly distributed, as shown in Fig. 2, with the highest numbers in the 10–20 and 50–60 age groups. Most transformers are of 50, 132 and 220 kV voltage class. In terms of power rating, most are between 10 and 100 MVA.

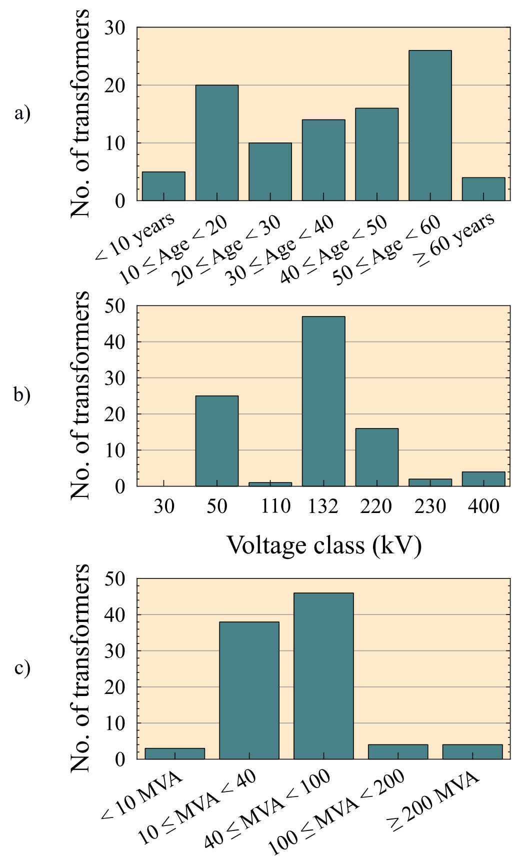

Fig. 2. Distribution of power transformers by (a) age, (b) voltage class, and (c) power rating.

In accordance with [5], a second screening was performed to check the reliability of the results by eliminating DGA results with missing data or abnormal gas concentrations. In this second screening, 13 DGA results were eliminated, leaving the working database with a total of 1161 DGA results from 95 transformers with OLTC. These DGA results were obtained between June 2004 and August 2023.

In the first part of this study, the DT shown in Fig. 1 was applied to the last DGA result for each transformer. The result was a prediction of whether or not oil contamination was present. A comparison was then made between the prediction made by the DT and the maintenance engineers' definition, focusing on the results of incorrect predictions. In the second part of this study, the DT was applied to the entire DGA database in order to observe the reason for the previous incorrect predictions. In addition, the results of transformer predictions where at least one did not match the maintenance engineers' definition were examined.

4. Results and discussion

Applying DT to the latest DGA results of the 95 power transformers resulted in an accuracy of 85.26% using equation 1, with fourteen cases of incorrect predictions, as shown in the confusion matrix in Fig. 3. The DGA results for the fourteen incorrect predictions are shown in Table I. The OC column indicates whether the contamination was present or not, as defined by the maintenance engineers. The DT column gives the prediction result, 0 for uncontaminated, 1 for contaminated according to EK, and 2 for contaminated according to IEC ratio.

In eleven of the fourteen incorrect cases, they were predicted to be uncontaminated when they were defined as contaminated. In ten of these cases, corresponding to transformers no. 1, 2, 3, 6, 7, 8, 10, 11, 12, and 13, according to the methodology presented in [9] that led to the DT, they would not be defined as contaminated by the OLTC gases because their C_2H_2 concentrations are very low. It is necessary to examine the remaining DGA results of these transformers to check whether these transformers should indeed be defined as contaminated. For example, the oil may have been treated prior to the last DGA result used and all gas concentrations would have decreased significantly. The last incorrect case of this type has a peculiarity, it corresponds to the transformer no. 5. Looking at the gas concentrations in Table I, it should have been predicted as contaminated. However, the methodology developed in [9] took into account a transformer connected to the transport network whose C_2H_2 concentration was high enough to define it as contaminated, but it was preferred to set it as uncontaminated in order to control its status with the lower limits defined in the IEC guide [4]. This situation caused the DT in Fig. 1 to create branches depending on the power rating of the transformer. By pruning these branches, transformer no. 5 would be defined as contaminated according to EK.

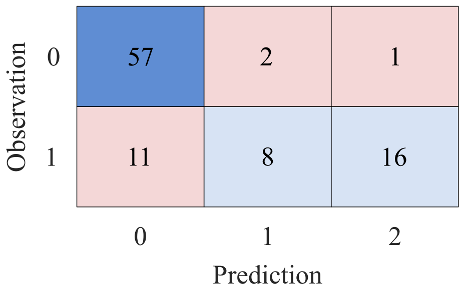

Fig. 3. Confusion matrix after applying the DT to the last DGA results.

The remaining three incorrect cases were predicted to be contaminated when they were defined as uncontaminated. These three transformers correspond to transformers 4, 9, and 14. Given the high concentration of C_2H_2, it is most likely that there would have been a gas leak from the OLTC compartment to the main tank. Therefore, in these cases, the trained model correctly identified the contamination, or at least these three transformers will be monitored for possible contamination.

Taking into account that four incorrect predictions could be considered correct, the accuracy of the DT when applied to the latest DGA results is between 85.26 and 89.47%.

Table I. - Prediction results based on the latest DGA results for transformers that do not match the maintenance engineer's definition.

Transformer No.	Installation Date	Power Rating (MVA)	DGA Date	H_2	C_2H_2	CH_4	C_2H_4	C_2H_6	CO	CO_2	OC	DT
1	1962	30	03/08/2023	5	1	3	31	3	188	2,656	Yes	0
2	1964	50	03/08/2023	17	4	25	228	19	515	4,410	Yes	0
3	1968	350	05/04/2022	9	1	14	155	26	353	3,279	Yes	0
4	1969	80	03/08/2023	58	120	13	23	1	126	2,109	No	2
5	1970	270	03/08/2023	208	409	38	79	23	167	3,321	Yes	0
6	1971	30	03/08/2023	0	3	3	19	8	28	2,099	Yes	0
7	1973	270	31/08/2021	31	1	155	45	321	728	2,645	Yes	0
8	1978	16	03/08/2023	16	13	5	27	20	60	2,112	Yes	0
9	1978	16	03/08/2023	12	21	2	161	3	152	2,050	No	1
10	1983	60	05/04/2022	25	6	7	120	31	293	1,691	Yes	0
11	1983	60	05/04/2022	24	13	3	18	5	305	2,219	Yes	0
12	1987	15	05/04/2022	71	1	1	1	1	85	1,156	Yes	0
13	1987	40	05/04/2022	27	11	4	27	10	302	2,910	Yes	0
14	2009	100	25/03/2022	576	100	87	68	68	52	295	No	1

Table II. - Results of the predictions on the transformers defined as uncontaminated by the maintenance engineers, where at least one of the predictions does not match.

Transformer No.	Installation Date	DGA Date																												
		2004	2005	2006	2007	2008	2009	2010	2011		2012	2013	2014	2015				2016	2017	2018			2019		2020		2021		2022	2023
4	1969																													2
9	1978		1																											1
14	2009																												1	
15	1962	1																												
16	1972			1		1																								
17	1990																						2							
18	2006							1																						
19	2006						1																							
20	2006						1																							
21	2007																						2							
22	2007																					1								

Green: prediction and definition match; Red: prediction and definition do not match.
1: contaminated according to EK; 2: contaminated according to IEC ratio.

Applying the DT to the entire DGA database gives an accuracy of 83.13%. The resulting confusion matrix is shown in Fig. 4.

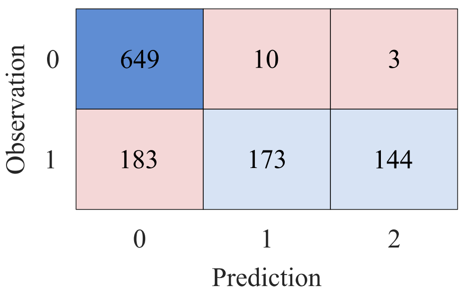

Fig. 4. Confusion matrix after applying the DT to all DGA results.

Table II shows the results of the predictions of all the DGA results of transformers defined as uncontaminated by the maintenance engineers, where at least one of the predictions does not match.

As mentioned above, the contamination predictions for transformers no. 4, 9, and 14 appeared in the last analysis carried out and they were therefore placed under surveillance in order to define the OC.

The predictions of transformers no. 15, 16, 18, 19, and 20 as contaminated were made using the initial DGAs carried out between 2004 and 2009, with the remaining predictions as uncontaminated using the following DGA results.

The C_2H_2 concentrations of 10.9, 11.3, and 18 ppm for transformers 17, 21, and 22 respectively are what caused the predictions to indicate OC. As the concentrations are very low and the following DGA results did not show an increase in C_2H_2 concentration, these predictions are incorrect.

Table III shows the results of the predictions of all the DGA results of transformers defined as contaminated by the maintenance engineers, where at least one of the predictions does not match.

Transformer no. 26 has the same situation as no. 5 above, its power rating is greater than 165 MVA and, according to the methodology developed [9], this type of power transformer is analysed with the lower limits of the IEC guide [4].

Transformers no. 24, 25, 27, 29, 30, 33, 34, 35, and 36 had very low C_2H_2 concentrations in the first DGA results. Given the C_2H_2 concentrations in these first analyses, it is most likely that the definition of OC in these transformers was after these DGA results, so the DT would have been correct.

The C_2H_2 concentrations are very variable in the last analyses for transformers no. 1, 2, 6, 8, 10, 11, 13, 23, 31, and 32, crossing the threshold of the OC definition several times up and down, hence the variability in the DT predictions.
As in the example above, in 2017 an oil treatment was performed on transformer no. 28, which caused the DT to fail in its prediction, the rest of the predictions are correct.

The two transformers whose predictions were all wrong are nos. 3 and 7, their C_2H_2 concentrations are between 0 and 4 ppm and 0 and 6 ppm respectively. After reviewing the OLTC test reports, which do not specify that there is communication between oil compartments, and the DGA results, it is considered that these transformers are not contaminated as indicated by the prediction results. Similarly for transformer no. 12, only the first two DGA results showed a high C_2H_2 concentration, the remaining analyses showed concentrations less than or equal to 4 ppm.

As shown in this section, the application of DT to the latest DGA results resulted in an initial prediction accuracy of 85.26% and led to three transformers being suspected of possible oil contamination. This means that when analysing future DGA results, the possibility of oil contamination will considered rather than the presence of an insulation fault which would lead to the removal of the transformers from service and the consequent PQ disturbances.

Table III. - Results of the predictions on the transformers defined as contaminated by the maintenance engineers, where at least one of the predictions does not match.

Transformer No.	Installation Date	DGA Date																													
		2004	2005		2006	2007	2008		2009	2010	2011	2012	2013	2014	2015	2016	2017			2018			2019		2020		2021		2022		2023
1	1962		G				G	G	R	R		G		G	R		R						R	R			R	R			R
2	1964	R			R	R	R			R	R	R		R	R	G	G			R	G		G				G	G			R
3	1968				R	R	R			R	R					R				R	R	R	R		R		R		R		
5	1970	G	G		G	R			G		G	G	R		R	R	G			R	G		R				G	R	R	G	R
6	1971	G									G	G	G		R		R			R											R
7	1973	R	R		R	R	R			R	R	R		R	R		R			R			R	R			R				
8	1978					R	R		G	G	G	R		R	R		G			R			R	R			G	G			R
10	1983		R		R		R	R	R	R	R	R		R		R	G			R			G		G		R		R		
11	1983	R			R		R			R		R		R		R				R	R		G		G		G		R		
12	1987	G	G		R		R			R		R		R		R	R			R					R	R	R		R		
13	1987		R	R					G		R	R				R	R			G					R	G	G		R		
23	1965		R				R		G		G	G	G		G		R			R	R		R				R	G			G
24	1969		R			R			R	R	R	R	R			R	R	R		R					G	G	G		G		
25	1971		R				G		G		G	G	G		G	G	G			G			G	G			G	G			G
26	1973	G				R			R		R	R	R		R	R	R			R	G		R				R	R			G
27	1973		R			G			R		G	R	G		G		G			G			G	G			G	G			G
28	1974		G			G	G			G		G		G	G	G	G	G	R	G			G	G			G	G			G
29	1974		G							R		R		G	G	G	G			G			G	G			G	G			G
30	1976	G			R		R			R		R		R	R		R			R	R		R				G	G			G
31	1978		G			G			R		G	R	R		R		R			R			R	R			R	R			G
32	1986		G			G			G		G	G	G			G				R	G	G			G	R	G		G		
33	1987				R	G			G		G	G	G		G		G			G			G	G			G	G			G
34	1988		G			R	R		R	R		R		G		G	G			G	G				G	G	G		G		
35	1988		G			R			R		R	R	G			G	G			G	G				G	G	G		G		
36	1990	G					R			G	G	G		G		G				G	G	G	G		G		G		G		

Green: prediction and definition match; Red: prediction and definition do not match.

An important issue to be considered is the prediction of transformers with power ratings above 165 MVA as non-contaminated, even with very high concentrations of C_2H_2. As this case occurred in the original methodology, but with concentrations slightly above the limits, it was defined as non-contaminated, but in the two cases discussed in this section, these two transformers should undoubtedly be defined as contaminated. It will therefore be necessary to re-train the model and see how new branches appear that take into account the power rating and the C_2H_2 concentration.

Another important issue is that two transformers are defined as contaminated by the maintenance engineers, but when making the predictions and reviewing the DGA results, they should not be defined as contaminated and should be treated with the lower limits of the IEC guideline.

The use of this new DGA database in the developed DT is a good source of results regarding the improvement needs of the trained model. Given the accuracies obtained, it is clear that the DT needs to be retrained. In order to do this, it will be necessary to apply the methodology described in [7] to the entire new database in order to determine the exact moment when the oil contamination from the OLTC gases is defined.

Finally, the latest versions of DTM [9] and DPM [10] for fault identification were applied to the latest DGA results of the 95 transformers, as shown in Figs. 5 and 6 respectively. The match or mismatch between the maintenance engineer's definition and the DT prediction is represented by a circle or star respectively. In addition, the results are presented in colours according to the engineers' definition, green for uncontaminated and red for contaminated.

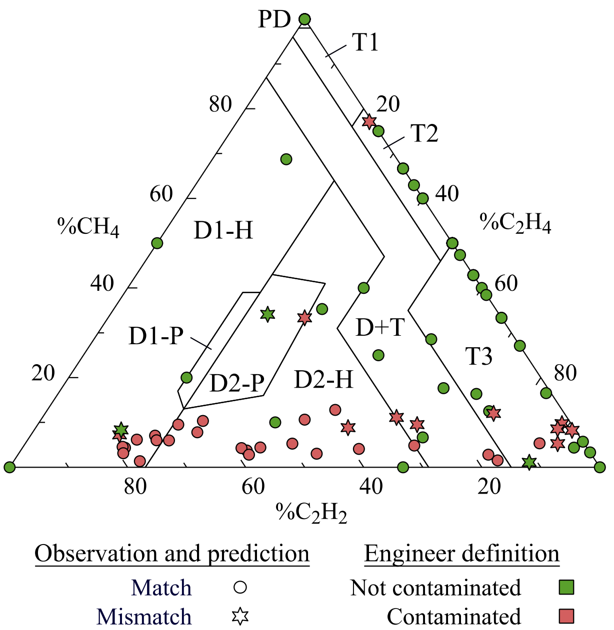

Fig. 5. Application of the DTM to the latest DGA results.

As can be seen in Figs. 5 and 6, most of the results of transformers defined as contaminated (predictions or

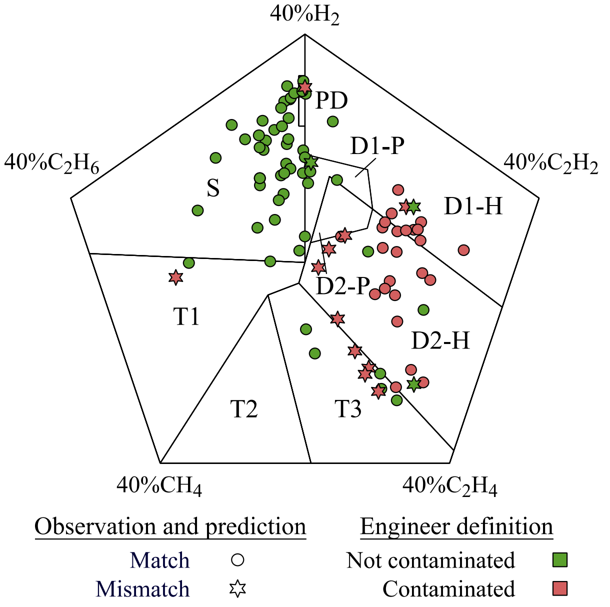

Fig. 6. Application of the DPM to the latest DGA results.

engineers' definition) are found above the zones indicating low (D1-H) and high (D2-H) energy discharges in the oil, mixed thermal and electrical faults (D+T) (only in the DTM), and thermal faults with a temperature higher than 700 ºC (T3). As discussed above, these types of faults generate C_2H_2 concentrations, so the identification methods diagnose the DGA results with these faults.

Therefore, it has been discussed throughout the study that the methods do not allow for the possibility of identifying oil contamination and that the DGA results are altered by the gas concentrations generated in the OLTC compartment.

5. Conclusion

This study presents the application of a DT based on the methodology developed in a previous work to define transformer oil contamination by OLTC gases. A new database of 1161 DGA results from 95 power transformers was used in the application of the DT. This new DGA database and the transformers to which they belong are different from the one used in the first study.

Applying the DT to all and the last DGA results gives an accuracy of 83.13 and 85.26% respectively. In the case of the last DGA results, this accuracy can be increased to 89.47%, since by analysing the incorrect predictions, four of them could be defined as correct. In addition, the possibility of OC was detected in three transformers, which were placed under surveillance and monitored for the evolution of their gas concentrations.

As a future study, the model will be retrained using the two DGA databases, the original and the one used in this work. This model retraining will take into account the results of this study, such as the influence of the DGA results of contaminated transformers where oil treatment has been carried out, or how to manage the definition of OC in transformers connected to the transmission network. After retraining the model, the decision tree run allows maintenance engineers to quickly decide which power transformers to focus on without having to review all the DGA results for all the transformers.

One of the limitations of the study was the unavailability of all OLTC models. Of the 95 transformers studied, the OLTC model was only available for 56 of them. For this reason, the results of the transformers defined as contaminated were not classified by OLTC type, as was done in the first paper. For future studies and retraining of the model, the full information from the OLTCs will be made available as far as possible from the inspections carried out on the OLTCs.

Distinguishing between the presence of a power transformer insulation fault and the detection of OC from OLTC gases allows a transformer to be taken out of service or not. This means that by avoiding the removal of a transformer from service, no PQ disturbances are generated by the energisation of the back-up transformer or by the transfer of the load to another part of the network.

Acknowledgement

This work was partially funded by the Contrato-Programa 2024 Government of Cantabria-Universidad de Cantabria. The authors acknowledge the support received from EDP Redes España.

References

[1] R. Kumar, B. Singh, R. Kumar, and S. Marwaha, “Recognition of underlying causes of power quality disturbances using stockwell transform,” IEEE Transactions on Instrumentation and Measurement (2020), vol. 69, no. 6, pp. 2798-2807.
[2] Y. Ma, X. Xiao, and Y. Wang, “Identifying the root cause of power system disturbances based on waveform templates,” Electric Power Systems Research (2020), vol. 180.
[3] R. Martinez, P. Castro, A. Arroyo, M. Manana, N. Galan, F.S. Moreno, S. Bustamante, and A. Laso, “Techniques to Locate the Origin of Power Quality Disturbances in a Power System: A Review,” Sustainability (2022), vol. 14, 7428.
[4] IEC 60599:2022 Mineral oil-filled electrical equipment in service. Guidance on the interpretation of dissolved and free gases analysis (2022).
[5] IEEE Std C57.104-2019 Guide for the interpretation of gases generated in mineral oil-immersed transformers (Revision of IEEE Std C57.104-2008) (2019).
[6] M. Duval and J. Buchacz, “Identification of Arcing Faults in Paper and Oil in Transformers—Part I: Using the Duval Pentagons,” IEEE Electrical Insulation Magazine (2022), vol. 38, no. 1, pp. 19-23.
[7] M. Duval and J. Buchacz, “Gas Formation from Arcing Faults in Transformers—Part II,” in IEEE Electrical Insulation Magazine (2022), vol. 38, no. 6, pp. 12-15.
[8] T. Breckenridge, S. Ryder, Overview of transformer and reactor procurement, CIGRE Green Books, Springer, Cham (2022).
[9] S. Bustamante, M. Manana, A. Arroyo, A. Laso, and R. Martinez, “Determination of Transformer Oil Contamination from the OLTC Gases in the Power Transformers of a Distribution System Operator,” Applied Sciences (2020), vol. 10, 8897.
[10] Advances in DGA Interpretation, JWG D1/A2.47, Technical Brochure No. 771, CIGRE (2019).