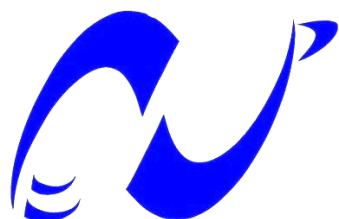
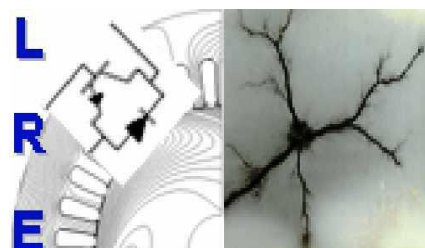


République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Ecole Nationale Polytechnique



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



Département d'Electrotechnique
Laboratoire de Recherche en Électrotechnique

Thèse de Doctorat

Présentée par :
Youcef BENMAHAMED

Master en Électrotechnique de L'ENP

Pour l'obtention du titre de
Docteur Troisième Cycle (D/LMD) en Électrotechnique

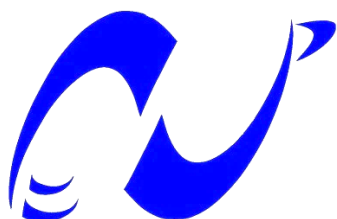
Thème

Application de méthodes intelligentes pour le diagnostic de l'huile de transformateur par analyse des gaz dissous

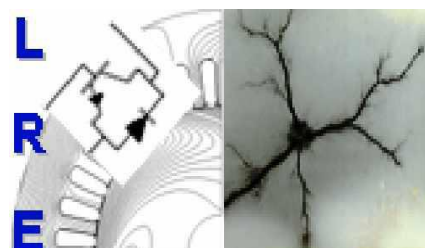
Président :	Ahmed Zaid Toudert	Professeur à l'ENP
Rapporteurs :	Teguar Madjid	Professeur à l'ENP
	Boubakeur Ahmed	Professeur à l'ENP
Examineurs :	Mekhaldi Abdelouahab	Professeur à l'ENP
	Moulai Hocine	Professeur à l'USTHB
	M'Ziou Nassima	Professeur à l'UMBB
Invité :	Ouagueni Mourad	Ingénieur d'Etudes à CREDEG/Sonelgaz

ENP 2019

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Ecole Nationale Polytechnique



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



Département d'Electrotechnique
Laboratoire de Recherche en Électrotechnique

Thèse de Doctorat

Présentée par :
Youcef BENMAHAMED

Master en Électrotechnique de l'ENP

Pour l'obtention du titre de
Docteur Troisième Cycle (D/LMD) en Électrotechnique

Thème

Application de méthodes intelligentes pour le diagnostic de l'huile de transformateur par analyse des gaz dissous

Président :	Ahmed Zaid Toudert	Professeur à l'ENP
Rapporteurs :	Teguar Madjid	Professeur à l'ENP
	Boubakeur Ahmed	Professeur à l'ENP
Examineurs :	Mekhaldi Abdelouahab	Professeur à l'ENP
	Moulai Hocine	Professeur à l'USTHB
	M'Ziou Nassima	Professeur à l'UMBB
Invité :	Ouagueni Mourad	Ingénieur d'Etudes à CREDEG/Sonelgaz

ENP 2019

ملخص

من أجل تحسين معدل تصنيف خماسي دوفال، تم استعمال خوارزميات آلة سند المتجهات (SVM) و ك أقرب الجيران (KNN) لتشخيص زيوت المحولات الكهربائية. تتكون قاعدة البيانات التي استخدمناها من تلك الموجودة في CEI TC 10 و CIGRE. وتتضمن 155 عينة تحتوي على تحليل الغازات الرئيسية الخمسة الناتجة عن انحلال الزيت (H_2 ، CH_4 ، C_2H_2 ، C_2H_4 و C_2H_6). تم تخصيص 105 عينة لمرحلة التدريب و 50 للاختبار. بالاعتماد على CEI و IEEE، تم تحديد خمس فئات من العيوب للتشخيص وهي PD، D1، D2، T1 و T2، T3. خوارزمية خماسي دوفال-PSO-SVM تم اختبارها باستعمال دالة نواة الخطية، كثير الحدود. تم النظر في عدة استراتيجيات تعلم: جميعا، واحدة ضد مجموعة واحدة ضد واحد. تم تحسين الإعدادات (C ، λ) من خلال تقنية تحسين سرب الجسيمات (PSO). توفر هذه الخوارزمية نسبة نجاح 88% باستخدام نواة حلقة كثير الحدود واستراتيجية تعلم واحد ضد واحد. أما بالنسبة لخوارزمية خماسي دوفال-KNN، تم اختباره من طرف ثلاثة أنواع من مسافة: الإقليدية، Cityblock و جيب التمام وعدد الجيران بين 1 و 105. وقد وجد أن نسبة النجاح أعلى من 82% لمسافة إقليدية وعدد الجيران مساو لـ 7. وأخيرا، تم اختبار أداء خوارزميات SVM-PSO و KNN والنظرية الافتراضية الساذجة من خلال تغيير نوع شعاع المدخل. لهذا، تم اعتماد خمسة أنواع وهي: تركيز الغازات الخمس الرئيسية ب جزء في المليون وتركيز الغازات بالنسبة المئوية، ونسب روجرز، ونسب مثلث دوفال ونسب دورنبرغ. تم الحصول على نسبة نجاح 98% عند استخدام خوارزمية PSO-SVM عن طريق حقن نسب مثلث دوفال كما ناقلا المدخلات ضد 92% لKNN.

كلمات البحث: محول الكهرباء، زيت العزل، التشخيص، تحليل الغازات المذابة، خماسي دوفال، آلة سند المتجهات، ك أقرب الجيران، النظرية

Abstract

In order to improve the classification rate of the Duval pentagon, SVM and KNN algorithms have been developed to diagnose transformer oil. The database we used is made up of those of IEC TC 10 and CIGRE. It contains 155 samples containing the analysis of the five main dissolved gases (H_2 , CH_4 , C_2H_2 , C_2H_4 and C_2H_6). 105 samples were reserved for learning and 50 for the test. Inspired by IEC and IEEE, five classes of defects were considered namely PD, D1, D2, T1 & T2, and T3. The pentagon algorithm of Duval-SVM-PSO was tested by kernel function: linear, polynomial, radial basic function and sigmoidal. Different learning strategies: together, one against all and one against one were considered. The parameters (C , λ) have been optimized by the particulate swarm optimization (PSO) technique. This algorithm offers an accuracy rate of 88% obtained using polynomial kernel function and one-against-one learning strategy. As for the pentagon algorithm of Duval-KNN, it was tested for three types of distance namely Euclidian, Cityblock and Cosine and for number of neighbors ranging between 1 and 105. The best accuracy rate of 82% was found for a Euclidean distance and number of neighbors equal to 7. Finally, the performances of the SVM-PSO, KNN and naive Bayes algorithms were tested by varying the type of the input vector. For this, five types were adopted namely, gas concentration in ppm, percentage gas concentration, Rogers ratios, Dornenburg ratios and Duval triangle reports. A success rate of 98% was obtained when using the PSO-SVM algorithm by injecting the Duval triangle reports as an input vector against 92% for the KNN one.

Key words: power transformer, insulation oil, diagnostics, dissolved gas analysis, Duval pentagon, Support Vector Machine, k-nearest neighbors, Naive Bayes.

Résumé

Dans le but d'améliorer le taux de classification du pentagone de Duval, des algorithmes SVM-PSO et KNN ont été élaborés pour diagnostiquer l'huile de transformateur. La base de données que nous avons utilisée est formée de celles de CEI TC 10 et CIGRE. Elle comporte 155 échantillons contenant l'analyse des cinq principaux gaz dissous (H_2 , CH_4 , C_2H_2 , C_2H_4 et C_2H_6). 105 échantillons ont été réservés pour l'apprentissage et 50 pour le test. Inspirés de IEC et IEEE, cinq classes de défauts ont été considérées à savoir PD, D1, D2, T1&T2, et T3. L'algorithme pentagone de Duval-SVM-PSO a été testé par une fonction noyau linéaire, polynomiale, à base radiale et sigmoïde. Différentes stratégies d'apprentissage à savoir « tous ensemble », « un contre ensemble » et « un contre un » ont été considérées. Les paramètres de marge et de conditionnement ont été optimisés par la technique d'optimisation par essaim particulière (PSO). L'algorithme pentagone de Duval-SVM-PSO offre un taux de réussite de 88% obtenu en utilisant la fonction noyau polynomiale et la stratégie d'apprentissage « un contre un ». Quant à l'algorithme pentagone de Duval-KNN, il a été testé pour trois types de distance à savoir « Euclidienne », « Cityblock » et « Cosine » et pour un nombre de voisins variant entre 1 et 105. Le meilleur taux de réussite de 82% a été trouvé pour une distance « Euclidienne » et un nombre de voisins égal à 7. Finalement, les performances des algorithmes SVM-PSO, KNN et naïve de Bayes ont été testées en variant le type du vecteur d'entrée. Pour cela, cinq types ont été adoptés à savoir, concentration de gaz en ppm, concentration gaz en pourcentage, rapports de Rogers, rapports de Dornenburg et rapports du triangle de Duval. Un taux de réussite de 98% a été obtenu lors de l'utilisation de l'algorithme PSO-SVM en injectant les rapports du triangle de Duval comme vecteur d'entrée contre 92% pour l'algorithme KNN.

Mots clés : transformateur de puissance, huile d'isolement, diagnostic, analyse des gaz dissous pentagone de Duval, Machines à Vecteurs Supports (SVM), k-plus proches voisins (KNN), naïve Bayes.

Remerciements

Avant tout, je remercie Allah, de m'avoir donné la force d'aller jusqu'au bout.

je tiens à remercier Messieurs **Madjid Tegar** et **Ahmed Boubakeur**, Professeurs à l'Ecole Nationale Polytechnique, pour avoir accepté de diriger et de suivre ce travail avec disponibilité, patience et bienveillance.

Je suis très honorée que Monsieur **Toudert Ahmed Zaid** Professeur à l'Ecole Nationale Polytechnique, ait accepté de présider le jury de soutenance de cette thèse.

Aussi, je tiens à remercier Monsieur **Abdelouahab Mekhaldi**, Professeur à l'Ecole Nationale Polytechnique, Monsieur **Hocine Moulai**, Professeur à l'Université des Sciences et de la Technologie Houari Boumediene, et Madame **Nassima M'Ziou** Professeur à l'Université de M'hamed Bougara de Boumerdès, d'avoir accepté d'examiner mes travaux.

Je tiens aussi à remercier également Monsieur **Mourad Ouagueni**, Ingénieur d'Etudes au CREDEG/Sonelgaz d'avoir accepté d'être membre invité dans le jury de soutenance.

Je tiens aussi à remercier tous les enseignants ayant contribué à ma formation et ce depuis le tronc commun jusqu'à la dernière année de graduation.

Dédicaces

Je dédie ce modeste travail,

À mes très chers parents, pour leur soutien et la patience dont ils ont fait preuve durant toutes les années de ma formation ;

À ma grande famille ;

Ainsi qu'à tous les électrotechniciens de ma promotion 2014 ;

À tous ceux qui m'aiment et qui me souhaitent le bonheur ;

Youcef

Table des matières

Résumé	
Liste des Figures	
Liste des Tableaux	
Introduction Générale	10

Chapitre I Huiles Isolantes pour Transformateurs

I.1 Introduction	14
I.2 Les transformateurs de puissances	14
I.2.1 Système d'isolation solide utilisé dans les transformateurs	16
I.2.2 Système d'isolation liquide utilisé dans les transformateurs	16
I.2.3 Les causes de vieillissements des huiles	17
I.3 Catégories d'huile et liquides isolants	17
I.3.1 Les hydrocarbures halogénés ou PCB	18
I.3.2 Les esters organiques (esters du pentaérythritol)	19
I.3.3 les huiles végétales (esters naturels)	19
I.3.4 Les huiles silicones	20
I.3.5 Les huiles minérales	20
1. Structure paraffinique (<i>Cp</i>)	21
2. Structure naphthénique (<i>Cn</i>)	21
3. Structure aromatique (<i>Ca</i>)	21
I.4 Les propriétés des huiles	22
I.4.1 Les propriétés électriques	22
1. La rigidité diélectrique et tension de claquage	22
a) Paramètre géométrique	23
b) Température	23
c) Teneur en eau	24
d) Nombre de Claquages	24
2. Permittivité relative (ϵ)	25
3. Facteur de perte diélectrique ($\text{tg } \delta$)	25
4. La conductivité / résistivité	25
5. Décharges partielles	26
6. Phénomène de gassing	26
I.4.2 Les propriétés physiques	26
1. La viscosité	26
2. Point d'écoulement	26
3. Le point d'éclair	26
4. Le point de feu	27
5. Capacité thermique	27
6. Conductivité thermique	27
7. Couleur	28
8. Masse volumique et coefficient de dilation	28
9. Aptitude de transfert de chaleur	29
I.4.3 Les propriétés physiques	29
1. La teneur en eau	29
2. L'acidité	29
3. Oxydation	30
I.5 Conclusion	30

Chapitre II Méthodes de Diagnostic Conventionnelles

II.1 Introduction	32
II.2 Méthode de Dornenburg	32
II.2.1 Méthodologie	32

II.2.2 Les inconvénients de la méthode de Dornenburg	35
II.3 Méthode de Rogers	35
II.3.1 Méthode de Rogers pour quatre ratios	35
II.3.2 Méthode de Rogers pour trois ratios	37
II.4 Méthode de la CEI 60599	39
II.5 Méthode de IEEE standard C57.104	40
II.6 Méthode des gaz-clés (Key Gas Method, David Pugh 1974)	41
II.7 Triangle de Duval	45
II.8 Pentagone de Duval	45
II.9 Cas d'application des différentes méthodes	49
II.10 Conclusion	51

Chapitre III Techniques de Classification et d'Optimisation

III.1 Introduction	53
III.2 Classifieur Machines à Vecteur Supports	53
III.2.1 Classifieurs SVM pour classification binaire	53
III.2.2 Classifieurs SVM multi-classe	58
III.2.3 Avantage et désavantages des SVM	60
III.3 Classifieur K plus Proches Voisins	61
III.3.1 Avantages et inconvénients	61
III.4. Classifieur Naïve Bayes	61
III.4.1 Construction de classifieur naïve Bayes	61
III.4.2 Avantages et inconvénients	62
III.5 Optimisation par les méthodes de métaheuristiques	63
III.5.1 Algorithme d'optimisation par les essaim de particules (PSO)	64
III.5.2 Principe de la technique de PSO	64
III.5.3 Configuration de l'algorithme PSO	64
1. Taille de la population	64
2. Initialisation de la position et la vitesse	65
3. Mise à jour de position et vitesse	65
III.5.4 Influence des paramètres sur l'algorithme de PSO	66
1. Facteur d'inertie	66
2. Coefficients d'accélération c_1 et c_2	67
3. Coefficient de constriction	67
4. Topologie de voisinage	68
5. Critère d'arrêt	68
III.6 Conclusion	69

Chapitre IV Application des Algorithmes SVM-PSO, KNN et Naïve Bayes dans le Diagnostic des Défauts

IV.1 Introduction	71
IV.2 Algorithme Pentagone de Duval-SVM-PSO	71
IV.3 Optimisation des paramètres de SVM par l'algorithme d'optimisation par essaim de particules (PSO)	72
IV.4 Algorithme pentagone de Duval-KNN	74
IV.5 Résultats de diagnostic	75
IV.6 Influence du type du vecteur d'entrée	78
IV.6.1 Algorithme KNN	82
IV.6.2 Algorithme naïve Bayes	85
IV.6.3 Algorithme SVM-PSO	88
IV.7 Conclusion	90
Conclusion Générale	93
Bibliographie	96
Annexes	103
Travaux effectués dans le cadre de cette thèse de Doctorat D/LMD	108

Liste des Figures

Figure I.1. Vue d'un transformateur de puissance [23]	15
Figure I.2. Structure de glucose et de cellulose [22]	16
Figure I.3. Structure de l'ester du pentaérythritol [29]	19
Figure I.4. Structure d'un ester naturel [29]	19
Figure I.5. Structure des huiles silicones [29]	20
Figure I.6. Structures des hydrocarbures de base dans d'huile minérale [35]	21
Figure I.7. Exemple d'électrodes cellulaires et sphériques appropriées (CEI 60156) [38]	23
Figure I.8. variation de la tension de claquage en fonction de la température pour différentes teneurs en eau [39]	24
Figure I.9. Tension de claquage d'une huile en fonction de la teneur en eau pour différentes températures [35]	24
Figure I.10. Conductivité thermique en fonction de température [45]	28
Figure II.1. Organigramme de la méthode de Dornenburg [6]	34
Figure II.2. Organigramme de la méthode de Rogers pour quatre ratios [55]	37
Figure II.3. Organigramme de la méthode Rogers utilisant trois ratios [56]	39
Figure II.4. Les gaz clés et leur proportions relatives [57]	42
Figure II.5. Triangle 1 de Duval	43
Figure II.6. Triangle 2 de Duval	44
Figure II.7. Triangle 4 de Duval	44
Figure II.8. Triangle 5 de Duval	45
Figure II.9. Pentagone I et II de Duval [9]	46
Figure II.10. Polygone irrégulier d'un échantillon	48
Figure II.11. Pentagone I de Duval	48
Figure II.12. Représentation du pentagone I de Duval utilisant la base de données CEI TC 10 ...	49
Figure II.13. Taux de réussite par type de défaut et par méthode	51
Figure III.1. Séparation linéaire de deux classes par SVM	54
Figure III.2. Séparation non linéaire de deux classes par SVM	55
Figure III.3. Changement d'espace	57
Figure III.4. Principe un-contre-un pour $w=3$	59
Figure III.5. Influence du nombre de voisins sur la classification	61
Figure III.6. Minima global et local	63
Figure III.7. Principe du déplacement possible entre deux points	65
Figure III.8. Voisinages social et géographique [90]	68
Figure IV.1. Distribution des coordonnées des centres de masse	72
Figure IV.2. Organigramme d'optimisation des paramètres de SVM	74
Figure IV.3. Organigramme récapitulatif des algorithmes élaborés	75
Figure IV.4. Performance du classifieur en fonction du nombre d'itérations	76
Figure IV.5. Impact de type de la distance et du nombre de voisins sur la performance du classifieur Pentagone de Duval-KNN	77
Figure IV.6. Représentation d'une boîte à moustaches	79
Figure IV.7. Boîte à moustaches (a) gaz en ppm, (b) gaz en pourcentage, (c) ratios de Triangle de Duval et (e) ratios de Rogers	81
Figure IV.8. Boîte à moustaches des ratios de Dornenburg (pour tous les classes)	81
Figure IV.9. Boîte à moustaches des ratios de Dornenburg par classe : (a) classe DP, (b) classe D1, (c) classe D2, (d) classe T1&T2 et (e) classe T3	81
Figure IV.10. Impact du nombre de voisins et le type de vecteur d'entrée sur la classification ...	82
Figure IV.11. Taux de réussite par type de défaut du classifieur KNN	84
Figure IV.12. Taux de réussite par type de défaut du classifieur naïve Bayes	88
Figure IV.13. Performance du SVM-PSO en fonction du nombre d'itérations pour différents types du vecteur d'entrée	89
Figure IV.14. Taux de réussite par type de défaut du classifieur SVM-PSO	89

Liste des Tableaux

Tableau I.1. Caractéristiques et applications des diverses catégories d'huiles diélectriques [27]	17
Tableau I.1. Normes d'essai de caractéristiques d'huile de transformateur [37]	22
Tableau II.1. Seuil des concentrations de gaz dans la méthode de Dornenburg [8]	33
Tableau II.2. Méthode de Dornenburg pour le diagnostic de défauts [8]	33
Tableau II.3. Gammes des ratios et leur code selon Rogers [7]	35
Tableau II.4 Diagnostic de défaut basé sur les codes de Rogers employant quatre ratios [7]	36
Tableau II.5. Gammes des ratios et leur code selon Rogers (méthode des trois Ratios) [7]	38
Tableau II.6. Code et défauts selon la méthode de Rogers (méthode des Trois ratios) [7]	38
Tableau II.7. Table d'interprétation d'AGD par la méthode de la CEI 60599 [5]	40
Tableau II.8. Conditions de concentration de gaz dissous dans l'huile selon IEEE [6]	41
Tableau II.9. Critères de diagnostic par la méthode des gaz clés [57]	41
Tableau II.10. Les gaz utilisés dans chaque triangle de Duval	43
Tableau II.11. Comparaison des types de défauts pour les deux versions de pentagone de Duval	46
Tableau II.12. Résultats de diagnostic pour 10 échantillons de test	49
Tableau II.13. Taux de réussite	50
Tableau IV.1. Résultats de diagnostic sur 10 échantillons testés	77
Tableau IV.2. Comparaison des taux de réussite des différents algorithmes	78
Tableau IV.3. Taux de réussite pour chaque vecteur d'entrée du classifieur KNN	84
Tableau IV.4. AGD de trois échantillons	86
Tableau IV.5. Sortie de classifieur naïve Bayes	87
Tableau IV.6. Taux de réussite pour chaque vecteur d'entrée du classifieur naïve de Bayes	87
Tableau IV.7. Paramètres ajustés donnant le meilleur taux de réussite pour chaque vecteur d'entrée	88
Tableau IV.8. Récapitulatif des résultats	90



Introduction Générale

Les transformateurs de puissance sont des éléments clés des réseaux électriques. En raison de leur coût onéreux, il est nécessaire de procéder à des inspections et maintenances périodiques afin d'en assurer un service efficace. La plupart des transformateurs de puissances sont remplis d'huile isolante qui assure une isolation électrique efficace. La circulation de l'huile à travers des radiateurs permet d'évacuer la chaleur produite par le circuit magnétique et les bobinages, assurant ainsi le refroidissement du transformateur.

La détection d'une initiation de dégradation de l'état d'une isolation, pouvant conduire à sa dégradation totale, nous permet de prévoir un incident, et donc d'éviter éventuellement une coupure généralisée de l'alimentation électrique et l'arrêt de fonctionnement d'usines et autres structures socio-économiques importantes. En effet, la défaillance d'une isolation utilisée dans le domaine de la haute tension, peut avoir des répercussions très graves sur les plans technique et surtout socio-économique.

L'utilisation de l'huile durant une longue durée sous plusieurs contraintes : électrique, thermique et chimique, engendre généralement sa dégradation. La durée de vie des transformateurs de puissance est principalement liée à la présence ou à l'absence de défauts. Il existe plusieurs techniques pour détecter les défauts liés à l'isolation diélectrique [1]. Parmi celles-ci, l'analyse des gaz dissous (AGD) est largement utilisée. Une telle technique de diagnostic est considérée comme l'une des méthodes prédictives les plus fiables utilisées pour réduire les risques de panne des transformateurs. L'avantage de l'AGD est de pouvoir être effectuée, soit directement à l'aide de capteurs introduits dans le transformateur en service (mode online), soit au laboratoire sur des échantillons prélevés (mode offline) [2].

En cas de défaut (surchauffe, carbonisation du papier, décharges de faible énergie, décharges de forte énergie, arcs), l'huile isolante de transformateur se décompose, provoquant la formation des gaz dissous. Les principaux gaz consistent en l'hydrogène (H_2), le méthane (CH_4), l'acétylène (C_2H_2), l'éthylène (C_2H_4), et l'éthane (C_2H_6) [3]. En outre, la décomposition du papier isolant produit le monoxyde de carbone (CO) et le dioxyde de carbone (CO_2) [4]. L'analyse de ces gaz permettrait de déterminer le type de défauts. Notons que la CEI [4] et IEEE [5] considèrent six défauts électriques et thermiques à savoir : décharges partielles (PD), décharges à faible énergie (D1), décharges à forte énergie (D2), défauts thermiques pour $T < 300$ °C (T1), défauts thermiques pour T de 300 °C à 700 °C (T2) et défauts thermiques pour $T > 700$ °C (T3).

Pour interpréter les résultats de l'AGD de l'huile de transformateur, différentes techniques ont été développées et regroupées en deux catégories de méthodes. Nous avons les méthodes conventionnelles et les méthodes basées sur l'Intelligence Artificielle.

Les méthodes conventionnelles se basent sur des rapports de concentration de gaz d'hydrocarbures précédemment cités [3]. A titre d'exemple, la méthode de Rogers [6] et celle de la CEI 60599 [4] utilisent trois rapports de concentration de gaz en ppm à savoir C_2H_2/C_2H_4 , CH_4/H_2 et C_2H_4/C_2H_6 . Celle du triangle de Duval emploie trois pourcentages de gaz consistant en % CH_4 , % C_2H_2 et % C_2H_4 [7]. Finalement, celle de Dornenburg [8] est basée sur les quatre rapports de concentration de gaz suivants CH_4/H_2 , C_2H_2/CH_4 , C_2H_4/C_2H_6 et C_2H_2/C_2H_4 .

Récemment, le pentagone de Duval a été développé comme nouvel outil complémentaire pour l'interprétation de l'AGD de l'huile de transformateur [9]. Il utilise les cinq principaux gaz d'hydrocarbures (en %) à savoir % H_2 , % CH_4 , % C_2H_2 , % C_2H_4 et % C_2H_6 . Outre les six défauts pris en compte par la CEI et IEEE, Duval a rajouté un défaut supplémentaire. Il s'agit du « stray gassing » (noté S) [10].

Ces méthodes conventionnelles sont faciles à utiliser. Cependant, leur précision de diagnostic nécessite une amélioration.

En utilisant les bases de données des gaz dissous de l'huile de transformateurs en exploitation, diverses techniques intelligentes ont été largement appliquées au diagnostic de défaut des transformateurs de puissance [11-15].

Dans [11], un classifieur probabiliste basé sur la fenêtre de Parzen (Parzen Windows, PW) a été proposé pour la classification des défauts de transformateurs. La technique d'Optimisation par Essaim Particulaire (Particle Swarm Optimization, PSO) a été utilisée pour optimiser les paramètres de PW.

Pour diagnostiquer des défauts multiples de transformateur et d'indiquer quantitativement la probabilité de sévérité de chaque défaut, une technique basée sur la logique floue a été développée dans [12].

Une approche utilisant le Réseau Bayésien est présentée dans [13] pour identifier les types de défauts et minimiser le risque de défaillances prématurées au niveau des transformateurs.

Shintemirov et al [14] ont décomposé la base de données disponibles en plusieurs classes selon la nature des défauts (thermique, électrique, etc.). Ils ont développé, par la suite, un algorithme génétique (Genetic Algorithm, GA) permettant d'extraire les caractéristiques de chaque classe et de les utiliser comme entrées dans un réseau de neurones artificiels (Artificial Neural Network, ANN), ainsi que dans la méthode des machines à vecteurs supports (Support Vector Machines, SVM) et dans celle du classificateur des plus proches voisins (K-Nearest Neighbors, KNN) afin d'identifier les défauts.

Dans [15], des réseaux d'ondelettes (Wavelet Networks, WNs) ont été utilisés pour le diagnostic de l'huile d'un transformateur de puissance au moyen de l'AGD. Le vecteur d'entrée contient six éléments ; les cinq principaux gaz en ppm ainsi que leur somme. Pour chaque réseau, la fonction gaussienne a été utilisée.

Différents travaux de recherche ont été effectués au Laboratoire de Recherche en Electro-technique à l'ENP, sur l'utilisation des méthodes intelligentes dans le diagnostic des propriétés de l'huile de transformateur telles que : les réseaux de neurones artificiels, la logique floue, les techniques des essaims [16-18].

Dans la continuité de ces travaux, notre thèse a pour objectif de contribuer dans l'amélioration du diagnostic des défauts pouvant survenir dans les transformateurs de puissance à partir de l'AGD [19-21]. Elle comporte quatre chapitres.

Le premier chapitre est consacré aux différentes huiles utilisées dans les transformateurs de puissances. Nous présentons les différentes propriétés de ces huiles et les normes internationales pour les mesurer tenant compte de plusieurs contraintes qui influent sur leurs propriétés : électriques, chimiques, thermiques, climatiques...

Dans le deuxième chapitre, nous présentons les méthodes de diagnostic conventionnelles : la méthode des ratios de Dornenburg, la méthode des ratios de Rogers et la méthode de la CEI 60599. Les méthodes graphiques (triangle et pentagone de Duval) sont aussi présentées. Nous terminons ce chapitre par l'application de quatre méthodes (Rogers, CEI 60599, le triangle de Duval et le pentagone de Duval) et comparons leur taux de réussite. La base de données que nous avons utilisée contient 50 échantillons.

Dans le troisième chapitre, nous décrivons les méthodes intelligentes de diagnostic suivantes : machines à vecteurs supports (SVM), k-plus proches voisins (KNN) et naïve Bayes, ainsi que la technique d'optimisation par essaim particulière (PSO).

Dans le quatrième et dernier chapitre nous combinons dans une première partie, le pentagone de Duval avec l'algorithme KNN et SVM-PSO. En ce qui concerne l'application de l'algorithme KNN, le type de distance et le nombre de voisins ont été variés de telle sorte à obtenir le meilleur taux de réussite possible. Les paramètres de marge et de conditionnement de l'algorithme SVM ont été optimisés par la technique d'optimisation de PSO.

Dans la deuxième partie de ce chapitre nous étudions l'influence du type de vecteur d'entrée sur le taux de réussite par l'application des trois classifieurs : KNN, SVM-PSO et naïve Bayes.

Nous terminons notre thèse par une conclusion générale représentant une synthèse globale de nos travaux de recherche.



Chapitre I
Huiles Isolantes pour
Transformateurs

I.1 Introduction

L'augmentation de la demande en énergie électrique nécessite des transformateurs de haute qualité aussi bien dans le transport que dans la distribution de l'électricité.

Les huiles isolantes sont largement utilisées en tant que moyen de refroidissement et d'isolement dans les transformateurs. Plusieurs types d'huiles ont été développées pour obtenir une excellente isolation. Elles doivent être surtout stables à l'oxydation et aux variations thermiques. Parmi les différentes huiles, l'huile minérale est la plus utilisée.

Nous présentons dans ce chapitre les principales propriétés des huiles isolantes tenant compte des différentes contraintes qui leur sont appliquées. Nous donnons en premier lieu un aperçu succinct sur les transformateurs de puissance.

I.2 Les transformateurs de puissances

Les transformateurs sont des machines électriques statique à induction électromagnétique, à deux ou plusieurs enroulements. Les deux parties principales qui forment les transformateurs sont :

- Le circuit magnétique qui sert à canaliser le flux magnétique. Il est feuilleté afin de réduire les pertes par hystérésis et les courants de Foucault.
- La partie électrique formée d'enroulements primaires et secondaires, pour produire un flux variable au primaire et de créer une tension induite au secondaire.

La CEI divise les transformateurs de puissance dans l'huile en trois catégories en fonction de la puissance apparente [22] :

- Transformateur de distribution : jusqu'à 2 500 kVA en triphasé ou 833 kVA en monophasé.
- Transformateur de moyenne puissance : de 2,5 MVA jusqu'à 100 MVA en triphasé et 33,3 MVA en monophasé.
- Transformateur de grande puissance : en triphasé au-delà de 100 MVA, en monophasé au-delà de 33,3 MVA.

La figure I.1 illustre l'ensemble des principales parties d'un transformateur de puissance.

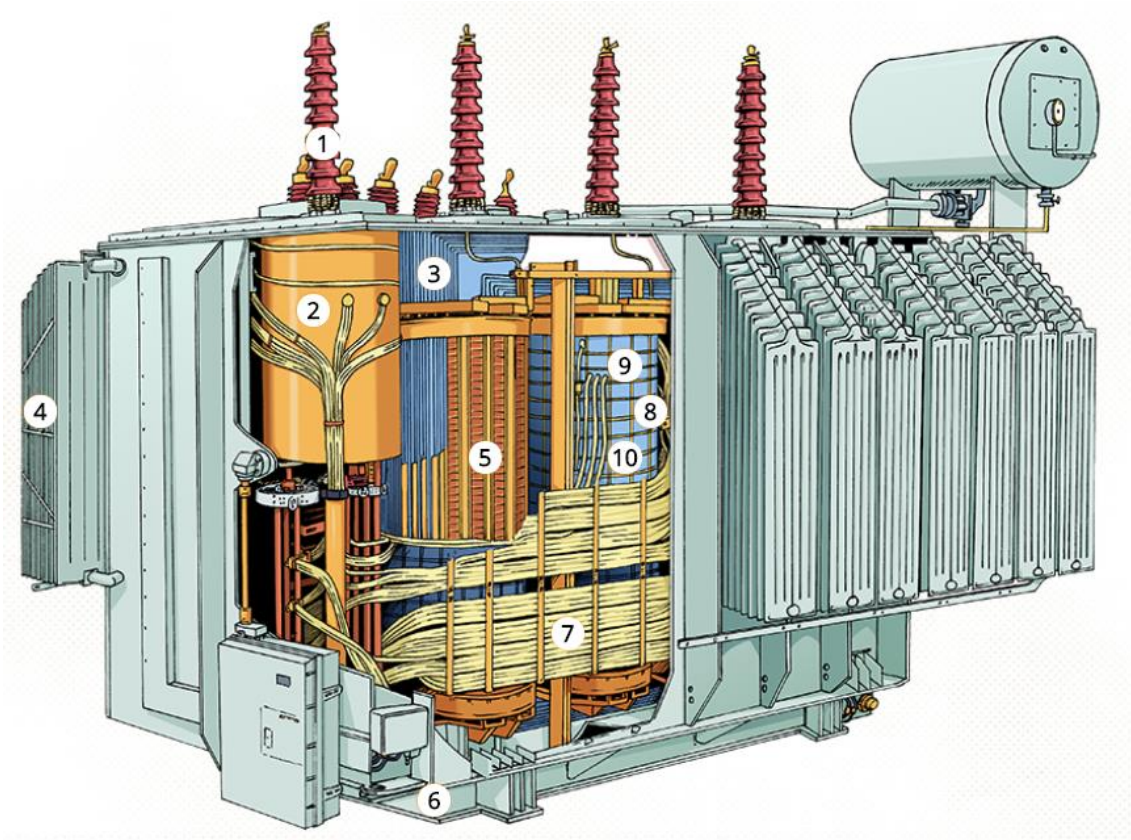


Figure I.1. Vue d'un transformateur de puissance [23].

- (1) Traversées et Borne en porcelaine
- (2) Changeur de prises
- (3) Circuit magnétique (CM)
- (4) Réfrigérants et protections
- (5) Isolation entre enroulements et par rapport à la masse
- (6) Masse / Cuve
- (7) Connexions des enroulements
- (8) Constitution des enroulements
- (9) Géométrie des enroulements
- (10) Isolation entre spires
- (11) réservoir d'expansion

I.2.1 Système d'isolation solide utilisé dans les transformateurs

Le papier et le carton sont employés dans les machines électriques pendant plus de cent années. L'isolant solide sert à l'isolation des enroulements, ainsi qu'entre les enroulements et le noyau. Le papier et le carton sont à base de cellulose (chaîne glucose extraites de bois ou de fibre de coton). Le bois est la source principale de cellulose. Le bois contient de 40 à 50% de cellulose, de 20 à 30% de lignine et de 10 à 30% d'hémicellulose et de polysaccharides à l'état sec [24].

Les structures du glucose et de la cellulose sont montrées sur la figure I.2 La cellulose est formée de longues chaînes comprenant environ 1200 anneaux de glucose reliés par des ponts d'oxygène [25]. La constante diélectrique du papier est deux fois plus élevée que celle de l'huile. Le papier et le carton seront donc utilisés dans les zones où règne un fort champ électrique pour renforcer l'isolation dans l'huile.

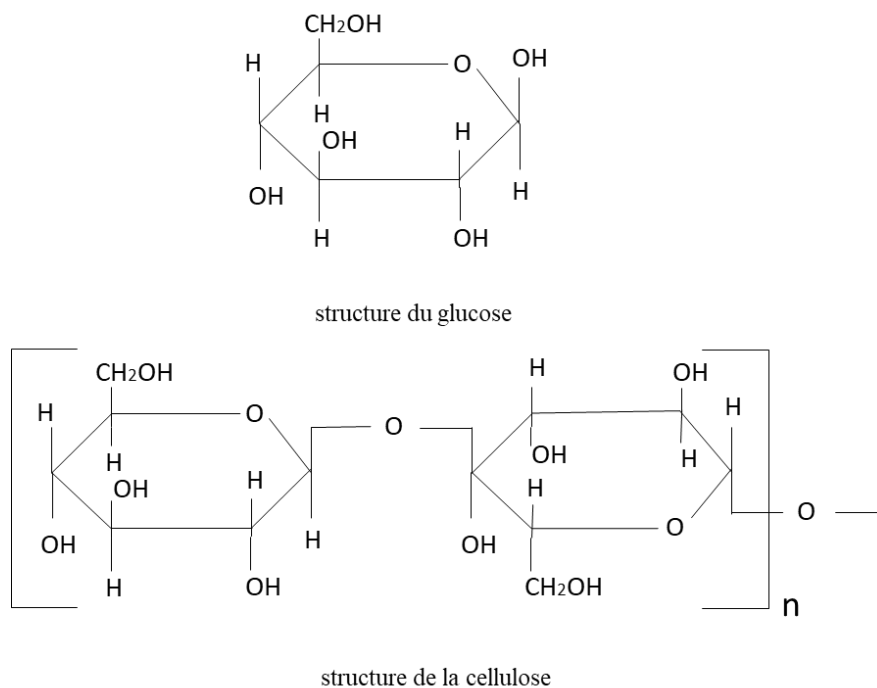


Figure I.2. Structure de glucose et de la cellulose [22]

I.2.2 Système d'isolation liquide utilisé dans les transformateurs

La technique de remplissage des transformateurs de puissance par les huiles isolantes a été développée depuis plus d'un siècle. Plusieurs types d'huiles ont été développés pour l'application dans les transformateurs de puissance. L'huile a trois principales fonctions [26] :

- Isoler électriquement ;

- Assurer l'évacuation de la chaleur produite par les pertes au niveau des enroulements vers les dispositifs de refroidissement (radiateurs externes) ;
- Ralentir la dégradation de l'isolation solide.

I.2.3 Les causes de vieillissement des huiles :

L'utilisation de l'huile durant une longue durée sous plusieurs contraintes : électrique, thermique, chimique, engendre généralement sa dégradation.

On appelle vieillissement, tout phénomène se traduisant par un changement lent et irréversible des propriétés du matériau. Les principaux types de vieillissement de l'huile de transformateur sont le vieillissement thermique, le vieillissement électrochimique et le vieillissement électrique. Le changement de propriétés affecte considérablement les performances de l'isolant et conduit à une réduction de sa durée de vie.

Le vieillissement thermique correspond à l'échauffement cyclique répété ou continu de longue durée, à des températures relativement élevées.

Le vieillissement électrochimique est caractérisé surtout par les conséquences, sous l'effet du champ électrique, de l'action de longue durée de certains agents chimiques accidentellement introduits dans l'huile ou provenant de la propre dégradation de cette dernière.

Le vieillissement électrique est quant à lui, surtout caractérisé par l'action de longue durée des décharges partielles ou tout phénomène d'ionisation dû au champ électrique. Il est caractérisé par l'augmentation des pertes diélectriques et une diminution de la résistance transversale.

I.3 Catégories d'huiles et de liquides isolants

Les types les plus courants d'huiles et de liquides isolants sont présentés dans le tableau I.1 par catégorie. On note que ces liquides sont généralement des liquides organiques, composés de molécules des séries d'hydrocarbures les plus stables : aliphatiques (alcanes), alicycliques (cyclanes), aromatiques ou de leur dérivé halogéné ; ce sont rarement des composés à groupements fonctionnels.

Tableau I.1. Caractéristiques et applications des diverses catégories d'huiles diélectriques [27]

Catégorie	Type	Application	Caractéristiques particulières
Huile minérales	Paraffiniques	Tous types : transformateurs, condensateur, câbles, etc.	- Bas point d'écoulement sans additif - Teneur en hydrocarbures aromatique > 1 2%
	Naphténiques		- Indice de viscosité de 11 5 à 120 (60 pour une huile naphténique)

	Aromatique	Transformateurs de distribution (remplacement des PCB)	- Point de feu > 300 °C
Hydrocarbure de synthèse	Polybutènes	Câbles d'énergie	- Pertes diélectriques très faible $\text{tg}\delta < 0,0005$ à 50 Hz et 90°C
	Alkylbenzènes	Traversées	- Rigidité électrique au choc de foudre élevée
	Alkylnaphtalènes	Condensateurs	
	Alkylbiphényles	Condensateurs tout film	- Absorption gazeuse sous décharges partielles
Esters organiques	Néocaprates de benzyle	Condensateur au papier	- Permittivité relative > 5 à 20°C à 50 Hz
	Diocylphtalate	Remplacement des PCB	- Permettent en mélange d'accroître le point de feu des diélectriques liquides
	Esters phosphates	Organe de coupure immergé	
	Esters du pentaérythrol	Transformateur de traction et distribution	- Point de feu > 300°C
Silicones liquides	Polydiméthylsiloxanes	Transformateur de traction	- Point de feu > 300°C - Indice de viscosité élevé : 170
	polyméthylphénylsiloxane	Liquide expérimental	- Absorption gazeuse sous décharge partielle
Liquides divers	Éthers-oxydes (alkyldiphényle-éther, ditolyléther)	Condensateurs	- Permittivité relative, à 20°C, supérieur à 3, adaptabilité
	Mélanges de liquides	Condensateurs	- Adaptabilité
	Gaz liquéfiés (Argon)	Enroulement supraconducteur et cryogénique	- Réduction ou suppression de pertes Joule.

I.3.1 Les hydrocarbures halogénés ou PCB

Les premiers hydrocarbures halogénés connus et désignés sous le terme générique d'askarels (en grec : résistant au feu) sont des polychrobiphényles (PCB). Ils n'ont pas de point de feu et sont ininflammable sous l'effet de la température, des décharges partielles ou lors d'un claquage électrique. Les PCB présentent une stabilité chimique, mais ils sont non biodégradables et donnent des produits toxiques à hautes températures. L'utilisation de ces huiles est limitée, puis a été interdite en Algérie par circulaire interministérielle datant du 04/12/1985 suivi du décret n° 87-182 du 18/08/1987 relatif aux PCB [28].

I.3.2 Les esters organiques (esters du pentaérythritol)

Les esters organiques ont été développés pour remplacer les PCB pour diverses applications (condensateur et transformateur), car ces huiles sont biodégradables et non toxiques. La figure I.3 illustre la structure d'un ester du pentaérythritol.

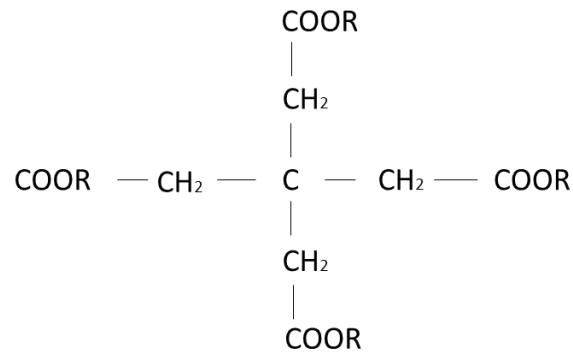


Figure I.3. Structure de l'ester du pentaérythritol [29]

La CEI 61099 donne les spécifications des huiles organiques à usage électrique [30]. Il existe aussi un guide de maintenance des esters pour transformateurs dans les matériels (CEI 61203 [31]). Le point fort de ces produits est la haute concentration d'eau qu'ils peuvent contenir. D'où, l'humidité devient un facteur moins limitant pour les propriétés électriques. Aussi, leur point de feu est supérieur à 300 °C. Cependant, leur point faible se situe au niveau de la viscosité qui est peu élevée et qui devient très importante aux basses températures. Leur coût relativement élevé est l'inconvénient majeur qui limite leur utilisation dans les transformateurs de puissance.

I.3.3 les huiles végétales (esters naturels)

Les huiles végétales sont des triglycérides et obtenues à partir de plantes ou de graines. Les huiles végétales ont été employées par homme depuis des siècles. La structure d'une huile végétale ou ester naturel est basée sur une chaîne d'alcool, à laquelle trois groupes naturels d'acide gras sont collés. Ces acides peuvent être les mêmes ou différents (figure I.4).

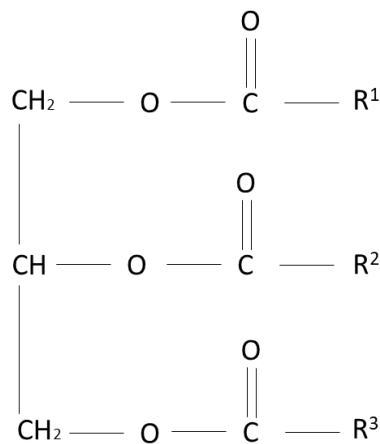


Figure I.4. Structure d'un ester naturel [29]

R^1 , R^2 et R^3 contiennent un nombre impair de carbones de 15 à 23.

Les huiles végétales sont considérées comme les premiers liquides isolants utilisés dans les dispositifs électriques. Plusieurs types d'huiles végétales existent, mais la plus utilisée est l'huile de ricin. Elle sert d'imprégnant dans l'isolation des condensateurs de stockage d'énergie [29].

L'avantage des huiles végétales est qu'elles sont biodégradables et peu toxiques et leur emploi dans les transformateurs ait été proposé récemment [31].

I.3.4 Les huiles silicones

Comme les esters organiques, les huiles silicones ont été développées aussi pour remplacer les PCB. Le produit le plus couramment employé est le polydiméthylsiloxane ou PDMS. Il est constitué d'une chaîne moléculaire linéaire, dont les éléments sont formés d'atomes de silicium et d'oxygène alternés. Les groupements de méthyle (CH_3) saturent les valences libres de l'atome de silicium. La formule générale est la suivante :

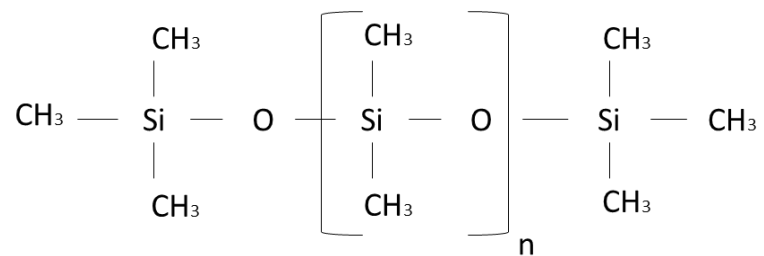


Figure I.5. Structure des huiles silicones [29]

La valeur de l'indice "n" caractérise leur degré de polycondensation, et il est d'environ 30.

La conservation des propriétés de l'isolant durant des périodes prolongées à haute température est le point fort de ce type d'huile. Cette conservation est obtenue grâce à une très grande stabilité à l'oxydation. Mais cette stabilité les rend difficilement biodégradables.

I.3.5 Les huiles minérales

Généralement, l'huile minérale est le liquide d'isolation le plus utilisé dans les transformateurs de puissance. Le choix d'une d'huile minérale est guidé par la norme CEI 60296 [33]. L'évaluation des huiles minérales en service est définie par la norme CEI 60422 [34]. Cependant, l'utilisation de d'huile minérale est aussi soumise à des conditions supplémentaires. Les huiles minérales sont obtenues par le raffinage des bruts pétroliers selon les trois types de structures à savoir paraffiniques, naphthéniques et aromatiques [29]. La figure I.6 illustre les trois structures de base de l'huile minérale.

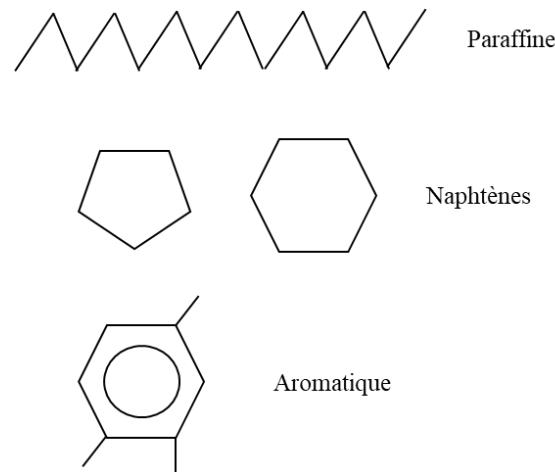


Figure I.6. Structures des hydrocarbures de base dans d'huile minérale [35]

1. Structure paraffinique (C_p)

La formule générale de ce groupe est C_nH_{2n+2} . Dans ce type d'huile la tendance paraffine ou la présence d'hydrocarbures paraffiniques prédomine. Ces hydrocarbures sont des molécules complexes à structure saturée et à chaîne droite ramifiée ou non, mais non cycliques (figure I.6). Les caractéristiques de l'huile de cette structure sont [36] :

- Faible densité pour une viscosité donnée ;
- Variation relativement faible de la viscosité en fonction de la température ;
- Faible volatilité pour une viscosité donnée ;
- Faible pouvoir solvant.

2. Structure naphténiqne (C_n)

La formule générale de ce groupe est C_nH_{2n} , et les molécules sont également connues sous le nom d'hydrocarbures saturés cycliques ou cycloalcanes en chaînes carbonées fermées (figure I.6). Les cycloalcanes sont moins stables à l'oxydation, mais présentent de meilleures propriétés d'écoulement aux basses températures et un meilleur pouvoir solvant que les n-alcanes [36].

3. Structure aromatique (C_a)

La formule générale de ce groupe est C_nH_{2n-6} . Ces huiles composées de fonctions à noyau benzéniques sont aussi appelées hydrocarbures insaturés. Ce type d'huile a une très faible résistance à l'oxydation, un indice de viscosité très faible et un pouvoir de solvant très élevé.

L'utilisation de ces hydrocarbures dans la fabrication des huiles est très limitée en raison de leur altérabilité. Ils sont facilement oxydables et provoquent la formation de dépôts résineux ou asphaltiques accompagnés de dérivés corrosifs [36].

I.4 Les propriétés des huiles

Le tableau I.2 résume les normes industrielles et les publications connexes concernant les procédures d'essais normalisés des caractéristiques de l'huile de transformateur. Ces tests sont regroupés en tests de routine, tests complémentaires et tests d'investigation spéciaux.

Tableau I.2. Normes d'essai de caractéristiques d'huile de transformateur [37]

Groupe de test	Nom du test	Norme de test
Tests de routine	Détermination de la couleur	ISO 2049
	Tension de claquage	CEI 60156
	Teneur en eau	CEI 60814
	Acidité (valeur de neutralisation)	CEI 62021
	Facteur de perte diélectrique	CEI 60247
	Contenu d'inhibiteur	CEI 60666
	Oxydation	CEI 61125 méthode C
Tests complémentaires	Tension interfaciale	ISO 6295
	Particules (nombre de particules)	CEI 6097
Tests d'investigation spéciaux	Stabilité à l'oxydation	CEI 61125
	Point d'éclair	ISO 2719
	Compatibilité	CEI 61125
	Point d'écoulement	ISO 3016
	Densité	ISO 3675
	Viscosité	ISO3104
	Biphényles polychlorés (PCB)	CEI 61619
	Soufre corrosif	DIN 51353

Le choix d'une huile isolante est déterminé à partir d'un nombre important de propriétés qu'on regroupe en trois catégories :

- propriétés électriques ;
- propriétés physiques ;
- propriétés chimiques.

I.4.1 Les propriétés électriques :

1. La rigidité diélectrique et tension de claquage

La rigidité diélectrique correspond au champ électrique maximal qu'on peut appliquer sans avoir de décharge disruptive. La tension de claquage d'un volume de liquide isolant se mesure entre deux électrodes dont la nature, la forme et la distance qui les séparent sont spécifiées [38] (Figure I.7). La norme CEI 60156, pour la mesure sous tension alternative, stipule que le liquide doit être soumis à un champ électrique quasi-uniforme avec une vitesse d'augmentation de la tension et une distance entre les électrodes fixées [35].

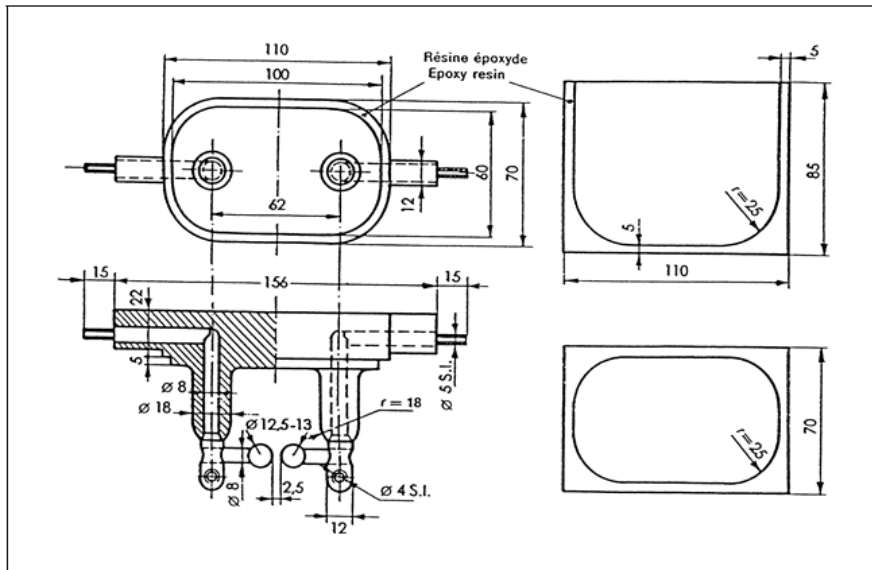


Figure I.7. Exemple d'électrodes cellulaires et sphériques appropriées (CEI 60156) [38]

$$E_C = \frac{U_C}{e}$$

où

E_C est la rigidité diélectrique en kV/mm.

U_C est la tension de claquage en kV.

e est la distance entre les deux électrodes en mm.

paramètres influant sur la rigidité diélectrique de l'huile :

La rigidité diélectrique dépend de plusieurs paramètres. Ces paramètres peuvent être géométriques, thermiques, climatiques, chimiques ... etc.

a) Paramètre géométrique :

La géométrie et la distance inter-électrodes sont très importantes, à cause de la rigidité qui varie avec leur forme et la distance qui les sépare. Plusieurs types d'électrodes ont été réalisés pour l'étude du claquage sous tension alternative à 50 Hz. La distance entre les électrodes varie de quelques millimètres à quelques centimètres [37-38].

b) Température :

La tension de claquage s'améliore avec la température [38], à cause de la croissance de solubilité de l'eau avec l'augmentation de température. La figure I.8 montre la variation de la tension de claquage en fonction de la température d'une huile minérale avec différents teneurs en eau.

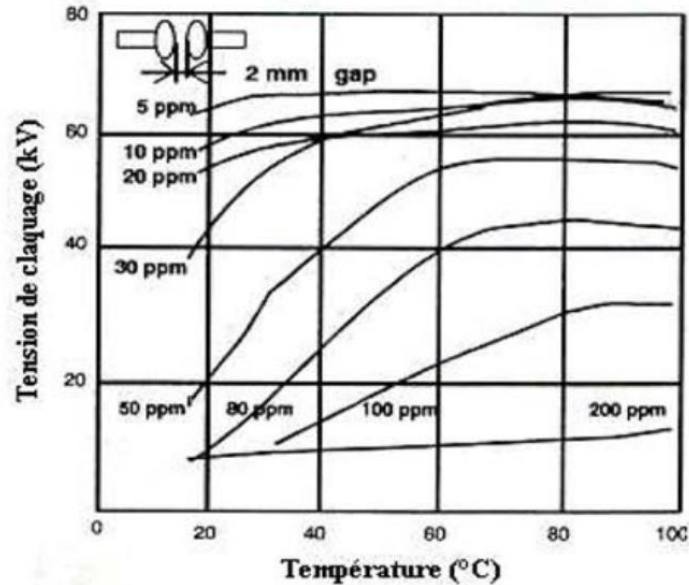


Figure I.7. variation de la tension de claquage en fonction de la température pour différentes teneurs en eau [39]

Il faut tenir compte du fait que la conductivité d'un liquide isolant augmente avec la température. Ainsi, la plupart des spécifications concernant les huiles indiquent une valeur de résistivité ou de $\tan\delta$ à 90 °C. De ce fait, si la température dépasse un certain seuil, la tension de claquage de l'huile risque de diminuer.

c) Teneur en eau

La présence d'un peu d'humidité dans les huiles diminue considérablement la rigidité diélectrique. La figure I.8 illustre l'influence de l'humidité sur la rigidité diélectrique à 50 Hz.

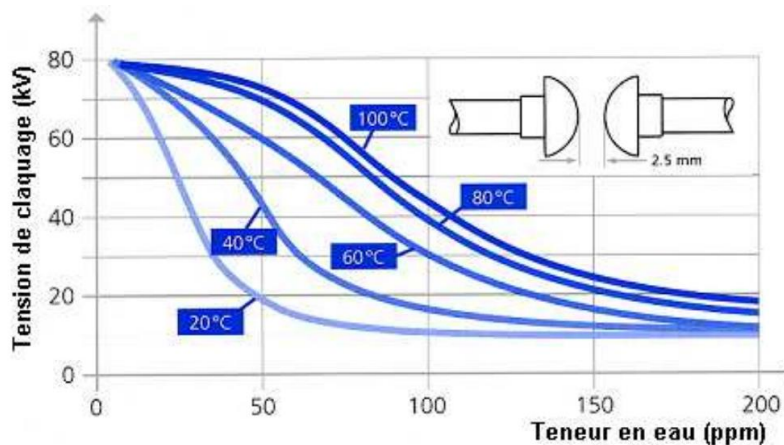


Figure I.8. Tension de claquage d'une huile en fonction de la teneur en eau pour différentes températures [35]

d) Nombre de claquages

Généralement, la tension de claquage de l'huile commence à augmenter en fonction de nombre de claquage, elle reste par la suite pratiquement constante et puis elle diminue. Ceci

peut être expliqué par la présence d'humidité dans l'huile. Cette humidité disparaît au fur et à mesure que l'huile subit des claquages (échauffement et évaporation de l'eau) [38], ou bien par l'élimination de bulles de gaz. De ce fait, des claquage répétés et nombreux provoquent la dégradation de l'huile et risquent de la rendre inutilisable.

2. Permittivité relative (ϵ)

La permittivité est la constante macroscopique fondamentale qui caractérise un diélectrique. Elle exprime la possibilité de polarisation sous l'action d'un champ électrique [27]. Elle est définie par la relation suivante :

$$\epsilon = \epsilon_0 \cdot \epsilon_r$$

ϵ est la permittivité absolue de l'huile (F/m).

ϵ_0 est la permittivité du vide ($\epsilon_0 = \frac{1}{36 \pi 10^9} = 8.85 \text{ pF/m}$).

ϵ_r est la permittivité relative de l'huile (sans unité).

La permittivité relative ϵ_r d'un isolant est définie comme le rapport entre la capacité d'un condensateur rempli de cet isolant et la capacité du même condensateur dans le vide. $\epsilon_r = \frac{C_r}{C_0}$

3. Facteur de perte diélectrique ($\text{tg } \delta$)

Le facteur de perte diélectrique ou le facteur de dissipation ($\text{tg}\delta$), qui est une valeur adimensionnelle, directement lié à la résistivité et la permittivité en tension alternative.

Plus δ est petit, plus les pertes diélectriques seront petites. L'augmentation de la température de diélectrique provoque l'augmentation des pertes diélectriques.

4. La conductivité / résistivité

La conductivité d'un liquide isolant est due à la présence de charges libres qui se déplacent sous l'effet d'un champ électrique provoquant ainsi un courant de conduction. Une bonne huile doit avoir une conductivité la plus faible possible ; autrement dit, elle doit avoir une très forte résistivité. Les huiles isolantes ont une conductivité de l'ordre de 10^{-11} à 10^{-13} S/m.

5. Décharges partielles

Les décharges partielles sont dues à la présence des bulles de gaz ou autres impuretés (humidité, particules ...) dans l'isolation. Ces impuretés, peuvent être sous l'action du champ électrique le siège de décharges localisées. En ce qui concerne les gaz pouvant être présents dans l'huile de transformateur nous pouvons citer l'hydrogène (H_2), et les hydrocarbures légers

tels que le méthane (CH₄), mais aussi en plus faibles quantités de l'éthane (C₂H₆), de l'éthylène (C₂H₄) et de l'acétylène (C₂H₂) [1]. On peut aussi avoir du monoxyde ou du dioxyde de carbone.

6. Phénomène de gassing

On appelle gassing (mm³/mm), la vitesse de variation du volume de gaz pendant 120 min d'application de la tension (10 kV à 80 °C). Si le volume augmente, l'huile aura un gassing positif alors que, s'il diminue il aura un gassing négatif.

I.4.2 Les propriétés physiques

1. La Viscosité

La viscosité est un facteur très important pour le transfert de chaleur. Plus le liquide est visqueux, plus il est difficile de le faire circuler dans les radiateurs pour le refroidissement des parties actives. La viscosité des hydrocarbures est corrélée à leur masse moléculaire : plus un produit est léger, plus sa viscosité est faible [31].

Elle s'exprime par la relation suivante :

$$CST = Ct - \frac{b}{t} \quad (1.1)$$

où

CST : Viscosité en centistokes ;

C : Constante de viscosité ;

t : Durée de l'écoulement en secondes ;

b : Coefficient du viscosimètre.

L'indice de viscosité (*VI*) est une caractéristique donnant la variation de la viscosité d'une huile en fonction de la température. Une huile est d'autant meilleure que sa variation de viscosité en fonction de la température est faible.

2. Point d'écoulement

L'utilisation de dispositifs électriques à l'extérieur nécessite de connaître la viscosité des isolants liquides à basse température (en hiver la température peut atteindre -25°C, et même -60°C).

Le point d'écoulement est défini comme la température maximale au-dessous de laquelle la circulation du liquide ne peut pas être établie, donc c'est la température à laquelle les liquides se figent. Il est déterminé suivant la norme ISO 3016 [42].

3. Le point d'éclair

Le bon fonctionnement de l'équipement électrique nécessite un point d'éclair suffisamment élevé (min 135 °C), mesuré conformément à la norme ISO 2719 (méthode

Pensky-Martens en vase clos). Le point d'éclair est la température minimale pour laquelle la concentration des vapeurs émises est suffisante pour produire une déflagration au contact d'une flamme ou d'un point chaud [43].

4. Le point de feu

C'est la température à laquelle la vapeur enflamme et soutient la combustion pendant au moins 5 secondes lorsque la source d'allumage est retirée.

5. Capacité thermique :

La capacité thermique (ou bien la chaleur spécifique) se définit comme la quantité de chaleur nécessaire pour augmenter d'un Kelvin la température d'un kilogramme d'une substance. La capacité thermique croît avec la température et décroît avec la masse volumique. Elle est exprimée en joules par kilogramme kelvin. Elle peut être calculée par la relation suivante [29] :

$$C_p = \frac{(1684 - 3.39T)}{\sqrt{\rho_{15}}} \quad (1.1)$$

où

C_p est la chaleur spécifique ($J/kg.K$) ;

T est la température (K) ;

ρ_{15} est la masse volumique à 15 °C ;

la capacité thermique des liquide à 20 °C varie de 1000 à 2300 J/(kg.K).

6. Conductivité thermique :

La conductivité thermique (λ) est inversement proportionnelle à la masse volumique et décroît avec la température. Elle exprime le flux thermique s'écoulant en régime permanent sous l'effet d'un gradient thermique entre deux isothermes du liquide. Elle est comprise entre 0.11 et 0.14 $W/(K.m)$. La conductivité thermique est influencée par la température (figure I.9).

La conductivité thermique est un paramètre très important car elle détermine la capacité de l'huile à évacuer de la chaleur [44].

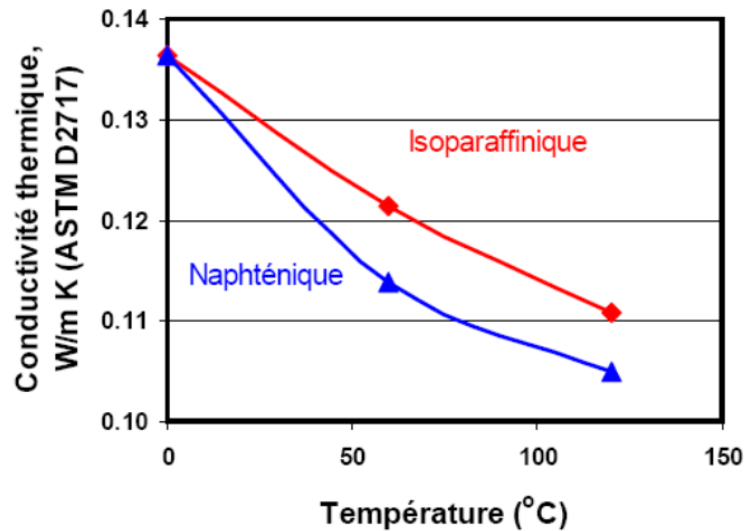


Figure I.9. Conductivité thermique en fonction de température [45]

7. Couleur :

La coloration d'une huile neuve est une indication de son degré de raffinage : les huiles très raffinées sont blanches ; une variation de la couleur dans des livraisons d'huile neuve peut indiquer un changement dans l'origine de l'huile ou dans la fabrication, sans que cela soit absolument certain. Dans le cas d'une huile usagée, une coloration foncée indique sa détérioration [46].

Les propriétés optiques présentent un intérêt particulièrement grand, car elles peuvent caractériser la constitution d'un liquide.

8. Masse volumique et coefficient de dilatation

La masse volumique ρ_θ est une caractéristique essentielle d'un produit. Elle dépend de sa composition chimique. La masse volumique est définie par le rapport d'une masse m de liquide sur son volume, mesurée à la température θ , elle est exprimée en kilogramme par mètre cube :

$$\rho_\theta = \frac{m}{V}$$

Où

m est la masse du liquide en kg ;

V est le volume de liquide en m^3 .

Selon la norme ISO 3675, la masse volumique est déterminée à 20 °C ; température à laquelle elle peut être comprise entre 800 et 2000 kg/m^3 . Les meilleures huiles ont une masse

volumique inférieure à celle de l'eau 1000 kg/m^3 . Ainsi, l'eau ne monte pas à la surface de l'huile en cas d'émulsion et ne conduit pas par conséquent à l'apparition d'arcs électriques. La masse volumique diminue quand la température augmente, elle est exprimée par la relation suivant [29] :

$$\rho_{\theta} = \rho_{20} \cdot (1 - \alpha_v \cdot [\theta - 20])$$

tel que

ρ_{θ} est la masse volumique de l'huile à la température θ ;

ρ_{20} est la masse volumique de l'huile à $20 \text{ }^{\circ}\text{C}$ (kg/m^3) ;

θ est la température de l'huile en $^{\circ}\text{C}$;

α_v est le coefficient de dilatation volumique (K^{-1}).

Le coefficient de dilatation volumique varie de $5 \cdot 10^{-4}$ à $10 \cdot 10^{-4} \text{ K}^{-1}$ et il est déterminé suivant la norme ASTM D 1903[47].

9. Aptitude de transfert de chaleur

Le dégagement de chaleur occasionné par les pertes dans le transformateur nécessite d'élaborer un système de refroidissement. Celui-ci consiste en un circuit de circulation d'un fluide qui doit être isolant. Deux types de fluide sont utilisables : l'air et l'huile. Celle-ci transporte mieux la chaleur et améliore la tenue diélectrique des cartons en les imprégnant.

I.4.3 Les propriétés chimiques

1. La teneur en eau

Il faut que la teneur en eau d'une huile soit la plus faible possible. Quand l'huile est neuve la teneur en eau est inférieure à 100 ppm ($1 \text{ ppm} = 100 \text{ mg/kg}$) selon la norme CEI 60814 [48].

$$\text{Teneur en eau } (\mu\text{g/g}) = \frac{m}{M}$$

m est la quantité d'eau titrée en microgrammes (μg) ;

M est la masse de liquide isolant en grammes (g).

2. L'acidité

L'acidité correspond aux concentrations des acides organiques et inorganiques libres présents dans l'huile et elle est exprimée en milligrammes d'hydroxyde de potassium (KOH) nécessaire pour neutraliser l'acide libre total contenu dans un gramme d'huile. Les acides proviennent généralement de la décomposition de l'huile, des produits d'oxydation et de sources atmosphériques externes. La présence de ces acides dans l'huile nuit au système d'isolation et

provoque la corrosion du réservoir d'expansion. Pour une huile neuve, elle est de l'ordre de 0,005 (mg KOH/g). Elle est mesurée par titrage calorimétrique (ISO 6618) [49], ou par titrage potentiométrique (CEI 62021-1) [50].

3. Oxydation

L'oxydation de l'huile provoque l'acidité et la formation de cambouis. Ceci peut être évité en employant des huiles avec une stabilité d'oxydation élevée réduisant au minimum le dépôt de cambouis et maximisant la durée de vie de l'isolation. La stabilité d'oxydation est mesurée selon la méthode C de la CEI 61125.

L'oxydation a des conséquences néfastes sur le bon fonctionnement des transformateurs car elle entraîne selon [51] :

- Une augmentation de la viscosité, d'où une diminution de la capacité à éliminer les dégagements de chaleur.
- Une augmentation des pertes diélectriques ;
- Une augmentation de la corrosivité vis-à-vis de la cellulose ;
- Une augmentation de la corrosivité vis-à-vis des parties métalliques.


Pour réduire l'effet de l'oxydation, des additifs anti-oxydants peuvent être incorporés à l'isolant liquide afin de réduire ou retarder sa dégradation.

I.5 Conclusion

Les contraintes électriques, thermiques, chimiques et environnementales auxquelles sont soumis les transformateurs en service peuvent engendrer, sous certaines conditions, la dégradation de leur système d'isolation.

Cette dégradation risque de conduire à l'apparition de différents défauts pouvant être identifiés selon la nature et la concentration des gaz dissous dans l'huile.

Dans le prochain chapitre nous présentons les différentes méthodes de diagnostic de l'état d'un transformateur à partir de l'analyse des gaz dissous (AGD).



Chapitre II
Méthodes de Diagnostic
Conventionnelles

II.1 Introduction

L'huile utilisée dans les transformateurs de puissance est considérée comme la mémoire des événements intervenus sur cet équipement électrique. De tels événements sont détectés en analysant et interprétant les gaz dissous, à partir de leurs concentrations et proportions, dans l'huile prélevée. En réalité, tout fonctionnement anormal du transformateur génère la production de gaz dissous dans l'huile.

L'analyse des gaz dissous (AGD) est une mesure importante dans la gestion du cycle de vie des transformateurs. L'interprétation correcte de l'analyse des gaz dissous peut fournir des informations précieuses sur l'état de santé de ces derniers. Elle permet, en fait, d'identifier et d'interpréter les défauts (thermiques et électrique) et l'origine de leur création, en vue d'éviter précocement un quelconque dysfonctionnement.

Pour passer des résultats d'analyse à une interprétation en termes de diagnostic, plusieurs méthodes ont été proposées [52]. Ce chapitre est dédié à l'application des méthodes les plus populaires consistant en les méthodes graphiques (triangle et pentagone) de Duval ainsi que celles des ratios de Dornenburg, des ratios de Rogers, de la CEI 60599 ainsi que la méthode des gaz clés et celle prescrite par la Norme IEEE C57.104.

Dans ce chapitre nous avons testé l'application des quatre méthodes suivantes : les triangle et pentagone de Duval, la méthode des ratios de Rogers et celle de la CEI 60599. Ces méthodes ont été choisies car elles ont le même nombre minimal de classe (6 en général) permettant de faire une étude comparative. Les autres méthodes ont un nombre de classes relativement petit.

Pour appliquer ces différentes méthodes, la base de données a été extraite de celles du comité technique CEI TC 10 [53] et de celle utilisée par Duval [54].

II.2 Méthode de Dornenburg

La méthode de Dornenburg est considérée comme l'une des plus anciennes méthodes (1967/1970) d'interprétation de l'analyse des gaz dissous. Trois défauts sont considérés dans cette méthode [8] :

- Décomposition thermique ;
- Décharges partielles ;
- Arc électrique.

II.2.1 Méthodologie

Cette méthode ne peut être appliquée que si la concentration de chaque gaz est deux fois supérieure à celle de seuil notée L1 et illustrée dans le tableau II.1 [8].

Tableau II.1. Seuil des concentrations de gaz dans la méthode de Dornenburg [8]

Gaz dissous	Concentration de seuil L1 (ppm)
H ₂	100
CH ₄	120
CO	350
C ₂ H ₂	35
C ₂ H ₄	50
C ₂ H ₆	65

Les quatre ratios de gaz utilisés sont : $R_1 = \frac{CH_4}{H_2}$, $R_2 = \frac{C_2H_2}{C_2H_4}$, $R_3 = \frac{C_2H_2}{CH_4}$ et $R_4 = \frac{C_2H_6}{C_2H_2}$.

Les valeurs des concentrations de gaz sont comparées aux concentrations L1 (présentées dans le tableau II.1) pour s'assurer s'il y a un problème et s'il y a une génération suffisante de chaque gaz. Les types de défauts sont déterminés selon la gamme de variation de chaque ratio (R_1 à R_4) et présentés dans le tableau II.2.

Tableau II.2. Méthode de Dornenburg pour le diagnostic de défauts [8]

Identification de défaut	$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{C_2H_2}{CH_4}$	$\frac{C_2H_6}{C_2H_2}$
Décomposition thermique	>1	< 0,75	< 0,3	> 0,4
Décharges partielles	< 0,1	Non significatif	< 0,3	> 0,4
Arc électrique	Entre 0,1 et 1	> 0,75	> 0,3	< 0,4

L'organigramme de la figure II.1 illustre l'application de la méthode des ratios de Dornenburg pour un échantillon d'huile quelconque [6].

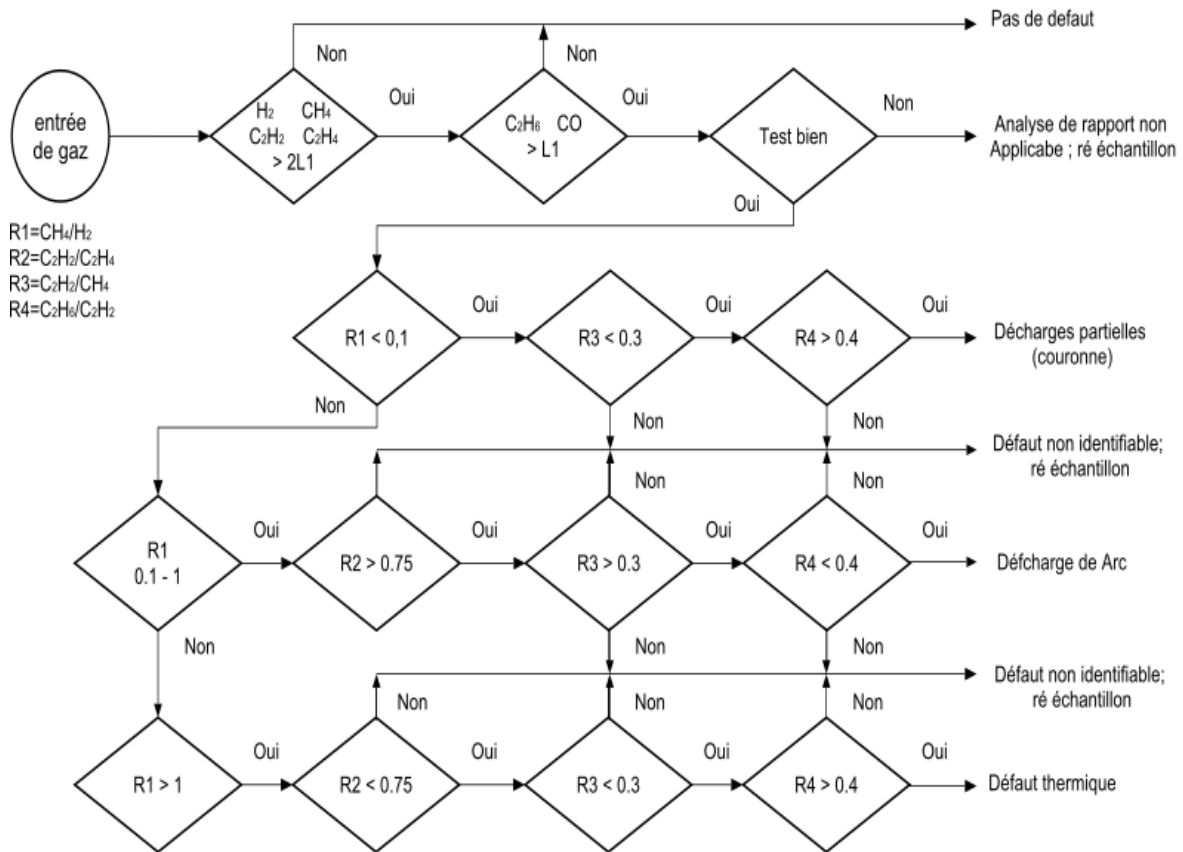


Figure II.1. Organigramme de la méthode de Dornenburg [6]

L'algorithme peut être récapitulé en quatre principales étapes données comme suit :

Étape 1 : Rassembler tous les échantillons contenant les six gaz nécessaires dans l'analyse des gaz dissous : H_2 , CH_4 , C_2H_4 , C_2H_2 , C_2H_6 et CO avec leurs concentrations exprimées en ppm.

Étape 2 : Pour un échantillon donné, si au moins une des concentrations des gaz H_2 , CH_4 , C_2H_4 et C_2H_2 est deux fois supérieure à (aux) valeur(s) limite(s) $L1$ correspondante(s) (Tableau II.1) et au moins un des autres gaz (C_2H_6 et CO) dépasse la (les) valeur(s) limite(s) $L1$ correspondante(s), l'échantillon est considéré défectueux. Il faut calculer les ratios et aller à l'étape 3 pour vérifier leur validité (ceci revient à vérifier si les ratios sont significatifs ou non). Dans le cas contraire, l'échantillon est considéré non défectueux ; aucune analyse et interprétation n'est préconisée.

Étape 3 : Déterminer la validité de la procédure de ratio : si au moins un des gaz de chaque ratio R_1 , R_2 , R_3 ou R_4 dépasse sa valeur limite (Tableau II.2) la procédure de ratio est valide ; les ratios sont donc significatifs. Dans ces conditions, il faut aller à l'étape suivante. Dans le cas contraire, les ratios ainsi obtenus ne sont pas significatifs.

Étape 4 : Il s'agit de la phase de diagnostic qui consiste à identifier le(s) défaut(s) conformément aux conditions (sur les ratios) présentées dans le tableau II.2.

Il est important de noter que les étapes 2 à 4 doivent être appliquées à l'ensemble des échantillons collectés.

II.2.2 Les inconvénients de la méthode de Dornenburg :

Comme inconvénient, la méthode de Dornenburg peut présenter des cas non interprétables. En outre, elle ne couvre pas tous les défauts possibles, du fait que son interprétation est très restreinte et ne concerne que les trois défauts suivants : décomposition thermique, décharges partielles et arc électrique.

II.3 Méthode de Rogers

Rogers a initialement développé quatre rapports CH_4/H_2 , C_2H_6/CH_4 , C_2H_4/C_2H_6 et C_2H_2/C_2H_4 avant de supprimer le ratio C_2H_6/CH_4 qui n'a pas servi suffisamment dans l'identification de défauts [7].

En général, la procédure de cette méthode est similaire à celle déjà utilisée par Dornenburg. La différence est que la méthode de Rogers peut être appliquée même lorsque les concentrations des gaz ne dépassent pas les valeurs limites de seuil rapportées dans le tableau II.1. Comme avec la méthode de Dornenburg, les ratios de Rogers peuvent être non interprétables.

II.3.1 Méthode de Rogers pour quatre ratios

Rogers a proposé plusieurs codes (0, 1, 2 et 5) pour les ratios. Chaque code correspond à une gamme de variation de chaque ratio. Les gammes des différents ratios ainsi que leurs codes sont illustrés dans le tableau II.3.

Tableau II.3. Gammes des ratios et leur code selon Rogers [7]

Ratios des gaz	Gamme	Codes
$R_1 = \frac{CH_4}{H_2}$	$R_1 \leq 0,1$	5
	$0,1 < R_1 < 1$	0
	$1 \leq R_1 < 3$	1
	$R_1 \geq 3$	2
$R_2 = \frac{C_2H_6}{CH_4}$	$R_2 < 1$	0
	$R_2 \geq 1$	1
$R_3 = \frac{C_2H_4}{C_2H_6}$	$R_3 < 1$	0
	$1 \leq R_3 < 3$	1
	$R_3 \geq 3$	2
$R_4 = \frac{C_2H_2}{C_2H_4}$	$R_4 < 0,5$	0
	$0,5 \leq R_4 < 3$	1
	$R_4 \geq 3$	2

La dernière colonne du tableau II.3 représente le code des quatre ratios pour chacune de leur plage de variation. La combinaison de codes pour les quatre ratios à la fois peut être liée à une interprétation diagnostic comme montrée dans le tableau II.4.

Tableau II.4 Diagnostic de défaut basé sur les codes de Rogers employant quatre ratios [7]

Cas	$\frac{CH_4}{H_2}$	$\frac{C_2H_6}{CH_4}$	$\frac{C_2H_4}{C_2H_6}$	$\frac{C_2H_2}{C_2H_4}$	Interprétation
1	0	0	0	0	Normale
2	5	0	0	0	Décharges partielles de faible énergie
3	1~2	0	0	0	Légère surchauffe < 150 °C
4	1~2	1	0	0	Surchauffe de 150 à 200 °C
5	0	1	0	0	Surchauffe de 200 à 300 °C
6	0	0	1	0	Surchauffe au conducteur
7	1	0	1	0	Échauffement causé par la circulation de courant dans les enroulements
8	1	0	2	0	Échauffement causé par la circulation de courant dans le noyau et le réservoir
9	0	0	0	1	Arc de faible énergie
10	0	0	1~2	1~2	Arc de grande énergie
11	0	0	2	2	Étincelle continue, arc
12	5	0	0	1~2	Décharges partielles

L'organigramme de la figure II.2 illustre l'application de la méthode de Rogers utilisant quatre ratios pour analyser et identifier les défauts d'un échantillon d'huile.

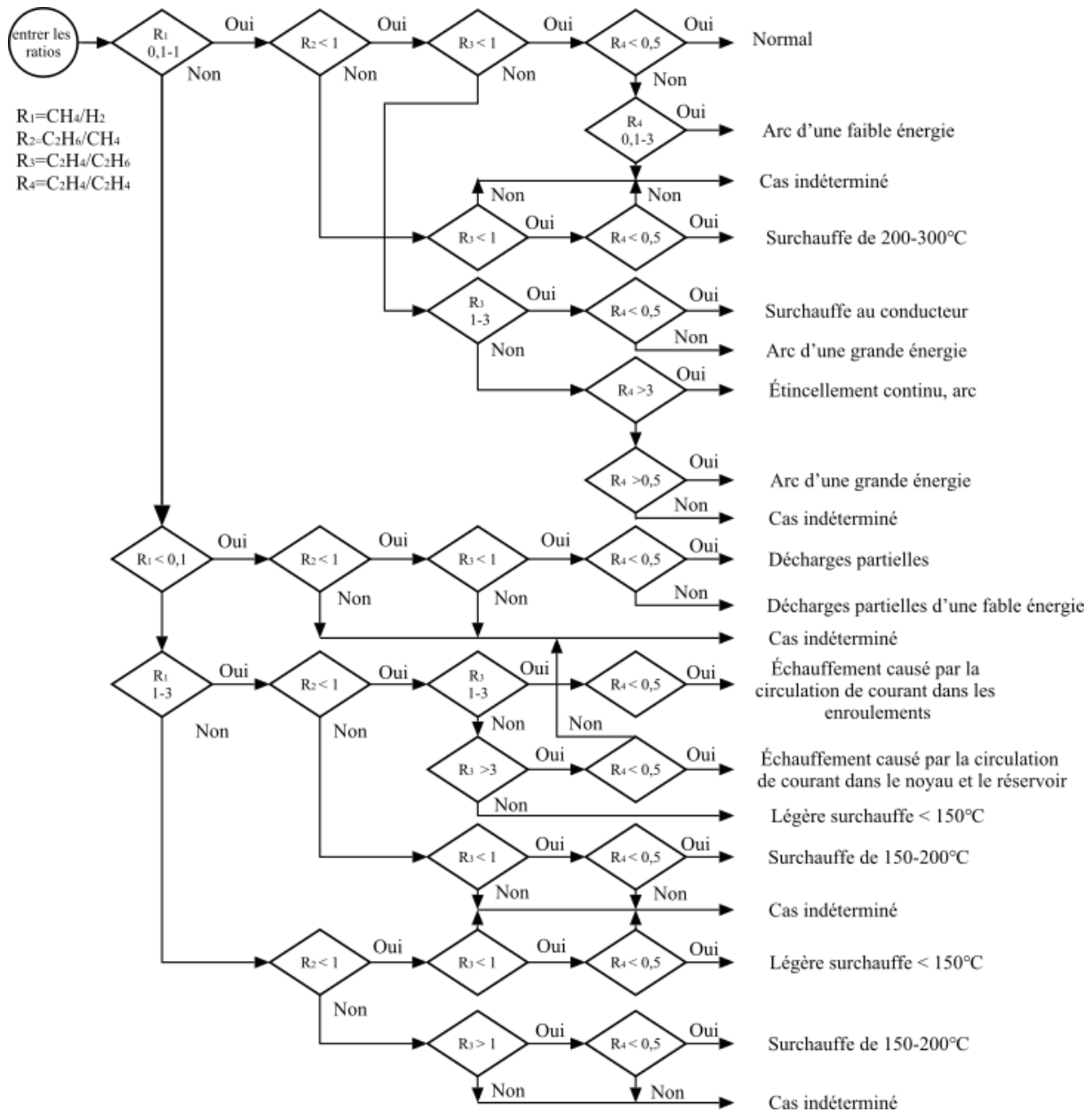


Figure II.2. Organigramme de la méthode de Rogers utilisant quatre ratios [55]

L'inconvénient de la méthode de Rogers pour quatre ratios

Sur l'ensemble de 54 combinaisons possibles (tableaux 2.3), uniquement 17 cas sont interprétables selon le tableau II.4. Les autres combinaisons des ratio R_1 , R_2 , R_3 et R_4 telles que (5,1,1,1) (1,1,1,1), etc. ne peuvent en aucun cas être interprétables. Ce qui limite l'utilisation de cette méthode.

II.3.2 Méthode de Rogers pour trois ratios

À la suite de la suppression du ratio C_2H_6/CH_4 , Rogers fait l'identification en se basant sur les trois ratios suivants CH_4/H_2 , C_2H_4/C_2H_6 et C_2H_2/C_2H_4 [7]. Ainsi, les nouveaux codes et

leurs plages de variation sont présentés dans le tableau II.5.

Tableau II.5. Gammes des ratios et leur code selon Rogers (méthode des trois Ratios) [7]

Ratios des gaz	Gamme	Code
$R_1 = \frac{C_2H_2}{C_2H_4}$	$R_1 < 0,1$	0
	$0,1 \leq R_1 \leq 3$	1
	$R_1 > 3$	2
$R_2 = \frac{CH_4}{H_2}$	$R_2 < 0,1$	1
	$0,1 \leq R_2 \leq 1$	0
	$R_2 > 1$	2
$R_3 = \frac{C_2H_4}{C_2H_6}$	$R_3 < 1$	0
	$1 \leq R_3 \leq 3$	1
	$R_3 > 3$	2

Le tableau II.6 indique l'ensemble des combinaisons des codes en vue de déterminer les défauts survenus.

Tableau II.6. Code et défauts selon la méthode de Rogers (méthode des Trois ratios) [7]

Cas	Code			Type de défaut
	R ₁	R ₂	R ₃	
1	0	0	0	Pas de défaut
2	0	1	0	Décharges Partielles (DP) avec faible énergie
3	1	1	0	Décharges Partielles (DP) avec grande énergie
4	1~2	0	1~2	Décharge électrique avec faible énergie
5	1	0	2	Décharge électrique avec grande énergie
6	0	0	1	Surchauffement avec température < 150°
7	0	2	0	Surchauffement avec température 150° < T < 300°
8	0	2	1	Surchauffement avec température 300° < T < 700°
9	0	2	2	Surchauffement avec température T > 700°

L'organigramme de la figure II.3 présente la procédure de la méthode de Rogers utilisant 3 ratios en vue d'identifier les défauts à partir de l'analyse des gaz dissous d'un échantillon d'huile.

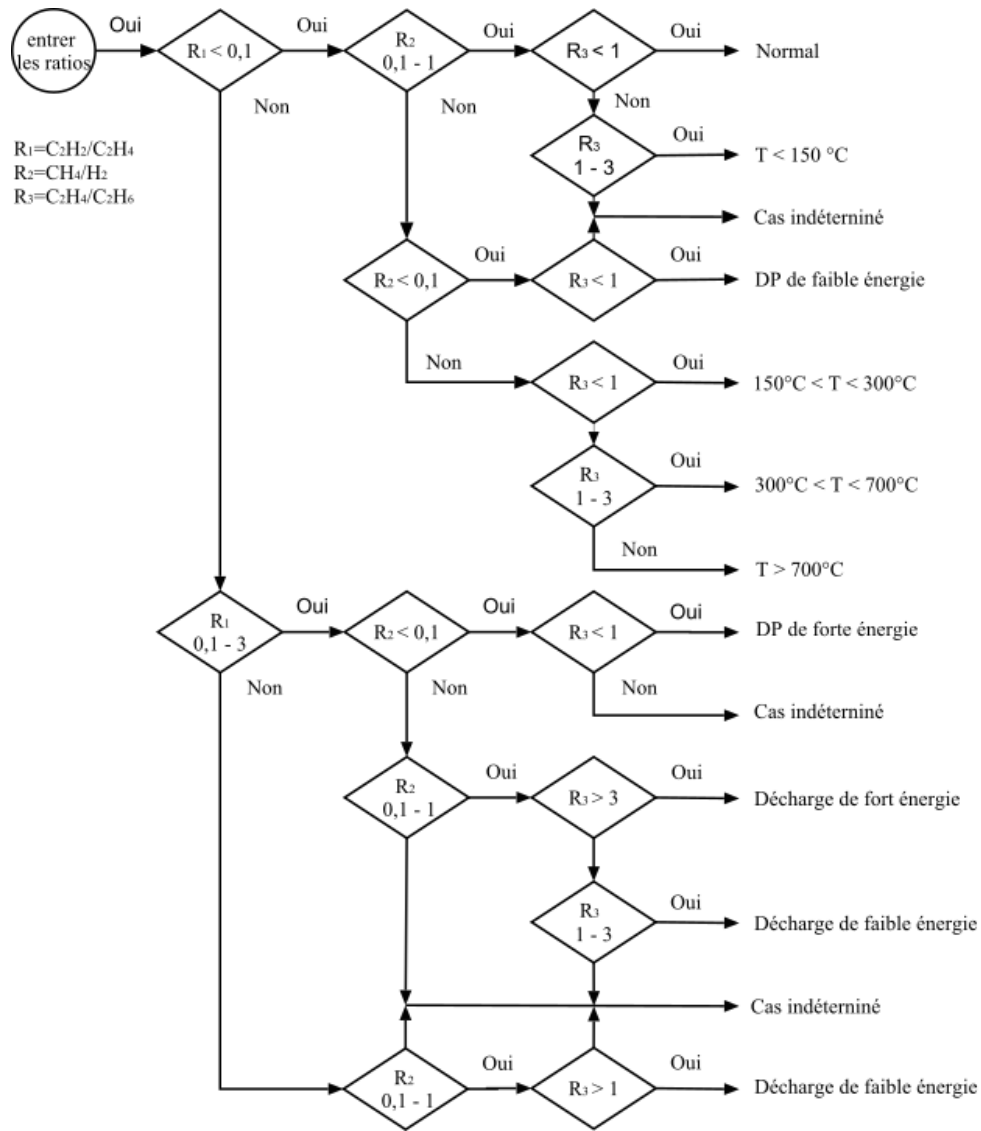


Figure II.3. Organigramme de la méthode Rogers utilisant trois ratios [56]

L'inconvénient de la méthode de Rogers pour trois ratios

Sur l'ensemble de 24 combinaisons possibles selon le tableau II.5, uniquement 11 cas sont interprétables selon le tableau II.6. Les autres combinaisons des ratio R1, R2 et R3 telles que (1,1,1) (1,0,0), etc. ne peuvent en aucun cas être interprétables. Ce qui limite l'utilisation de cette méthode.

II.4 Méthode de la CEI 60599

Cette méthode se base sur les mêmes rapports de gaz utilisés dans celle de Rogers [5]. Cependant, six défauts sont considérés et sont désignés par les indications suivantes :

- DP : décharges partielles
- D1 : décharges électriques de faible énergie
- D2 : décharges électriques de forte énergie

T1 : défaut thermique $T < 300 \text{ °C}$

T2 : défaut thermique $300 \text{ °C} < T < 700 \text{ °C}$

T3 : défaut thermique $T > 700 \text{ °C}$

Nous conserverons ces indications données par la CEI dans les méthodes qui vont suivre.

Dans le tableau II.7 nous présentons la Table d'interprétation d'Analyse des Gaz Dissous par la méthode de la CEI 60599.

Tableau II.7. Table d'interprétation d'AGD par la méthode de la CEI 60599 [5]

Cas	Défaut caractéristique	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
DP	Décharges partielles	NS ¹⁾	<0,1	<0,2
D1	Décharges de faible énergie	> 1	0,1 – 0,5	> 1
D2	Décharges de forte énergie	0,6-2,5	0,1 – 1	> 2
T1	Défaut thermique T <300 °C	NS ¹⁾	>0,1,	<1
T2	Défaut thermique 300°C < T < 700 °C	<0,1	> 1	1 – 4
T3	Défaut thermique T > 700 °C	0,2 ²⁾	> 1	> 4

¹⁾ NS = Non significatif, quelle que soit la valeur.

²⁾ une valeur croissante de C₂H₂ peut indiquer que le point chaud est à une température supérieure à 1000 °C.

II.5 Méthode de IEEE standard C57.104

Cette méthode est appelée en Anglais : Total Dissolved Combustible Gas (TDCG). La IEEE standard C57.104 prend en considération tous les gaz combustibles. Cependant, en l'appliquant, il n'est pas facile de déterminer l'état de fonctionnement normal d'un transformateur et s'il n'a aucun historique précédent de gaz dissous [6].

Un critère de quatre conditions de concentrations de sept gaz ainsi que la concentration globale des gaz dissous combustible (TDGC) correspondante a été élaboré pour classer les risques de défauts dans un transformateur [6]. La concentration de CO₂ n'est pas incluse dans la valeur de TDCG, car ce n'est pas un gaz combustible (Tableau II.8).

Tableau II.8. Conditions de concentration de gaz dissous dans l'huile selon IEEE [6]

Etat	Limites de concentration de gaz clés [$\mu\text{L/L}$ (ppm)]							
	H ₂	CH ₄	C ₂ H ₂	C ₂ H ₄	C ₂ H ₆	CO	CO ₂	TDGC
Condition 1	100	120	1	50	65	350	2500	720
Condition 2	101-700	121-400	2-9	51-100	66-100	351-570	2500-4000	721-1920
Condition 3	700-1800	401-1000	10-35	101-200	101-150	571-1400	4001-10000	1921-4630
Condition 4	>1800	>1000	>35	>200	>150	>1400	>10000	>4630

L'interprétation des conditions de 1 à 4 est décrite comme suit :

Condition 1 : Pour un TDCG situé au-dessous du niveau indiqué le transformateur est en bonne condition d'exploitation. Tout gaz dépassant les niveaux indiqués dans le tableau II.9 mérite une attention particulière.

Condition 2 : Un TDCG situé dans cette marge indique un niveau de gaz dissous plus grand que normal. N'importe quel gaz combustible individuel dépassant les niveaux spécifiques devrait inciter un complément d'enquête.

Condition 3 : Un TDCG situé dans cette marge indique un niveau élevé de décomposition de l'isolation en papier et/ou de l'huile. N'importe quel gaz combustible individuel dépassant les niveaux spécifiques mérite une attention particulière.

Condition 4 : Un TDCG dépassant cette valeur indique une décomposition excessive de l'isolation (papier/huile). Il faut dans ce cas procéder à un arrêt immédiat. En effet, le maintien en exploitation du transformateur pourrait mener à une panne de celui-ci.

II.6 Méthode des gaz-clés (Key Gas Method, David Pugh 1974)

Cette méthode emploie différents niveaux de gaz, ou « gaz clés » pour détecter des défauts. La méthode de gaz clés peut être considérée comme modification de la méthode de TDCG. Le tableau II.9 résume les gaz principaux et leurs indications de défaut [57].

Tableau II.9. Critères de diagnostic par la méthode des gaz clés [57]

Gaz clés	Défauts	Gaz détectés	Quantité de gaz en %
C ₂ H ₄	Surchauffe de l'huile	Les produits de décomposition incluent l'éthylène et le méthane, ainsi que de plus petites quantités d'hydrogène et d'éthane. Des traces d'acétylène peuvent être détectées si le défaut est grave ou comporte des contacts électriques.	C ₂ H ₄ : 63% C ₂ H ₆ : 20%
CO	Surchauffe de la cellulose	De grandes quantités d'anhydride carbonique et d'oxyde de carbone sont développées à partir de la cellulose surchauffée. Des gaz d'hydrocarbure, tels	CO : 92%

		que le méthane et l'éthylène, seront formés si le défaut implique une structure imprégnée d'huile.	
H ₂	Décharges partielles (effet de couronne)	Les décharges électriques de faible énergie produisent l'hydrogène et le méthane, avec de petites quantités d'éthane et d'éthylène. Les quantités comparables de monoxyde de carbone et de dioxyde de carbone peuvent résulter des décharges dans la cellulose.	H ₂ : 85% CH ₄ : 13%
H ₂ et C ₂ H ₂	Arcs électrique	Formation de quantités importantes d'hydrogène et d'acétylène, avec de petites quantités de méthane et d'éthylène. Il peut y avoir également formation de CO et CO ₂ s'il y a de la cellulose à l'endroit du défaut.	H ₂ : 60% C ₂ H ₂ : 30%

La figure II.4 illustre les valeurs de gaz caractéristiques normalisés dans les quatre types généraux de défaut.

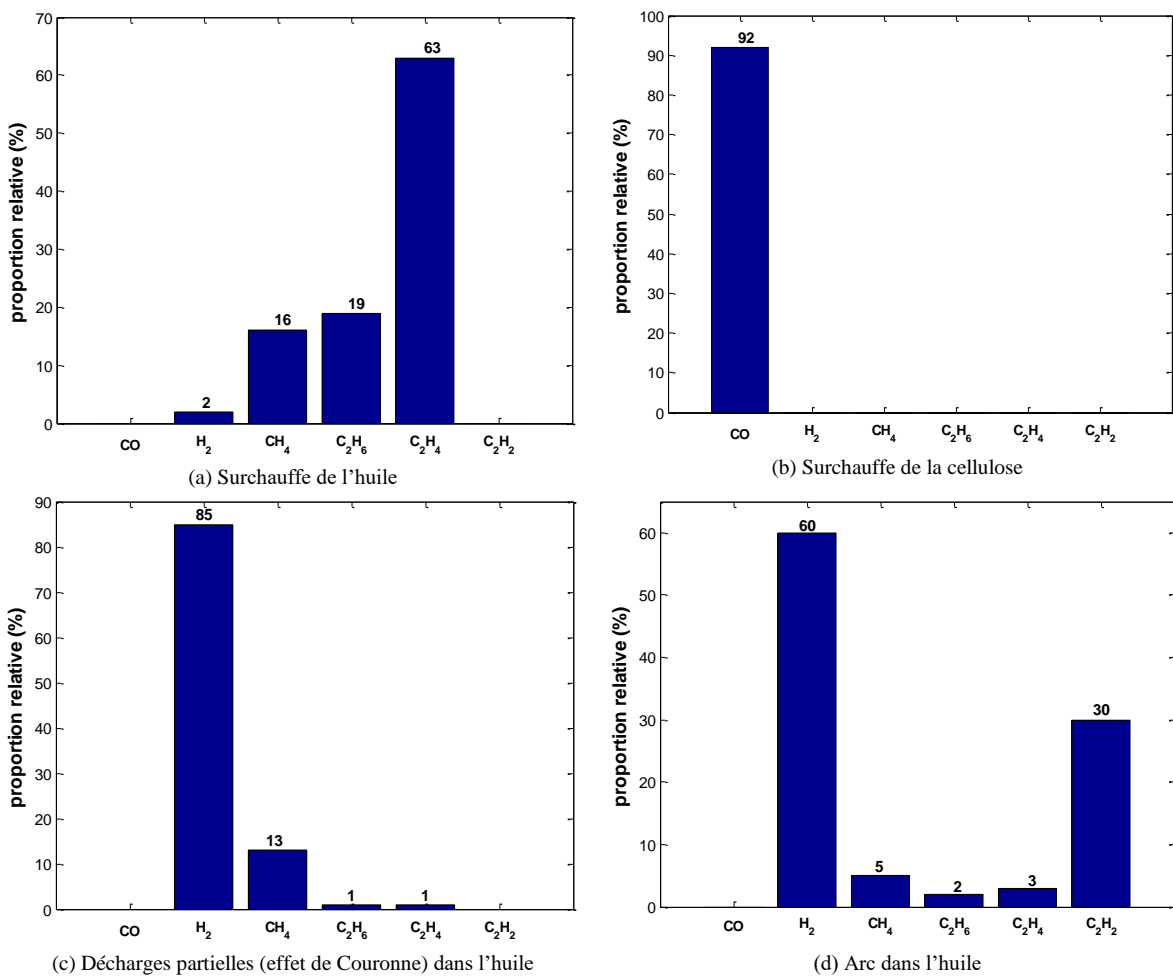


Figure II.4. Les gaz clés et leur proportions relatives [57]

II.7 Triangle de Duval

M. Duval a développé une représentation (graphique) triangulaire basée sur l'utilisation de trois gaz. Sept versions de triangle ont été développées par M. Duval, chaque version appliquée pour un type d'huile ou un cas bien défini [58]. Le tableau II.10 illustre les combinaisons des gaz utilisés dans chaque version de triangle.

Tableau II.10. Les gaz utilisés dans chaque triangle de Duval

	H ₂	CH ₄	C ₂ H ₂	C ₂ H ₄	C ₂ H ₆
Triangle 1		X	X	X	
Triangle 2		X	X	X	
Triangle 3		X	X	X	
Triangle 4	X	X			X
Triangle 5		X		X	X
Triangle 6	X	X			X
Triangle 7		X		X	X

Les triangles 3,6 et 7 ont été développés pour les huiles végétales et les huiles silicones. Dans la suite nous ne présentons que les versions développées pour les huiles minérales.

Triangle 1 de Duval : il a été développé en 1974 pour le diagnostic des transformateurs et des câbles dont l'isolation contient de l'huile minérale. Le triangle 1 de Duval utilise les gaz CH₄, C₂H₂ et C₂H₄, et les défauts détectés sont : DP, D1, D2, T1, T2, T3 et DT qui représente les défauts mixtes thermique et électrique (figure II.5).

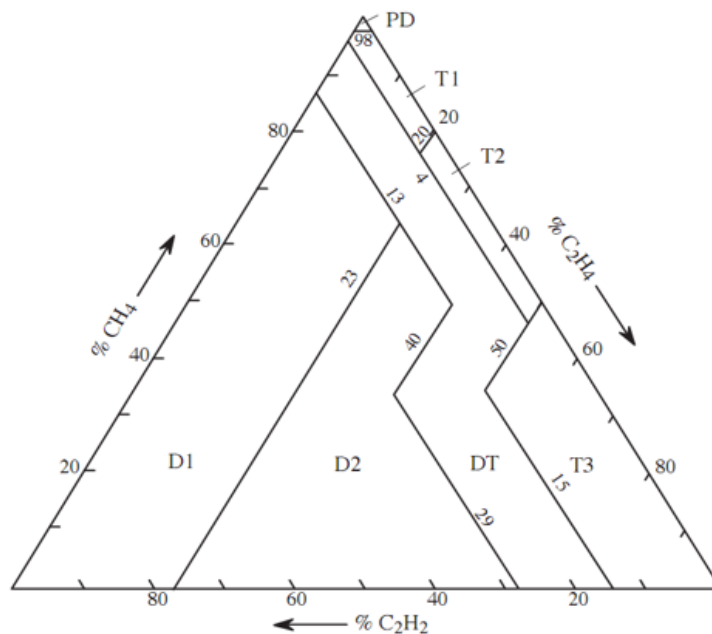


Figure II.5. Triangle 1 de Duval

Triangle 2 de Duval : il est applicable dans le cas des changeurs de prises (des transformateurs) remplis d'huile minérale. Les gaz CH_4 , C_2H_2 et C_2H_4 sont employés. Les six zones qui définissent ce triangle sont : D1, T2, T3, N (opération normale), X1 (défaut T3 ou T2 en progrès, ou arc grave anormal D2) et X3 (arc anormal D1 ou défaut thermique en progrès).

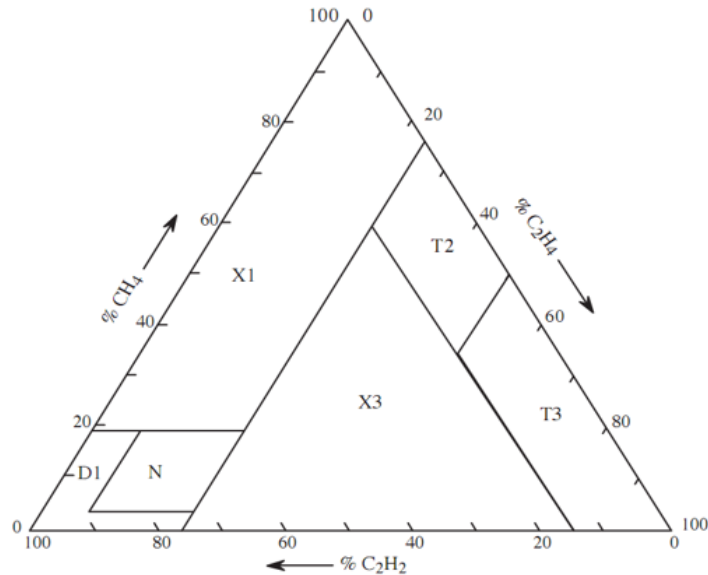


Figure II.6. Triangle 2 de Duval

Triangle 4 de Duval : Il emploie les gaz (H_2 , CH_4 et C_2H_6). Cette version est développée spécifiquement pour les défauts de basse énergie ou de température (DP, T1 et T2), afin d'obtenir plus d'informations sur ces défauts dans les huiles minérales. Il est caractérisé par cinq zones : DP, S (stray gassing), C (point chaud avec la carbonisation du papier avec $T > 300^\circ\text{C}$), O (surchauffe $T < 250^\circ\text{C}$) et la zone N/D (non déterminé). Le triangle 4 de Duval représentée sur la figure II.7.

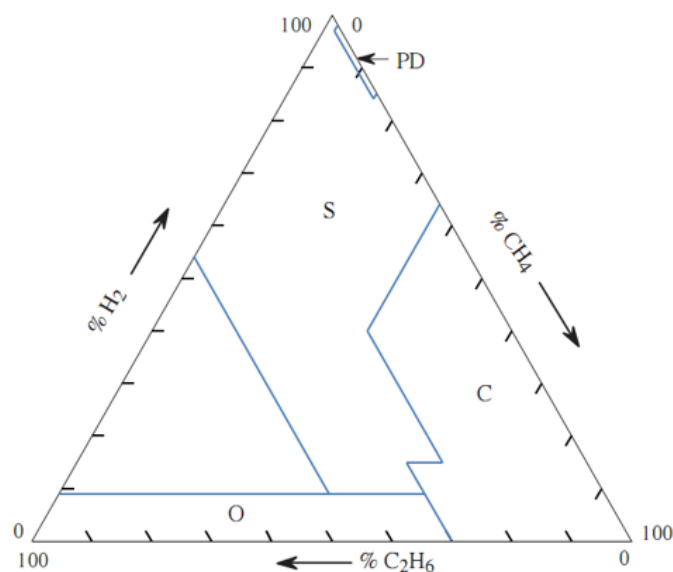


Figure II.7. Triangle 4 de Duval

Triangle 5 de Duval : Il emploie les gaz (CH_4 , C_2H_4 et C_2H_6) qui sont formés plus spécifiquement pour les défauts thermiques engendrant une surchauffe de l'huile minérale et la carbonisation du papier. Sept zones forment le triangle 5 et sont (figure II.8) : PD, T2, T3, et S, C, O, et la zone N/D comme définies dans le cas du triangle 4.

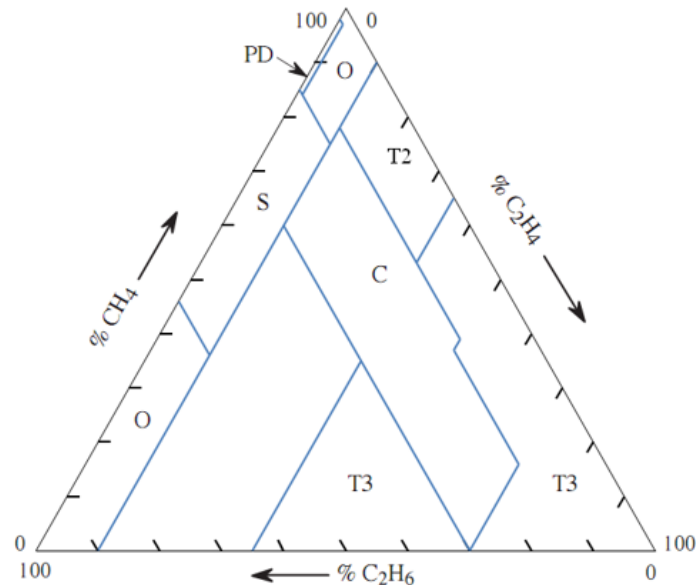


Figure II.8. Triangle 5 de Duval

II.8 Pentagone de Duval

Cette méthode a été développée en 2014 par Duval [9]. Tout comme le triangle, le pentagone de Duval est une représentation graphique permettant la visualisation des différents défauts.

Nous distinguons deux types de pentagone : le Pentagone I et le Pentagone II.

Dans le cas du Pentagone I, outre les six défauts électriques et thermiques considérés par la CEI et IEEE : PD, D1, D2, T1, T2, T3, Duval a inclus un septième défaut nommé « Stray gassing » [9], désigné par l'indication S et normalisé par la CEI en 2015 (CEI 60599) [10] (Tableau II.7).

Les défauts du pentagone II sont : DP, D1, D2, le défaut thermique T3-H dans l'huile uniquement, les défauts thermiques T3-C, T2-C et T1-C avec carbonisation de papier (C), la surchauffe de l'huile T1-O < 250°C (O) et le stray gassin S (Tableau II.11). Les gaz de décomposition de l'huile faisant l'objet de cette étude sont l'hydrogène (H_2), le méthane (CH_4), l'acétylène (C_2H_2), l'éthylène (C_2H_4), et l'éthane (C_2H_6). La méthode des pentagones I et II de Duval utilise les pourcentages des concentrations de ces gaz. Cette méthode est basée sur le principe de séparation des zones de défaut par des hyperplans [8]. Chaque type de pentagone de Duval

peut identifier sept défauts, la différence entre ces deux types de pentagone sont récapitulées dans le tableau II.11 et la figure II.9.

Tableau II.11. Comparaison des types de défauts pour les deux versions de pentagone de Duval

Pentagone I de Duval	Pentagone II de Duval
PD : décharges partielles	PD : décharges partielles
D1 : décharges à faible énergie	D1 : décharges à faible énergie
D2 : décharges à forte énergie	D2 : décharges à forte énergie
T3 : défauts thermique > 700°C	T3-H : défaut thermique T3-H dans huile uniquement
T2 : défauts thermique 300°C à 700°C	C : défauts thermique T3-C, T2-C et T1-C avec carbonisation de papier
T1 : défauts thermique < 300°C	O : surchauffe T1-O < 250°C
S : stray gassing	S : stray gassing

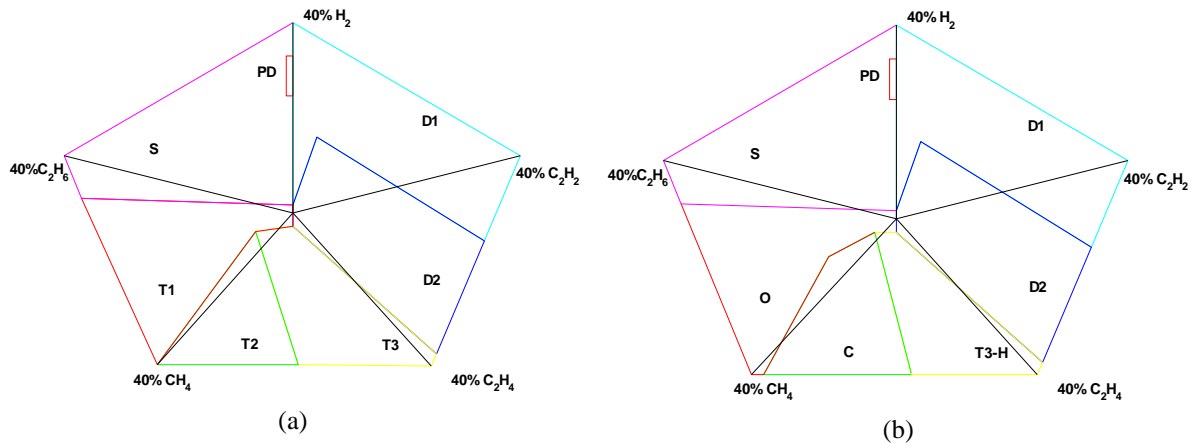


Figure II.9. Pentagone I et II de Duval [9]

Pour un échantillon donné, les séquences de calcul suivantes décrivent l'algorithme de cette méthode (valable pour les deux types de pentagone jusqu'à l'étape 5) :

Etape 1 : Calcul des %H₂, %CH₄, %C₂H₂, %C₂H₄ et %C₂H₆, tel que,

$$\%H_2 = H_2 \text{ (en ppm)} * 100 / \text{sum};$$

$$\%CH_4 = CH_4 \text{ (en ppm)} * 100 / \text{sum};$$

$$\%C_2H_2 = C_2H_2 \text{ (en ppm)} * 100 / \text{sum};$$

$$\%C_2H_4 = C_2H_4 \text{ (en ppm)} * 100 / \text{sum};$$

$$\%C_2H_6 = C_2H_6 \text{ (en ppm)} * 100 / \text{sum};$$

$$\text{sum} = H_2 + CH_4 + C_2H_2 + C_2H_4 + C_2H_6.$$

Exemple : H₂=100 ppm, CH₄=62 ppm, C₂H₂=250 ppm, C₂H₄=140 ppm et C₂H₆=90 ppm, le pourcentage relatif de chaque gaz sont respectivement : 16.87%, 9.50%, 38.34%, 21.47% et 13.80 %.

Etape 2 : Traçage de chaque pourcentage relatif à un gaz donné sur son axe respectif.

Etape 3 : Calcul des coordonnées x_i et y_i ($i=0$ à $n-1$; $n=5$ est le nombre de gaz) des pourcentages des cinq gaz :

$$\left\{ \begin{array}{l} X_0 = \% H_2 \cos \frac{\pi}{2} \\ X_1 = \% C_2H_6 \cos\left(\frac{\pi}{2} + \alpha\right) \\ X_2 = \% CH_4 \cos\left(\frac{\pi}{2} + 2\alpha\right) \\ X_3 = \% C_2H_4 \cos\left(\frac{\pi}{2} + 3\alpha\right) \\ X_4 = \% C_2H_2 \cos\left(\frac{\pi}{2} + 4\alpha\right) \end{array} \right.$$

$$\left\{ \begin{array}{l} Y_0 = \% H_2 \sin \frac{\pi}{2} \\ Y_1 = \% C_2H_6 \sin\left(\frac{\pi}{2} + \alpha\right) \\ Y_2 = \% CH_4 \sin\left(\frac{\pi}{2} + 2\alpha\right) \\ Y_3 = \% C_2H_4 \sin\left(\frac{\pi}{2} + 3\alpha\right) \\ Y_4 = \% C_2H_2 \sin\left(\frac{\pi}{2} + 4\alpha\right) \end{array} \right.$$

$$\alpha = \frac{2\pi}{5}$$

Dans notre exemple, (x_i, y_i) avec $i=0$ à 4 sont les coordonnées respectives pour H_2 , CH_4 , C_2H_2 , C_2H_4 , et C_2H_6 et sont : $(0, 20.73)$, $(- 5.59, -7.69)$, $(36.47, 11.84)$, $(12.62, -17.37)$ et $(- 1312, 4.26)$.

Etape 4 : Calcul de la surface du polygone irrégulier comme suit :

$$A = \frac{1}{2} \sum_{i=0}^{n-1} (X_i Y_{i+1} - X_{i+1} Y_i) \quad (3)$$

Le polygone irrégulier de l'exemple précédant est illustré sur la figure II.10.

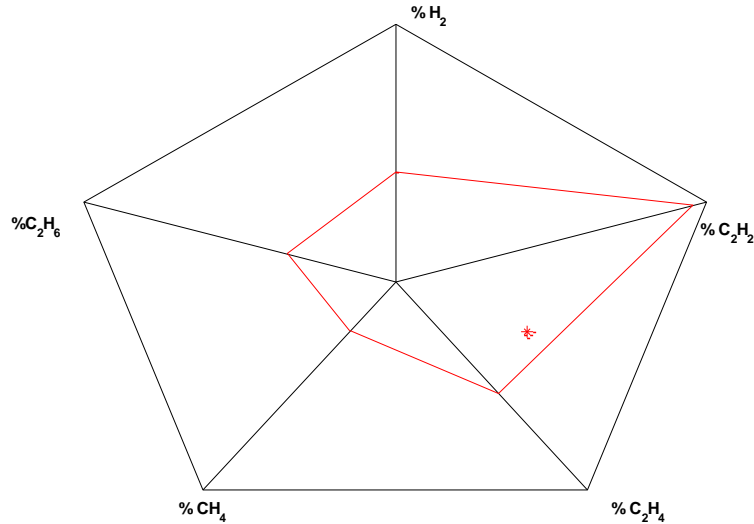


Figure II.10. Polygone irrégulier d'un échantillon

Etape 5 : Calcul des coordonnées (Cx, Cy) du centre de masse du pentagone irrégulier, utilisant les équations suivantes :

$$C_x = \frac{1}{6A} \sum_{i=0}^{n-1} (X_i + X_{i+1})(X_i Y_{i+1} - X_{i+1} Y_i) \quad (4)$$

$$C_y = \frac{1}{6A} \sum_{i=0}^{n-1} (Y_i + Y_{i+1})(Y_i X_{i+1} - Y_{i+1} X_i) \quad (5)$$

Le centre de masse de notre exemple (16.07, -7.78) est représenté avec un point rouge sur la figure II.10.

Etape 6 : Nous donnons la localisation du centre de masse calculé dans la zone de défaut qui lui correspond dans le pentagone I de Duval (Figure II.11). Il en sera de même dans le cas du pentagone II en respectant sa structure.

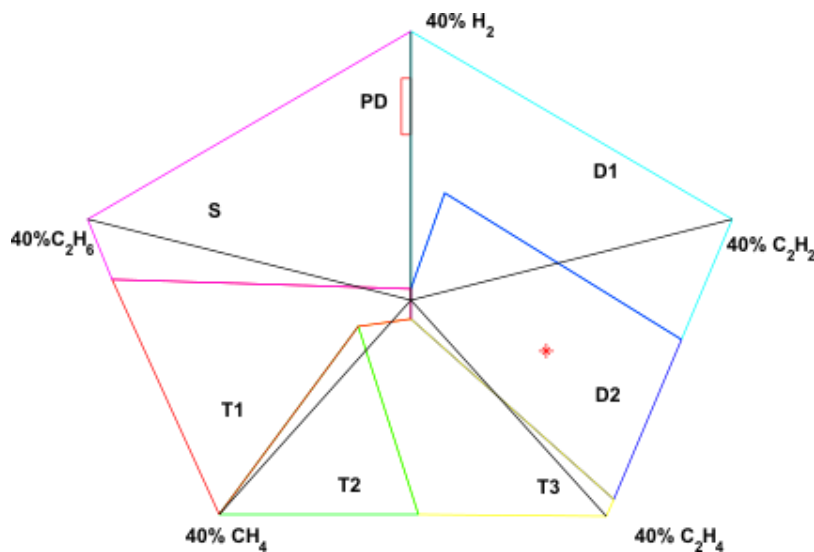


Figure II.11. Pentagone I de Duval

En reproduisant le pentagone I de Duval (cas d'exemple choisi) pour tous les échantillons, nous obtenons la figure II.12 illustrant la distribution des coordonnées des différents centres de masse.

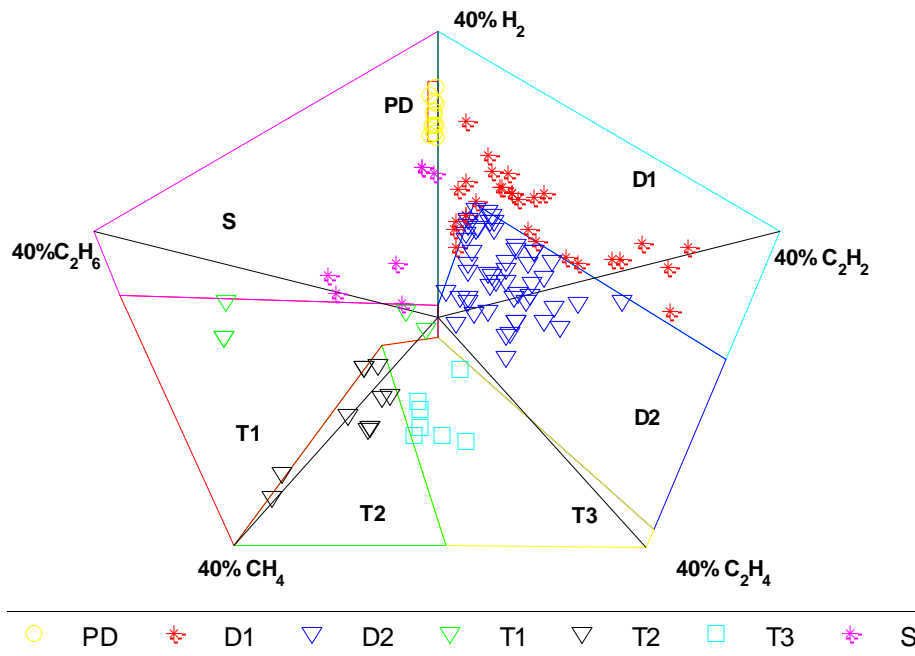


Figure II.12. Représentation du pentagone I de Duval utilisant la base de données CEI TC 10

II.9 Cas d'Application des différentes méthodes

Pour tester les méthodes du triangle de Duval, du pentagone I de Duval, des ratios de Rogers et de la CEI 60599, nous avons utilisé une base de données de 50 échantillons composée de celles de la CEI TC 10 [53] et de Duval [54].

Le tableau II.12 illustre les résultats de diagnostic basé sur l'analyse des gaz dissous dans 10 échantillons d'huile minérale (2 échantillons par type de défauts : DP, D1, D2, T1 & T2, et, T3). L'ensemble des résultats obtenus pour les 50 échantillons est présenté dans l'annexe 1.

Tableau II.12. Résultats de diagnostic pour 10 échantillons de test

Cas	H ₂ (ppm)	CH ₄ (ppm)	C ₂ H ₂ (ppm)	C ₂ H ₄ (ppm)	C ₂ H ₆ (ppm)	Inspection	Triangle de Duval	Pentagone I de Duval	Rogers	CEI 60599
1	16000	3600	0	14	670	PD	PD	S*	NF*	UD*
2	6600	1000	19	2	38	PD	D1*	PD	UD*	UD*

3	60	10	4	4	4	D1	D1	S*	UD*	D1
4	78	20	28	13	11	D1	D1	D1	UD*	UD*
5	7150	14440	1760	2110	97	D2	D2	D1*	D2	D1*
6	755	229	460	404	32	D2	D2	D2	D2	D1*
7	3420	7870	33	6990	15000	T1&T2	T1&t2	T1&T2	T3*	T3*
8	3675	6392	5	7691	2500	T1&T2	T3*	T3*	T3*	T1&T2
9	150	22	11	60	9	T3	T3	D1*	T1&T*	T1&T*
10	2500	10500	6	13500	4790	T3	T3	T3	D2*	UD*

La méthode de triangle de Duval offre le meilleur résultat de diagnostic, avec 8 échantillons sur 10 qui sont bien classés.

Le pentagone I de Duval présente uniquement 5 types de défauts en accord avec l'inspection. Cependant, les méthodes de Rogers et de la CEI 60599 n'indiquent que deux échantillons sur 10 qui concordent avec l'inspection.

Le taux de réussite final, correspondant au pourcentage de cas identiques à l'inspection, des 50 échantillons est présenté pour chaque méthode dans le tableau II.13. Le triangle de Duval fournit un taux de réussite de 90% (45 échantillons bien classés), suivi par le pentagone de Duval avec 40 échantillons qui sont en accord avec les inspections (taux de réussite de 80%) et finalement par les méthodes de Rogers et la CEI 60599 avec un taux de réussite de 52% et 50% respectivement.

Tableau II.13. Taux de réussite

	Triangle de Duval	Pentagone de Duval	Rogers	CEI 60599
Taux de réussite (%)	90	80	52	50

La figure II.13 illustre le taux de réussite par type de défaut pour les quatre méthodes de diagnostic appliquées et permet d'avoir une meilleure vision sur les résultats obtenus à partir de chaque méthode et de prendre les décisions adéquates pour parer aux défaillances engendrées par chaque défaut.

Comme nous pouvons le constater, il est favorable d'utiliser la méthode du triangle de Duval pour les défauts DP, D1, D2 et T3, et celle de la CEI pour le défaut T1&T2. Le pentagone I de Duval fournit le même résultat que le triangle de Duval en ce qui concerne le défaut DP.

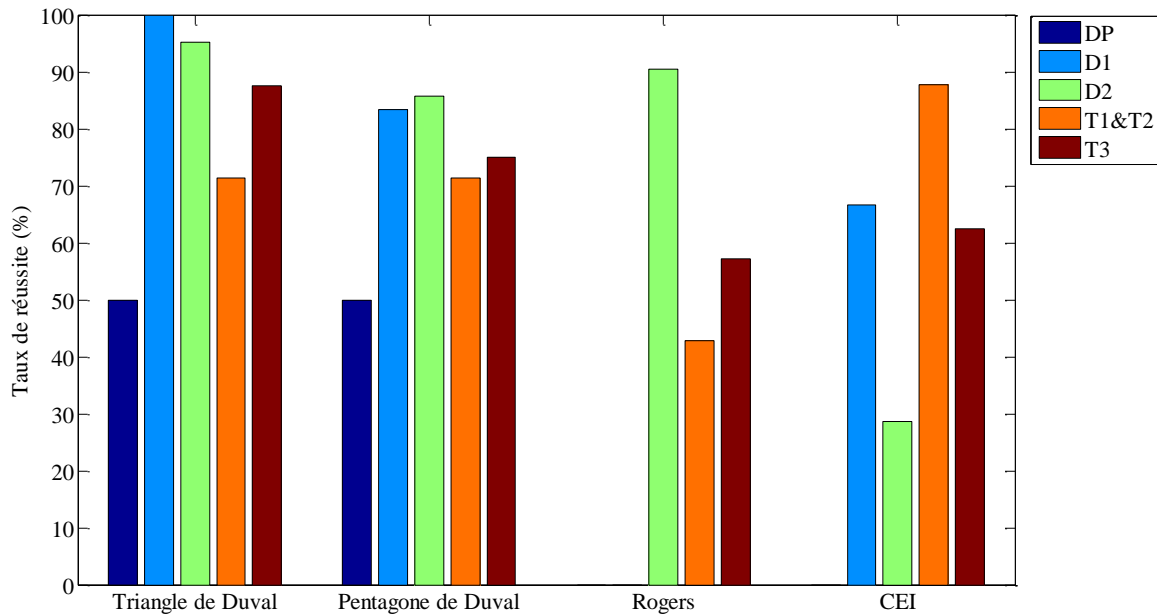


Figure II.13. Taux de réussite par type de défaut et par méthode

II.10 Conclusion

Ce chapitre a porté sur le diagnostic des transformateurs de puissance par les méthodes conventionnelles de Dornenburg, Rogers, CEI, triangle et pentagones de Duval, IEEE « TDCG » et Gaz clés.

Les avantages et les inconvénients de ces méthodes ont été suffisamment résumés. Nous avons comparé les résultats de diagnostic des quatre méthodes suivantes : pentagone I de Duval, triangle de Duval, méthode de Rogers et la CEI 60599.

Les limites de ces méthodes imposent de les combiner à des systèmes plus performants de diagnostic tels que les techniques d'intelligence artificielle (IA). La combinaison des méthodes conventionnelles avec les techniques (IA) fera l'objet du quatrième chapitre.



Chapitre III

Techniques de Classification et d'Optimisation

III.1 Introduction

Les algorithmes d'apprentissage pour la classification sont une branche de l'intelligence artificielle (IA). Ils sont appliqués dans de nombreux domaines en raison de leur capacité à traiter tous les types de données. Plusieurs algorithmes ont été développés aussi pour améliorer l'exactitude de la classification.

Nous présentons dans ce chapitre les fondements théoriques des trois algorithmes de classification que nous avons utilisés, à savoir : le classifieur machine à vecteurs de supports (Support Vector Machines « SVM »), le classifieur « k plus proches voisins » (K Nearest Neighbors « KNN ») et le classifieur bayésien naïf (naive Bayes). Nous présentons aussi la technique d'optimisation par essaim particulière (Particle Swarm Optimization « PSO ») utilisée pour ajuster les paramètres du classifieur SVM.

III.2 Classifieur Machines à Vecteurs Supports

Les machines à vecteurs supports ou séparateurs à vastes marges (en anglais Support Vector Machines, SVM) sont des techniques d'apprentissage statistique supervisé. Ces techniques sont destinées à résoudre des problèmes de discrimination et de régression traitant des données de grande dimension. Les Classifieurs SVM ont supplanté les réseaux de neurones et les autres techniques d'apprentissage [59]. Les SVM sont initiées par V. Vapnik en 1995 [60]. Elles sont largement répandues en apprentissage statistique et ont eu beaucoup de succès dans quasiment tous les domaines où elles ont été appliquées. Elles ont été utilisées, entre autres, dans le diagnostic médical (analyse de séquences d'images IRM, détection d'arythmie cardiaque, évaluation des risques de cancer, ...), dans la reconnaissance de forme (traitement d'image et de la parole), etc. [61-64].

Les classifieurs SVM ont été d'abord développées pour les problèmes de classification à deux groupes ou classes (appelée également classification binaire) [60], avant qu'elles soient étendues aux classes multiples ou multi-classe [65]. Les Classifieurs SVM génèrent un problème d'optimisation quadratique [66]. Résoudre le problème d'optimisation quadratique conduit à trouver l'hyperplan de séparation entre deux classes avec une marge maximale. Les données peuvent être linéairement ou non linéairement séparables.

III.2.1 Classifieurs SVM pour classification binaire

- **Cas de données linéairement séparables**

Pour un ensemble de données $T = \{x_k, y_k\}_k^m$, où x_k est le vecteur d'entrée, $y_k \in \{1, -1\}$ désigne l'étiquette de la classe de x_k et m le nombre total de données, la fonction de décision $f(x)$ de SVM peut être écrite comme suit [60] :

$$f(x_k) = \langle w, x_k \rangle + b = \sum_{k=1}^m w_k \cdot x_k = 0 \quad (3.1)$$

w est le vecteur de poids et b le terme de polarisation (pour déterminer la position de séparation de l'hyperplan).

Un hyperplan définitivement séparant satisfait les contraintes suivantes :

$$f(x_k) \geq 1 \text{ si } y_k = 1 \Rightarrow x_k \text{ appartient à la classe 1}$$

$$f(x_k) \leq -1 \text{ si } y_k = -1 \Rightarrow x_k \text{ appartient à la classe -1.}$$

Ces contraintes peuvent être fusionnées, comme suit :

$$y_k (\langle w, x_k \rangle + b) \geq 1 \quad \text{pour } k = 1, 2, \dots, m \quad (3.2)$$

Comme le montre la figure III.1, la position de l'hyperplan de séparation est définie par w et b . La marge maximale (correspondant à la distance maximale entre l'hyperplan et les données les plus proches) est évaluée par :

$$\frac{2}{\|w\|}.$$

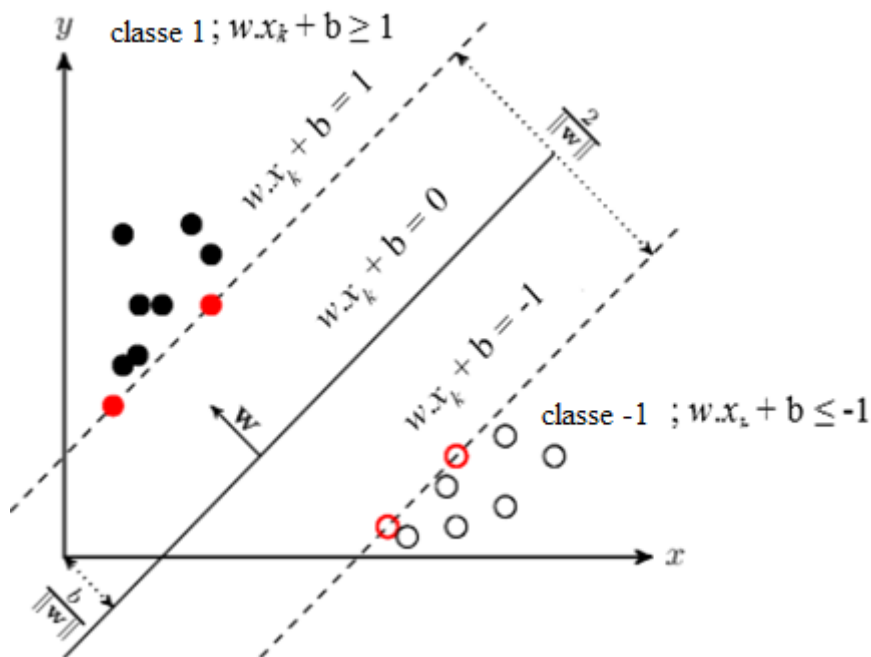


Figure III.1. Séparation linéaire de deux classes par SVM

L'hyperplan de séparation doit satisfaire les contraintes suivantes :

$$\begin{cases} y_k (< w \cdot x_k > + b) \geq 1 & k = 1, 2, \dots, m \\ \min \frac{1}{2} \|w\|^2 \end{cases} \quad (3.3)$$

$$\|w\|^2 = w^T \cdot w$$

Le principe de Lagrange est utilisé pour trouver la solution optimale pour l'hyperplan. Le Lagrange permet de transformer l'équation (3.3) en :

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{k=1}^m \alpha_k [y_k (< w \cdot x_k > + b) - 1] \quad (3.4)$$

α_k sont les coefficients de Lagrange ($\alpha_k > 0$).

Selon la condition d'optimalité :

$$\begin{cases} \frac{\partial L(w, b, \alpha)}{\partial w} = 0 \\ \frac{\partial L(w, b, \alpha)}{\partial b} = 0 \end{cases} \quad (3.5)$$

La résolution des équations (3.5) permet d'obtenir :

$$\begin{cases} w = \sum_{k=1}^m \alpha_k \cdot x_k \cdot y_k \\ \sum_{k=1}^m \alpha_k \cdot y_k = 0 \end{cases} \quad (3.6)$$

Un tel problème peut être résolu en utilisant la représentation double de Wolfe [67] :

$$\left\{ \begin{array}{l} \text{maximiser } L(w, b, a) \\ \text{sujet à } w = \sum_{k=1}^m a_k \cdot x_k \cdot y_k \\ \sum_{k=1}^m a_k \cdot x_k = 0 \\ \forall k, a_k \geq 0 \end{array} \right. \quad (3.7)$$

Ce système d'équations est équivalent à :

$$\left\{ \begin{array}{l} \text{maximiser } \sum_{k=1}^m a_k - \frac{1}{2} \sum_{k,j} a_k a_j y_k y_j \langle x_k, x_j \rangle \\ \text{sous les contraintes } \forall k, a_k \geq 0, \text{ et } \sum_{k=1}^m a_k \cdot x_k = 0 \end{array} \right. \quad (3.8)$$

- **Cas de données non linéairement séparables**

Dans le cas où les classes sont non linéairement séparables (Figure III.2), nous introduisons la marge d'erreur (ξ_k) dans l'équation de l'hyperplan (3.9), comme suit :

$$\left\{ \begin{array}{l} y_k (\langle w \cdot x_k \rangle + b) \geq 1 - \xi_k \quad k = 1, 2, \dots, m \\ \forall k, 0 < \xi_k < 1 \end{array} \right. \quad (3.9)$$

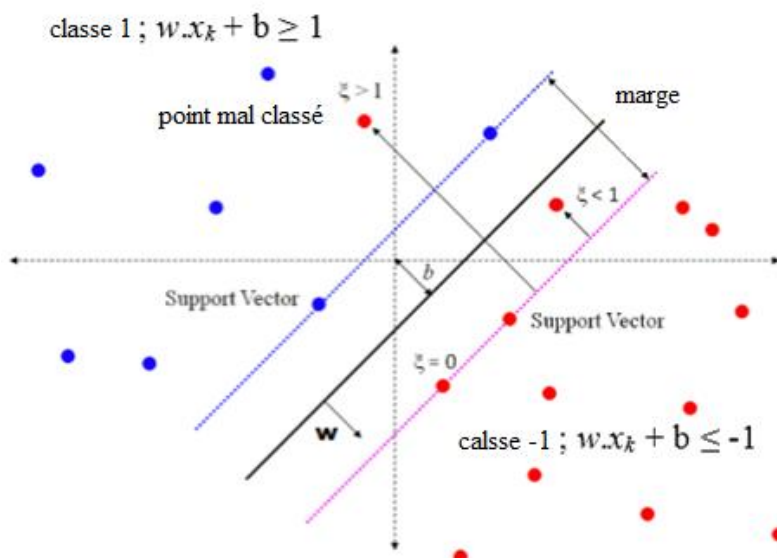


Figure III.2. Séparation non linéaire de deux classes par SVM

"C" est le paramètre de marge ou bien contrainte de régularisation (le coût attribué à un échantillon mal classé). Le problème d'optimisation devient :

$$\left\{ \begin{array}{l} \text{minimiser } \frac{1}{2} \|w\|^2 + C \sum_{k=1}^m \zeta_k \\ \text{sujet à } y_k (w \cdot x_k + b) \geq 1 - \zeta_k \\ \forall k, 0 < \zeta_k < 1 \quad k = 1, 2, \dots, m \end{array} \right. \quad (3.10)$$

De ce fait, dans le cas non linéairement séparable, la SVM ne peut pas trouver aucun hyperplan séparable. L'astuce consiste à trouver une transformation non linéaire ϕ qui représente les points originaux dans un espace de dimension beaucoup plus grande ou il sera plus facile de trouver un séparateur linéaire, comme montré sur la figure III.3.

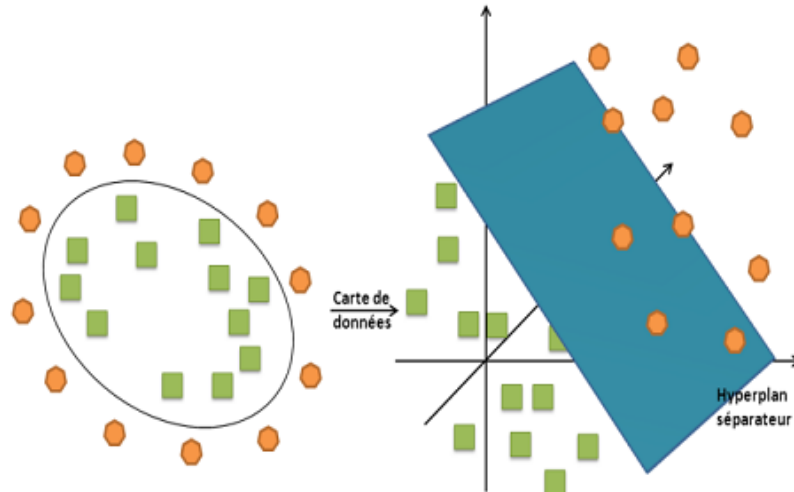


Figure III.3. Changement d'espace

Cette idée est introduite par Boser et *al.* [68]. Le but de cette idée est de placer les observations dans un espace de Hilbert T de dimension plus élevée. L'espace T ainsi obtenu est appelé espace des caractéristiques ou espace transformé. En appliquant la transformation hermitienne sur le système d'équation (3.8), le problème d'optimisation devient :

$$\left\{ \begin{array}{l} \text{maximiser } \sum_{k=1}^m \alpha_k - \frac{1}{2} \sum_{k,j} \alpha_k \alpha_j y_k y_j \langle \phi(x_k), \phi(x_j) \rangle \\ \text{sous les contraintes } : \forall k, 0 \leq \alpha_k \leq C, \text{ et } \sum_{k=1}^m \alpha_k \cdot x_k = 0 \end{array} \right. \quad (3.11)$$

Il reste à déterminer la transformation ϕ . Cependant, cette transformation pourrait être compromise si la dimension de l'espace de $\phi(x)$ est très importante. Pour contourner ce

problème, Aizerman [69] a introduit une méthode dite "astuce de noyau" (en anglais : kernel trick). En effet, sous certaines hypothèses sur ϕ , le produit scalaire $\langle \phi(x_i), \phi(x_j) \rangle$ peut être calculé à l'aide d'une fonction symétrique k (noyau). Le noyau k doit satisfaire les conditions de Mercer [70] (k doit être symétrique et semi-définie positive).

$$k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \quad (3.12)$$

Il suffit donc de remplacer le produit scalaire par la fonction noyau non linéaire. De ce fait, les classifieurs SVM peuvent trouver un hyperplan séparateur linéaire. Le système (3.11) se transforme alors sous la forme suivante :

$$\left\{ \begin{array}{l} \text{maximiser } \sum_{k=1}^m \alpha_k - \frac{1}{2} \sum_{k,j} \alpha_k \alpha_j y_k y_j k(x_k, x_j) \\ \text{sous les contraintes } : \forall k, 0 \leq \alpha_k \leq C, \text{ et } \sum_{k=1}^m \alpha_k x_k = 0 \end{array} \right. \quad (3.13)$$

En pratique, quelques familles de fonctions de noyau paramétrables sont couramment utilisées. Il revient simplement à l'utilisateur d'effectuer des tests pour déterminer celle qui convient le mieux pour son application. Dans la littérature, on trouve les fonctions noyaux suivantes [71]:

- La fonction linéaire : $K(x, y) = x \cdot y'$
- La fonction polynomiale : $K(x, y) = (c + x \cdot y')^d$
- La fonction à base radiale (Gaussienne) : $K(x, y) = \exp\left(-\frac{\|x - y'\|^2}{2\sigma^2}\right)$
- La fonction sigmoïde : $K(x, y) = \tanh(\alpha_0(x \cdot y') + \beta_0)$

III.2.2 Classifieurs SVM multi-classes

Dans la pratique, la plupart des problèmes de classification contiennent un nombre de classes supérieur à deux. Il existe plusieurs approches (stratégies) pour résoudre le problème multi-classes consistant en "tout-ensemble" (en anglais : all-together), "un-contre-un" (one-against-one) et "un-contre-tout" (one-against-all). En comparant ces trois stratégies, il s'est avéré que "un-contre-un" produit une meilleure exactitude et exige moins de nombre de vecteurs de soutien, et par conséquent temps inférieur de calcul [72]. De ce fait, cette stratégie a été adoptée dans cette présente étude. Dans ce qui suit, nous présentons le processus d'apprentissage utilisant la stratégie un-contre-un.

La stratégie un-contre-un a été initialement présentée par Knerr et al [73]. Elle a été associée à SVM pour la première fois par Mayoraz et Alpaydin [72]. Cette stratégie construit $w(w-1)/2$ classifieurs (w étant le nombre de classes) où chacun est formé par des données de deux classes. Donc il faut trouver $w(w-1)/2$ fonctions de décisions. La solution du problème multi-classes est de le reformuler afin d'obtenir une classification binaire (en deux classes). Chaque classifieur

sépare les données en deux catégories ; la première comporte une seule classe, les autres données sont fusionnées pour former la deuxième classe. La figure III.4 illustre les différents scénarios qui en résultent, pour un problème de 3 classes. En fait, la stratégie un-contre-un représente un cas spécial des méthodes de décomposition proposée par Dietterich et al [74] pour résoudre le problème de multi-classes. La figure III.4 montre comment trouver les fonctions de décisions étape par étape.

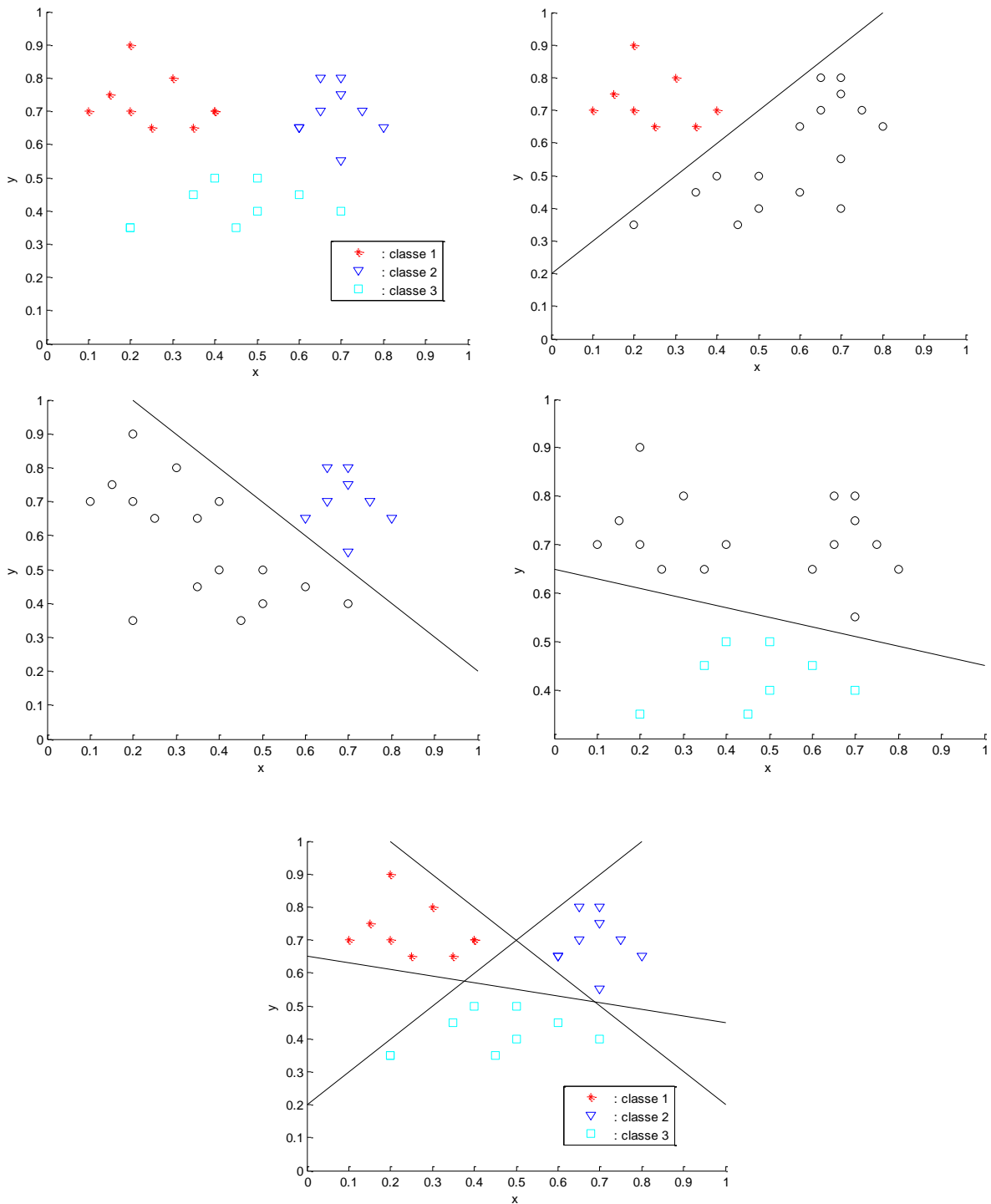


Figure III.4. Principe un-contre-un pour $w=3$

Pour un échantillon inconnu x_i , le classement est donné par le vote majoritaire ou par un graphe à cycle de décision (arbre de décision) si chaque classe reçoit le même nombre de votes.

III.2.3 Avantages et désavantages des SVM

L'algorithme SVM offre une solution au problème sous forme d'une fonction convexe. Cette dernière contient, en fait, un seul optimum global, correspondant à la minimisation du risque structurel. Il s'agit de rechercher l'hypothèse permettant les meilleures capacités de généralisation à partir d'un espace d'hypothèses donné. Ce dernier dépend du choix de la fonction noyau possédant souvent des paramètres libres, comme la largeur de bande avec un noyau gaussien. Cependant, la recherche de la valeur optimale de ces paramètres libres ne correspondre forcément à la recherche du minimum d'une fonction convexe. En effet, la recherche des valeurs idéales des paramètres optimaux de la fonction noyau, appelée communément sélection de modèles, est critique [75-77]. La sélection d'une mauvaise fonction noyau ou de mauvaises valeurs de ses paramètres peut conduire à un sur-apprentissage. Outre ce majeur désavantage, la durée de la phase d'entraînement des algorithmes SVM est très importante, notamment lorsqu'il s'agit de sélectionner les modèles et lorsque la taille de la base d'apprentissage est grande.

III.3 Classifieur K Plus Proches Voisins

L'algorithme des k plus proches voisins, en anglais « k Nearest Neighbors (KNN) », est classé parmi les plus simples algorithmes intelligents. KNN ne nécessite aucune phase d'apprentissage. Il est basé sur le calcul des distances entre les points d'échantillonnage aux voisins les plus proches de l'ensemble des points attribués [78]. La décision est basée sur le vote majoritaire des k voisins les plus proches. De nombreux types de distance peuvent être utilisés pour décider des voisins les plus proches.

La Figure III.5 illustre le principe ainsi que l'influence du choix du nombre de voisins. Le choix des 3 voisins les plus proches permet de classer l'étoile comme carré (objets à l'intérieur du petit cercle en ligne continue). Cependant, l'étoile est classée comme triangle si l'on considère 5 plus proches voisins (objets à l'intérieur du grand cercle en ligne discontinu). En effet, le choix du nombre de voisins k est un facteur principal durant le processus de classification.

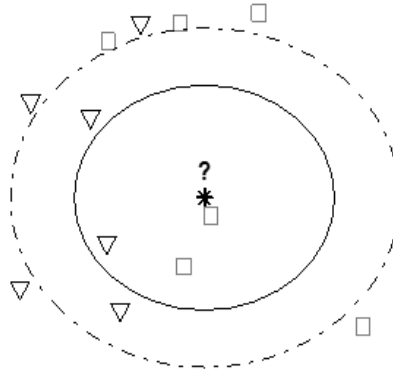


Figure III.5. Influence du nombre de voisins sur la classification

III.3.1 Avantages et inconvénients

Cette technique est facile à implémenter et appliquer à tout type de données, y compris celle complexes telles que des informations géographiques, des textes, des images, du son... Ainsi, elle est robuste au bruit. L'introduction de nouvelles données n'exige pas une reconstruction d'un modèle. La classe est attribuée à un objet avec facilité et clarté, une fois les plus proches voisins exhibés.

Les performances de cette méthode dépendent du choix de la distance, du nombre de voisins et du mode de combinaison des réponses des voisins. Les résultats pourraient être de mauvaise qualité si le nombre d'attributs pertinents est faible relativement au nombre total d'attributs, car la proximité sur les attributs pertinents sera noyée par les distances sur les attributs non pertinents [79]. En l'absence d'apprentissage, tous les calculs se font dans la phase de classification, ce qui peut prendre énormément du temps.

III.4 Classifieur Naïve Bayes

Le classifieur en question appartient à une famille d'algorithmes fondés sur le théorème de Bayes. L'attribution est de prédire la valeur des paramètres du modèle en termes de probabilité. En effet, la classification bayésienne est utilisée comme méthode d'apprentissage probabiliste.

III.4.1 Construction de classifieur naïve Bayes

La construction de classifieur est simple par l'appliquant de théorème de Bayes. Ceci concerne la probabilité conditionnelle et marginale de deux événements aléatoires. Pour un ensemble de données $D : \{(x_i, w(x_i)) : i = 1, \dots, n\}$, avec $w(x_i) \in \{w_1, \dots, w_c\}$.

x_i représente un échantillon avec ses variables caractéristiques $x_i = (x_{i1}, \dots, x_{ip})$;

w représente une classe quelconque ;

$w(x_i)$ est la classe de l'échantillon x_i ;

n dénote le nombre des échantillons et c le nombre de classes.

La décision du classifieur naïf de Bayes est définie par la probabilité postérieure $P(w_k|x)$ donnée par l'équation suivante :

$$P(w_k|x_i) = \frac{P(x_i|w_k) P(w_k)}{P(x_i)} \quad (3.20)$$

$k=1, \dots, c$.

$P(x_i|w_k)$: la probabilité conditionnelle dont l'expression est :

$$P(x_i|w_k) = \prod_{k=1}^c P(x_i|w_k) \quad (3.21)$$

$P(w_k)$: la probabilité antérieure de chaque classe qui peut être estimée selon deux possibilités :

$$P(w_k) = \frac{\text{le nombre des échantillons de la classe } k}{\text{le nombre total des échantillons}}$$

$P(w_k) = \frac{1}{c}$ C'est à dire que tous les classes possèdent la même probabilité antérieure. Pour les

deux possibilités, on a :

$$\sum_{k=1}^c P(w_k) = 1 \quad (3.22)$$

$P(x_i)$ est la densité sans conditions qui normalise les postérieurs.

$$P(x_i) = \sum_{k=1}^c P(x_i|w_k)P(w_k) \quad (3.23)$$

La règle de décision est simple ; elle consiste à choisir l'hypothèse la plus probable. De ce fait, il s'agit de la règle du maximum à posteriori ($\max P(w_k|x)$) [80].

Si, on suppose, par exemple, que le calcul final de la probabilité postérieure, pour un problème de classification de trois classes w_1 , w_2 et w_3 , a donné l'inégalité suivante : $P(w_2|x) > P(w_3|x) > P(w_1|x)$, l'échantillon x appartient, dans ces conditions, à la classe 3.

III.4.2 Avantages et inconvénients

La simplicité du classifieur Naïf Bayes consiste en l'avantage majeur de la classification bayésienne de nature probabiliste. L'hypothèse d'indépendance des variables permet de se limiter à l'utilisation de la variance de chacune d'entre elle pour chaque classe, sans pour autant calculer de matrice de covariance. Etant donné sa naïveté, le classifieur bayésien exige

relativement peu de données d'entraînement pour estimer les moyennes et les variances des différentes variables nécessaires à la classification. Or, son efficacité pourrait être compromise lors de la manipulation de bases de données vastes et/ou complexes. En outre, la classification bayésienne donne parfois de mauvais résultats si les données utilisées sont dépendantes entre elles, ou fortement corrélées.

III.5 Optimisation par les méthodes de métaheuristiques

Lorsqu'une fonction possède plusieurs optimaux locaux, les techniques d'optimisation traditionnelles risquent de converger vers l'un de ces optimaux locaux. Les méthodes métaheuristiques (ou stochastiques) sont des algorithmes d'optimisation globale. Comme avantages, ces méthodes ne nécessitent ni le point de départ, ni la connaissance de la dérivée de la fonction objectif pour atteindre la solution optimale.

Le terme métaheuristique a été inventé par Fred Glover en 1986, lors de la conception de la recherche taboue [81]. Les métaheuristiques peuvent être considérées comme des algorithmes stochastiques itératifs. Un problème d'optimisation peut être adopté avec ou sans contrainte. La plupart des métaheuristiques utilisent des processus aléatoires pour la collecte de l'information. Ces algorithmes manipulent une ou plusieurs solutions à la recherche de l'optimum. La figure III.6 illustre la différence entre un optimum globale et local (M_1 optimum globale et M_2 optimum local).

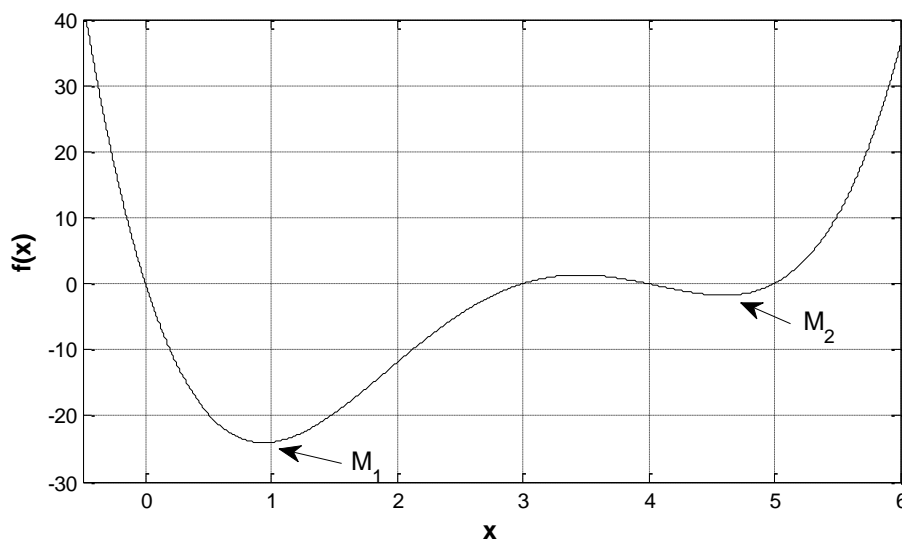


Figure III.6. Minima global et local.

Les métaheuristiques sont souvent inspirées de processus naturels qui relèvent de la physique (algorithme du recuit simulé, par exemple), de la biologie (algorithmes génétiques) ou encore de l'éthologie (algorithmes de colonies de fourmis, optimisation par essaim particulaire, etc.).

III.5.1 Algorithme d'optimisation par essaim de particules (PSO)

La technique d'optimisation par essaim particules, faisant partie de la métaheuristique, a été initialement introduite en 1995 par Kennedy et Eberhart [82]. Elle était destinée à simuler la migration et la collecte d'un troupeau d'oiseaux ou de poissons lorsqu'ils recherchent de la nourriture. Chaque oiseau rappelle ses expériences de vol et partage les distances les plus courtes entre l'aliment et lui-même. Ce comportement social, appelé intelligence d'essaims, guide tous les oiseaux afin de déterminer la distance la plus courte à la nourriture [83].

III.5.2 Principe de la technique PSO

Un essaim est un ensemble de particules positionnées dans l'espace de définition de la fonction objectif. Le principe de l'algorithme consiste à trouver la solution optimale suite à des déplacements de ces particules dans l'espace de définition. Chaque particule est localisée par ces trois attributs, en l'occurrence, sa position, sa vitesse et son voisinage.

L'algorithme PSO a subi plusieurs améliorations. L'objectif est d'améliorer la qualité de la solution en réduisant au maximum le temps de calcul. Yann Cooren [84] a proposé un nouveau mode d'initialisation de particules ; la distribution doit être régulière sur le long de l'espace de recherche. Cette nouvelle idée d'initialisation assure une couverture plus uniforme de particules. Une nouvelle stratégie hybride de déplacement de particules avec un algorithme d'estimation de distribution a été mise en œuvre pour garder la diversité au sein de l'essaim, tout au long du traitement.

Mendes et al [85] ont proposé une nouvelle topologie dite FIPS (Fully Informed Partical Swarm). Outre de la meilleure voisine et la meilleure expérience propre à la particule, cette utilise également une partie des informations de chaque voisine.

Clerc et Kennedy [86] ont proposé un model adaptatif de PSO qui utilise un nouveau paramètre " χ " appelé le coefficient de constriction. Cette idée sera éclairée dans la section III.5.4.3.

III.5.3 Configuration de l'algorithme PSO

1. Taille de la population

Etant donné l'absence de règles relatives au choix du nombre de particules allouées à la résolution du problème, nous l'avons déterminé en fonction de deux paramètres, à savoir la taille de l'espace de recherche et le rapport entre les capacités de calcul de la machine et le temps maximum de recherche (ou temps d'exécution de l'algorithme). Ce dernier dépend du nombre de résolutions de la fonction objectif, correspondant au nombre de particules multiplié par le nombre d'itérations. De ce fait, pour réduire le nombre de résolutions de la fonction objectif, il faut réduire la taille de l'essaim ou le nombre d'itérations. Mais un trop petit nombre d'essaims risque de ne pas converger vers l'optimum global.

2. Initialisation de la position et la vitesse

Il faut définir l'espace de recherche des particules de la fonction objectif pour passer à l'étape de génération de la population. Les particules sont placées au hasard (aléatoirement) selon une distribution uniforme dans l'espace de recherche. Dans un espace de recherche de dimension d , la particule i de l'essaim est modélisée par son vecteur de position $X_i = (X_{i1}, X_{i2}, \dots, X_{id})$, et par son vecteur vitesse $V_i = (V_{i1}, V_{i2}, \dots, V_{id})$.

3. Mise à jour de position et vitesse

La qualité de la position de la particule est déterminée par la valeur de la fonction objectif en cette position (en ce point). Cette particule garde en mémoire la meilleure position par laquelle elle est déjà passée, que l'on note $p_{best} = (p_{i1}, \dots, p_{i2}, \dots, p_{id})$. Par ailleurs, la meilleure position atteinte par ses particules voisines est notée $g_{best} = (g_{i1}, \dots, g_{i2}, \dots, g_{id})$.

À chaque itération, les particules mettent à jour leurs positions et vitesses de déplacement en tenant compte non seulement de leurs meilleures positions, mais aussi celles de son voisinage. La figure III.7 illustre le déplacement d'une particule.

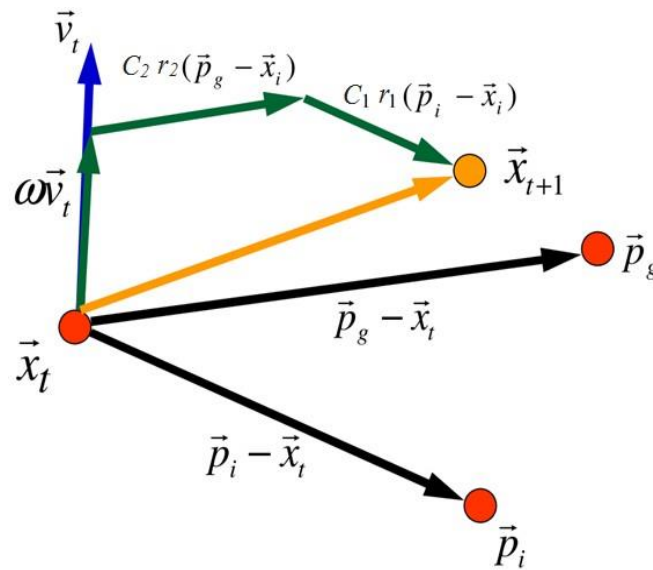


Figure III.7. Principe du déplacement possible entre deux points

La nouvelle vitesse est calculée à partir de la formule suivante [82] :

$$V_i(k+1) = wV_i(k) + c_1r_1(k)(p_{best}(k) - X_i(k)) + c_2r_2(k)(g_{best}(k) - X_i(k)) \quad (3.24)$$

Cette nouvelle vitesse permet la détermination de la position suivante comme suit [réf] :

$$X_i(k+1) = X_i(k) + V_i(k+1) \quad (3.25)$$

avec : $X_i(k)$, $X_i(k+1)$: la position de la particule P_i à l'itération k et $k+1$, respectivement ;

$V_i(k)$, $V_i(k+1)$: les vitesses de la particule P_i à l'itération k et $k+1$ respectivement ;

$P_{best}(k+1)$: la meilleure position obtenue par la particule P_i à l'itération $k+1$;

$g_{best}(k+1)$: la meilleure position obtenue par l'essaim à l'itération $k+1$;

c_1, c_2 : des constantes représentant les coefficients d'accélération ;

r_1, r_2 : des nombres aléatoires tirés de l'intervalle $[0,1]$;

$w(k)$: le poids inertiel.

Le pseudo-code de l'algorithme de PSO est donné comme suit :

Algorithme d'optimisation par essaim de particules

Initialisation de la taille de population

Initialisation aléatoire des positions et des vitesses de chaque particule

pour chaque particule i , $p_i = x_i$

***tant que** le critère d'arrêt n'est pas atteint faire*

***pour** $i = 1$ à N faire*

Déplacement de particule

$$\begin{cases} V_i(k+1) = W V_i(k) + c_1 r_1(k)(p_{best}(k) - p_i(k)) - c_2 r_2(k)(g_{best}(k) - p_i(k)) \\ x_i(k+1) = x_i(k) + V_i(k+1) \end{cases}$$

Evaluation des positions

***si** $f(x_i)$ est meilleur que $f(p_i)$*

***Alors** $p_i = x_i$*

fin si

***si** $f(g_{best})$ est meilleur que $f(p_{best})$*

***Alors** $g_{best} = p_{best}$*

fin si

fin pour

fin tant que

III.5.4 Influence des paramètres sur l'algorithme de PSO

Les paramètres affectant sur la convergence de l'algorithme de PSO vers l'optimum globale dans l'espace de recherche, consistent en :

1. Facteur d'inertie

Le facteur d'inertie $w(k)$ est introduit par Shi et Eberhart [87]. La force de vitesse de la particule est équilibrée grâce au facteur d'inertie $w(k)$. Shi et Eberhart ont trouvé qu'une meilleure convergence est obtenue pour $w(k) \in [0,8 ; 1,2]$. Selon [88], $w(k)$ peut être évalué, dans l'intervalle $[0,9 ; 0,4]$, par la formule suivante :

$$w(k) = w_{min} + (w_{max} - w_{min}) \left(\frac{k}{max_{iter}} \right) \quad (3.26)$$

w_{max} : Valeur maximale de w (= 0,9) ;

w_{min} : Valeur minimale de w (= 0,4) ;

$iter$: Itération courante ;

max_{iter} : Nombre maximum d'itérations à effectuer.

D'après la formule précédente, le facteur d'inertie $w(k)$ diminue au cours de l'exécution. En effet, au début des itérations, les particules établiront de grands déplacements. Ceci permettra d'étudier une large zone de l'espace de recherche. Quand le nombre d'itérations est augmenté, l'algorithme pourrait converger. A cet effet, il serait préférable de ralentir le déplacement, et donc diminuer le facteur d'inertie $w(k)$, de la particule pour ne pas rater l'optimum global.

2. Coefficients d'accélération c_1 et c_2

Les deux paramètres c_1 et c_2 sont des constantes. Ils influent sur la direction que va prendre la particule à chaque itération vers la meilleure position (globale ou locale). Ces constantes appartiennent à l'intervalle $[0 ; 2]$ (c'est-à-dire $0 < c_1$ et $c_2 \leq 2$). Les deux nombres $r_1(k)$ et $r_2(k)$ sont générés aléatoirement, à chaque itération, dans l'intervalle $[0 ; 1]$ par une distribution uniforme [87]. Par voie de conséquence, les deux quantités $r_1(k) * c_1$ et $r_2(k) * c_2$ sont des variables au cours des itérations. Leurs valeurs sont comprises entre 0 et 2 pour chaque itération.

3. Coefficient de constriction

L'équilibrage de l'algorithme pendant le processus de recherche est assuré par la combinaison des trois paramètres ω , c_1 et c_2 . Clerc et Kennedy [86] prouvent que la convergence peut avoir lieu en remplaçant les trois paramètres (ω , c_1 et c_2) par le coefficient de constriction χ . L'algorithme ainsi obtenu est connu sous le nom de canonical PSO. Ce dernier ne nécessite pas la connaissance de la valeur de la vitesse maximale. La nouvelle équation de la mise à jour de la vitesse est telle que [86] :

$$V_i(k+1) = \chi (V_i(k) + \phi_1(k)(p_{best}(k) - X_i(k)) - \phi_2(k)(g_{best}(k) - X_i(k))) \quad (3.27)$$

avec :

$$\chi = \frac{2}{\phi - 2 + \sqrt{\phi^2 - 4\phi}} \quad (3.28)$$

où : $\phi = \phi_1 + \phi_2$, $\phi > 4$.

Les valeurs optimales de ϕ_1 et ϕ_2 déterminées par Clerc et Kennedy sont : $\phi_1 = \phi_2 = 2,05$ soit $\phi = 4,1$, ce qui donne un coefficient de constriction $\chi = 0,7298844$.

4. Topologie de voisinage

Le réseau de communication entre les particules a pour rôle l'orientation des particules durant le processus de déplacement. Le type de communication a une influence sur les performances de l'algorithme. La communication entre les particules voisines peut être faite par deux types de structures : géométrique et social. La figure III.8 montre la différence entre un voisinage géographique et un voisinage social.

- **Voisinage géographique :** Il faut définir le nombre de voisins à l'avance. À chaque itération, la distance entre la particule " i " et les autres particules de l'essaim doivent recalculer à partir de la distance entre elles, prédéfinie dans l'espace de recherche ; cette distance change à chaque itération. Cette topologie dynamique nécessite un temps de calcul énorme [85].
- **Voisinage social :** Au début de l'algorithme, les voisinages sont prédéfinis et ne sont pas modifiés par la suite. Le but d'utiliser ce type de voisinage est d'améliorer la rapidité de convergence de l'algorithme. A l'étape de convergence, un voisinage social tend à devenir un voisinage géographique [89].

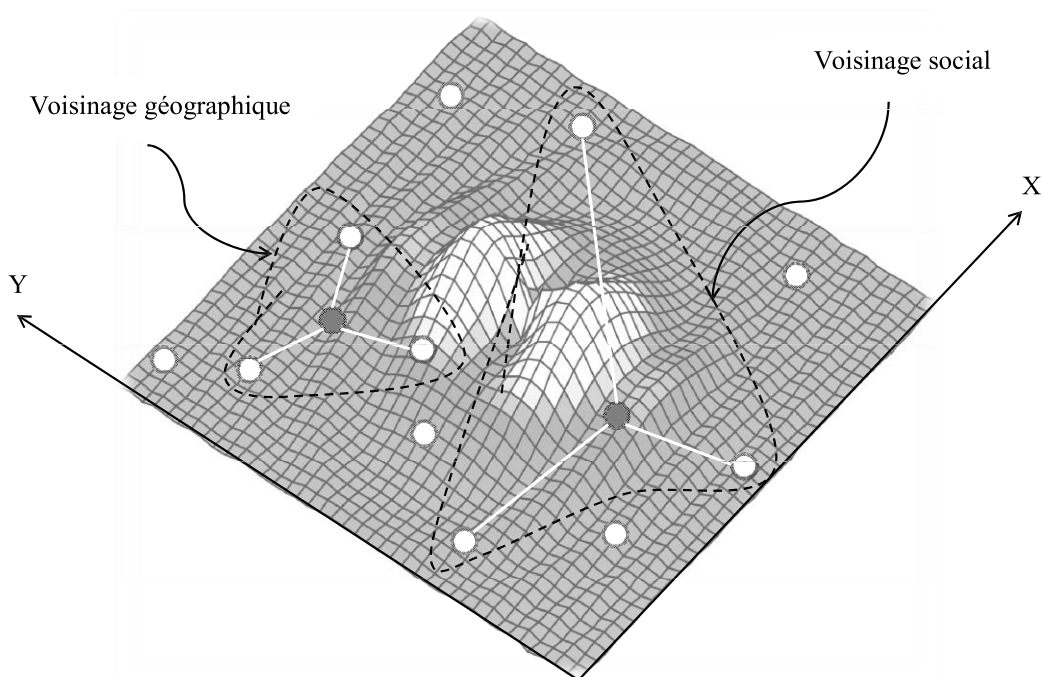


Figure III.8. Voisinages social et géographique [90]

5. Critère d'arrêt

En général, les critères d'arrêt qui sont sélectionnés consistent en [91] :

- Critère statique : le nombre maximum d'itérations ;

– Critère dynamique : la stagnation de la vitesse ;

Etant donné que nous n'avons pas eu un problème de stagnation de vitesse dans les algorithmes que nous avons élaborés, nous nous sommes basées sur le premier critère d'arrêt.

III.6 Conclusion

Dans ce chapitre, nous avons présenté les fondements théoriques essentiels pour comprendre le principe des méthodes d'apprentissage supervisé (SVM, KNN et naïve bayésienne). Nous verrons dans le quatrième chapitre l'utilisation de ces trois algorithmes aux problèmes de diagnostic des huiles de transformateurs. Nous verrons aussi comment adapter l'optimisation par essaims de particules à l'algorithme SVM pour le réglage des paramètres de ce dernier afin d'obtenir un meilleur diagnostic des huiles de transformateur de puissance.



Chapitre IV

Application des Algorithmes SVM-PSO, KNN et Naïve Bayes dans le Diagnostic des Défauts

IV.1 Introduction

L'objectif de ce chapitre est d'identifier les types de défauts en se basant sur l'analyse des gaz dissous de l'huile de transformateur de puissance. Basés sur les méthodes intelligentes et métaheuristiques, cinq algorithmes dont trois hybrides et deux indépendants ont été élaborés. Il s'agit de Pentagone de Duval-SVM-PSO, Pentagone de Duval-KNN, SVM-PSO, KNN et naïve Bayes. Le taux de réussite de chaque algorithme dépend de plusieurs paramètres tels que le nombre de voisins et le type de distance entre eux pour KNN, la fonction de noyau et les paramètres de marge et de conditionnement pour SVM, ...

Pour l'algorithme de SVM-PSO, la méthode d'optimisation métaheuristique PSO a été utilisée pour le réglage des paramètres de marge et de conditionnement de SVM. L'influence de nombre de voisins et le type de distance ont été pris en considération pour maximiser le taux de réussite de KNN. Par ailleurs, les paramètres de l'algorithme probabiliste naïve Bayes ne nécessitent aucun ajustement préalable.

IV.2 Algorithme de pentagone Duval-SVM-PSO

Le pentagone de Duval permet de calculer les coordonnées (C_x, C_y) du centre de masse. La distribution des coordonnées des différents centres de masse (Figure IV.1) est obtenue en reproduisant le pentagone de Duval pour tous les échantillons, sans ses frontières, conformément aux étapes 1 à 6 explicitement présentées dans la section II.8 du chapitre II. Ces étapes peuvent être récapitulées comme suit :

Étape 1 : Calcul de %H₂, %CH₄, %C₂H₂, %C₂H₄ et %C₂H₆.

Étape 2 : Traçage de chaque pourcentage sur son axe respectif.

Étape 3 : Calcul des coordonnées des pourcentages des cinq gaz.

Étape 4 : Calcul de la surface du polygone irrégulier.

Étape 5 : Calcul des coordonnées (C_x, C_y) du centre de masse du pentagone irrégulier.

Étape 6 : Localisation du centre de masse calculé dans la zone de défaut qui lui correspond dans le pentagone de Duval.

Basée sur une inspection visuelle, la répartition des coordonnées des centres de masse est non linéairement séparable. Les coordonnées des centres de masse, qui ne sont pas linéairement séparables, sont introduites dans un vecteur d'entrée noté $x_k = (C_x, C_y)$. Ce dernier sera injecté dans l'algorithme SVM-PSO élaboré. Nous obtenons ainsi une classification en 2D suivant cinq défauts, comme le montre la figure IV.1. L'objectif principal de l'algorithme hybride pentagone de Duval-SVM-PSO est d'améliorer le taux de réussite de diagnostic du pentagone de Duval

dans la classification des défauts. L'algorithme SVM-PSO a été choisi du fait qu'il est capable de résoudre aussi bien les cas de classification linéaires et que ceux non linéaires.

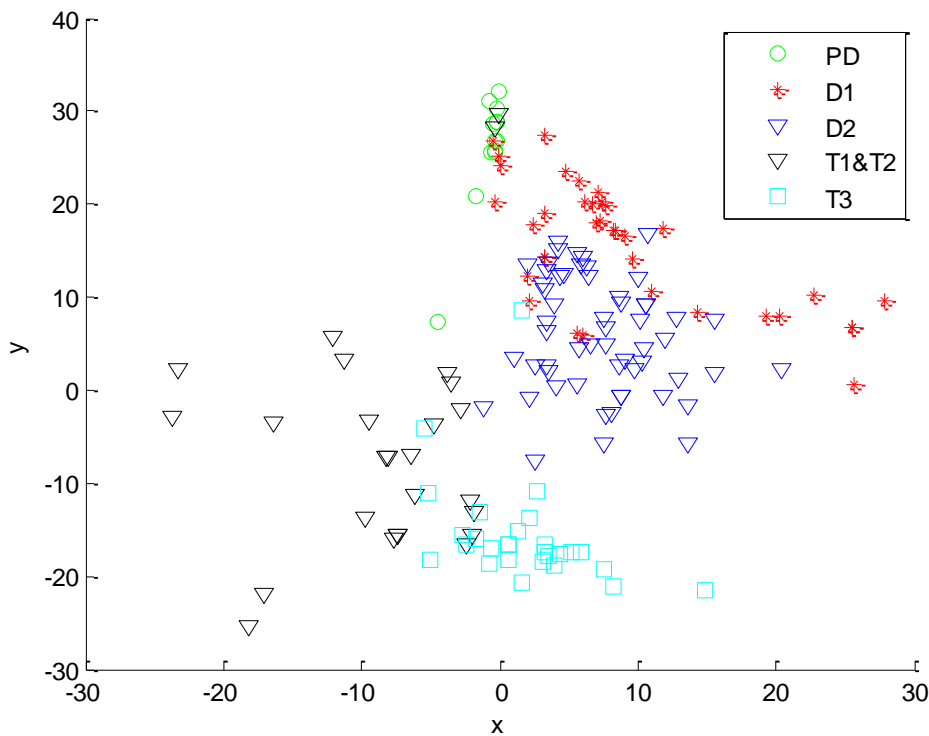


Figure IV.1. Distribution des coordonnées des centres de masse

IV.3 Optimisation des paramètres de SVM par l'algorithme d'optimisation par essaim de particules (PSO)

Afin de maximiser le taux de réussite de diagnostic, les paramètres de marge (C) et de conditionnement (λ) du modèle SVM ont été optimisés par l'algorithme PSO. Cette technique, faisant partie de la métaheuristique, simule le comportement social d'un groupe d'oiseaux ou de poissons lors d'une recherche de la nourriture [82, 83].

Les différentes étapes pour l'optimisation des paramètres SVM utilisant l'algorithme PSO sont décrites comme suit :

1. Initialisation : PSO génère, d'une manière aléatoire, une population de N particules ayant des positions et des vitesses aléatoires dans l'espace de recherche de dimensions d . En effet, la $i^{\text{ème}}$ particule possède une position $P_i = (P_{i1}, P_{i2}, P_{i3} \dots P_{id})$ et une vitesse (déplacement) $V_i = (V_{i1}, V_{i2} \dots V_{id})$; P_i et V_i sont des matrices de $N * d$. Notons que chaque position et déplacement sont représentés par des couples (paramètre de marge, paramètre de conditionnement).

2. Apprentissage : L'algorithme SVM est exécuté pour chaque particule en utilisant les paramètres de marge et de conditionnement à partir de la position actuelle de cette particule. Cette opération est répétée N fois, soit pour toutes les particules de la population. Pour cela, 68% de la base de données (correspondant à 105 sur 155 échantillons) est utilisée pour effectuer

le processus d'apprentissage.

3. Test : L'algorithme SVM est testé sur le reste de la base de données (32%, soit sur les 50 échantillons restants). La fonction objectif représentant le taux de réussite (ou d'exactitude) de diagnostic de chaque particule est calculée selon l'équation (4.1). Notons que le taux de réussite correspond au nombre d'échantillons classés avec succès sur le nombre d'échantillons total formant l'ensemble de test. Il est donné par :

$$f_i = \frac{y_t - y_{fi}}{y_t} \times 100 \quad (4.1)$$

y_t est le nombre de données de test et y_{fi} le nombre d'échantillons mal classés.

À partir de l'équation (4.1), la meilleure position (p_{ibest}) de chaque particule "i" est sélectionnée. Cela permet de déterminer la meilleure position globale (g_{ibest}) de l'ensemble constitué de N particules.

4. Mise à jour : La vitesse et la position de chaque particule sont mises à jour conformément aux équations (4.2) et (4.3) suivantes [26] :

$$V_i(k+1) = wV_i(k) + c_1r_1(k)(p_{best}(k) - X_i(k)) - c_2r_2(k)(g_{best}(k) - X_i(k)) \quad (4.2)$$

$$X_i(k+1) = X_i(k) + V_i(k+1) \quad (4.3)$$

w est le poids d'inertie, c_1 et c_2 sont deux constantes positives, appelées coefficients d'accélération, et $r_1(k)$ et $r_2(k)$ sont des variables aléatoires entre 0 et 1.

5. Critère d'arrêt : Les étapes 2 à 4 sont répétées jusqu'à atteindre le nombre maximal d'itérations fixé à l'avance.

6. Exploitation : Les paramètres optimaux de SVM ainsi obtenus sont utilisés pour la classification.

Les six étapes d'optimisation peuvent être récapitulées par l'organigramme de la figure IV.2.

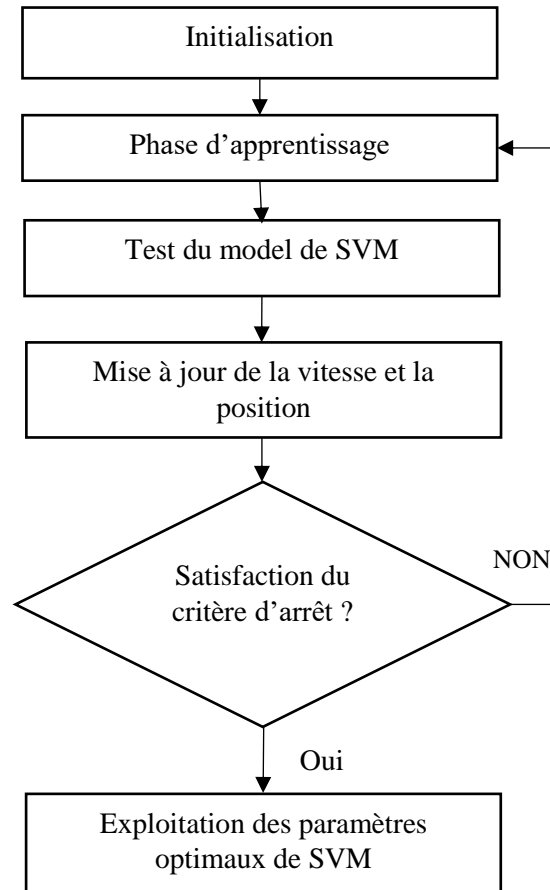


Figure IV.2. Organigramme d'optimisation des paramètres de SVM

IV.4 Algorithme pentagone Duval-KNN

Comme pour l'algorithme pentagone de Duval-SVM-PSO, dans le pentagone de Duval-KNN, les coordonnées des centres de masse sont utilisées pour établir la classification par l'algorithme de KNN.

L'organigramme de la figure IV.3 récapitule les différents algorithmes élaborés dans ce chapitre à savoir le pentagone de Duval, pentagone de Duval-SVM-PSO et finalement pentagone de Duval-KNN.

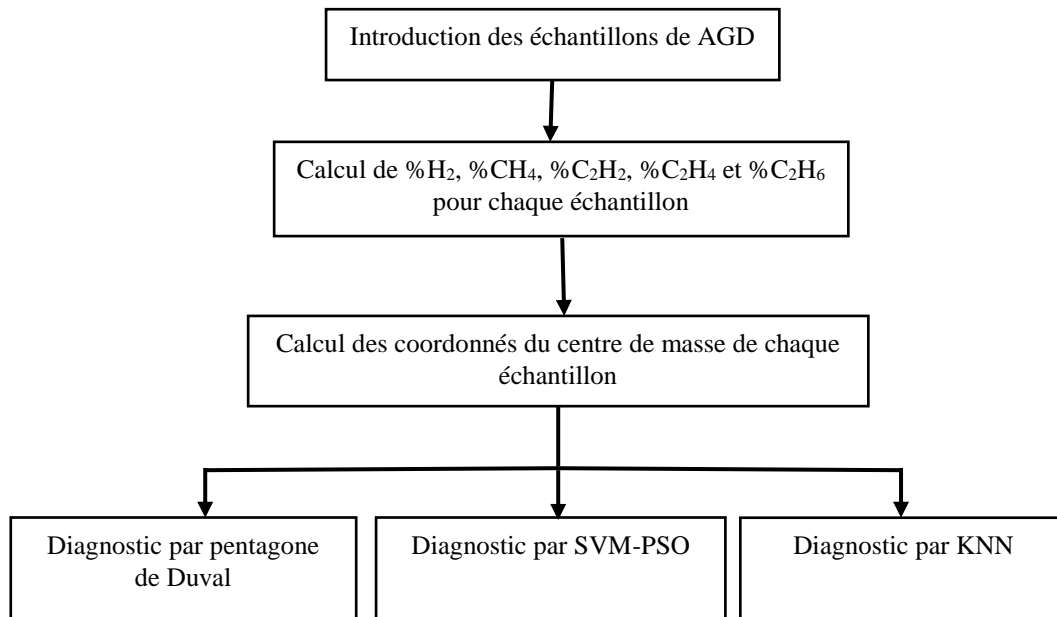


Figure IV.3. Organigramme récapitulatif des algorithmes élaborés

IV.5 Résultats de diagnostic

La base de données que nous avons utilisée est présentée dans l'Annexe 2. Elle est constituée de 50 échantillons. Dans le cas du test de l'algorithme du pentagone de Duval, nous avons obtenu un taux de réussite de 80%.

Quant au classifieur Pentagone de Duval-SVM-PSO, trois tailles de population à savoir 15, 30 et 60 ont été choisies. Pour chaque taille, l'algorithme a été exécuté pour 80 itérations. Ce nombre maximal d'itérations a été adopté comme critère d'arrêt, du fait qu'il nous a permis d'obtenir largement la convergence de cet algorithme.

Dans le but de se rapprocher à (voire d'atteindre) l'optimum global en termes du couple (paramètre de marge, paramètre de conditionnement) soit (C, λ) , l'algorithme a été répété 100 fois. Par ailleurs, l'algorithme élaboré a été exécuté pour quatre fonctions « noyau » (Kernel) consistant en : la fonction « noyau » linéaire, polynomiale, à base radiale et sigmoïde. Pour chaque fonction « noyau », trois stratégies d'apprentissage à savoir, "tous ensemble", "un contre ensemble" et "un contre un" ont été adoptées. Pour ne pas alourdir le manuscrit, nous nous limitons, dans ce chapitre, à présenter les résultats donnant le meilleur taux de réussite.

La figure IV.4 présente l'évolution de la performance de ce classifieur en fonction du nombre d'itérations, et ce pour toutes les tailles de population considérées. Pour une taille donnée de population, et avec l'augmentation du nombre d'itérations, le taux de réussite augmente depuis sa valeur initiale générée aléatoirement, jusqu'à atteindre la valeur maximale

de 88%. Cette dernière est commune pour les trois tailles de population ainsi adoptées. Il est à noter que la convergence de cet algorithme s'accroît avec l'augmentation de la taille de population. En effet, pour la taille la plus élevée (60) de population, l'algorithme a atteint le taux de 88% en uniquement 3 itérations, contre 8 et 43 pour une taille de 30 et 15 respectivement. En fait, le taux de réussite de 88% (soit une amélioration de 8% par rapport à celui obtenu lors de l'utilisation du pentagone de Duval) est le plus élevé puisqu'il correspond aux valeurs optimales des paramètres de marge et de conditionnement de SVM ($C=27,922021$ et $\lambda=0,00047$). Ce taux a été trouvé en utilisant la fonction noyau "polynomiale de degré 1" et la stratégie d'apprentissage "un contre un".

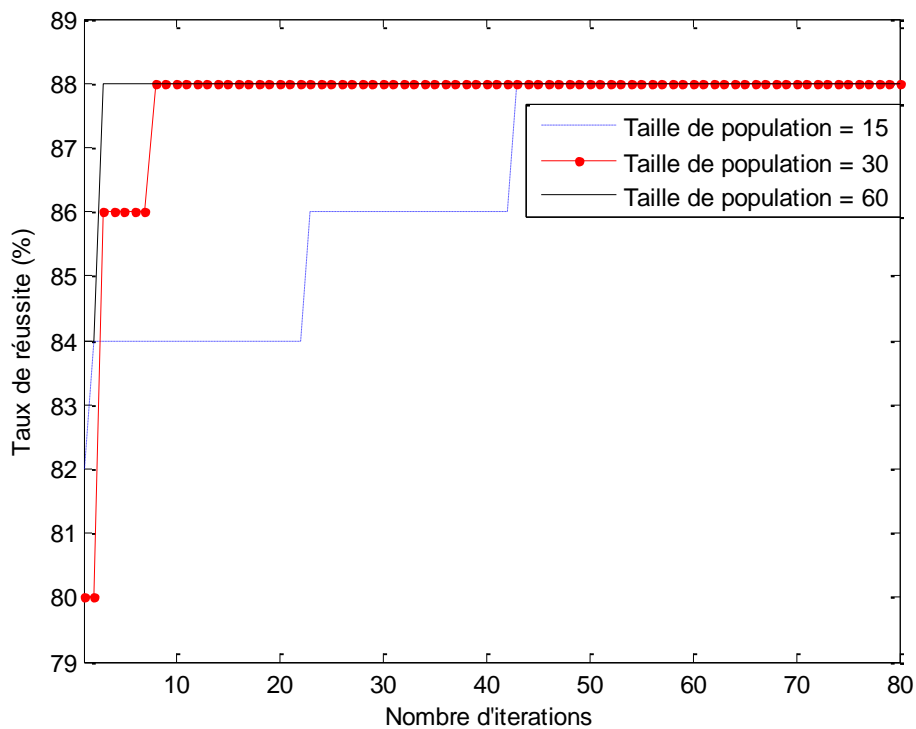


Figure IV.4. Performance du classifieur en fonction du nombre d'itérations

Dans le cas de l'algorithme Pentagone de Duval-KNN, plusieurs types de distance ont été utilisés à savoir "Euclidienne", "Cityblock" et "Cosine". Par ailleurs, le nombre de voisins k a été varié de 1 à 105. La valeur maintenue est celle permettant d'obtenir un meilleur taux de réussite. La figure IV.5 montre la performance du classifieur Pentagone de Duval-KNN. Un taux de 82% a été obtenu aussi bien pour la distance "Euclidienne" avec un nombre de voisins $k = 7$ ou 10, que pour celle "Cityblock" avec $k = 10$, contre 70% pour $k = 20$ obtenu lors de l'utilisation de la distance "Cosine". Dans notre classification utilisant l'algorithme Pentagone de Duval-KNN, nous avons choisi la distance "Euclidienne" avec $k = 7$.

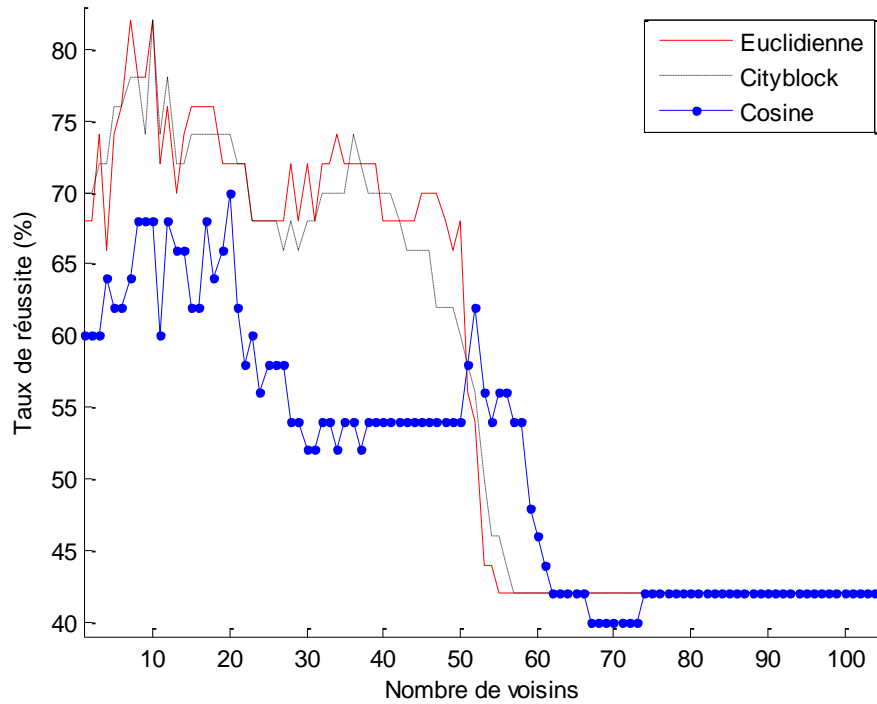


Figure IV.5. Impact de type de la distance et du nombre de voisins sur la performance du classifieur Pentagone de Duval-KNN

Les résultats de la classification présentés dans le tableau IV.1 sont relatifs à l'utilisation du noyau "polynomial" (d'ordre 1) et de la stratégie d'apprentissage "un contre un" pour l'algorithme Pentagone de Duval-SVM-PSO d'une part, et la distance "Euclidienne" avec $k=7$ pour l'algorithme Pentagone de Duval-KNN d'autre part. De tels choix ont permis l'obtention de meilleures précisions de classification. Nous illustrons les résultats contenant 10 échantillons dans le tableau IV.1. Les résultats relatifs aux 50 échantillons testés sont présentés en Annexe 2.

Tableau IV.1. Résultats de diagnostic sur 10 échantillons testés

Cas	H ₂ (ppm)	CH ₄ (ppm)	C ₂ H ₂ (ppm)	C ₂ H ₄ (ppm)	C ₂ H ₆ (ppm)	Inspection	Pentagone de Duval	Pentagone de Duval -SVM- PSO	Pentagone de Duval - KNN
1	16000	3600	0	14	670	PD	S	PD	PD
2	6600	1000	19	2	38	PD	PD	PD	PD
3	60	10	4	4	4	D1	S	D1	D1
4	78	20	28	13	11	D1	D1	D2	D2
5	7150	14440	1760	2110	97	D2	D1	D2	D2
6	755	229	460	404	32	D2	D2	D2	D2
7	3420	7870	33	6990	15000	T1&T2	T3	T3	T3

8	3675	6392	5	7691	2500	T1&T2	T1&T2	T&T2	T3
9	150	22	11	60	9	T3	D1	D2	D2
10	2500	10500	6	13500	4790	T3	T3	T3	T3

De ce tableau, l'algorithme Pentagone de Duval-SVM-PSO fournit 7 types de défaut (correspondant aux cas 1, 2, 3, 5, 6, 8 et 10) en parfait accord avec les inspections de la CEI TC 10 [53] et de [54], contre 6 (correspondant aux cas 1, 2, 3, 5, 6 et 10) pour l'algorithme Pentagone de Duval-KNN. Cependant, le pentagone (méthode graphique) de Duval présente uniquement 5 types de défauts (correspondant aux cas 2, 4, 6, 8 et 10) en accord avec les inspections susmentionnées.

Sur l'ensemble des 50 cas (tableau présenté en Annexe 2), les types de défauts similaires aux inspections sont 44, 41 et 40 obtenus lors de l'application de l'algorithme Pentagone de Duval-SVM-PSO, Pentagone de Duval-KNN et Pentagone de Duval respectivement. Les précisions de diagnostic obtenues en appliquant la méthode graphique (Pentagone de Duval) et les algorithmes hybrides sont montrés dans le tableau IV.2. Les résultats indiquent que la combinaison Pentagone de Duval avec la SVM-PSO permet d'améliorer la méthode du Pentagone de Duval de 8%, contre 2% lors de l'application de l'algorithme Pentagone de Duval-KNN. Ainsi, il est fortement recommandé d'utiliser SVM-PSO pour une meilleure amélioration du pentagone de Duval.

Tableau IV.2. Comparaison des taux de réussite des différents algorithmes

	Pentagone de Duval	Pentagone de Duval-KNN	Pentagone de Duval-SVM-PSO
Taux de réussite (%)	80	82	88

IV.6 Influence du type du vecteur d'entrée

Nous commençons d'abord par une représentation graphique qui explique la distribution de la base de données. La technique des boîtes à moustaches ou bien le diagramme en boîte de Tukey (en anglais boxplot) a été introduire par John Tukey [92]. Chaque boîte contient une médiane, un quartile, un minimum et un maximum. La figure IV.6 illustre la représentation en boîte à moustache.

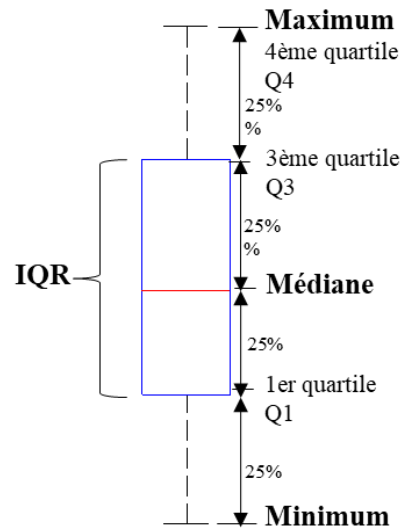
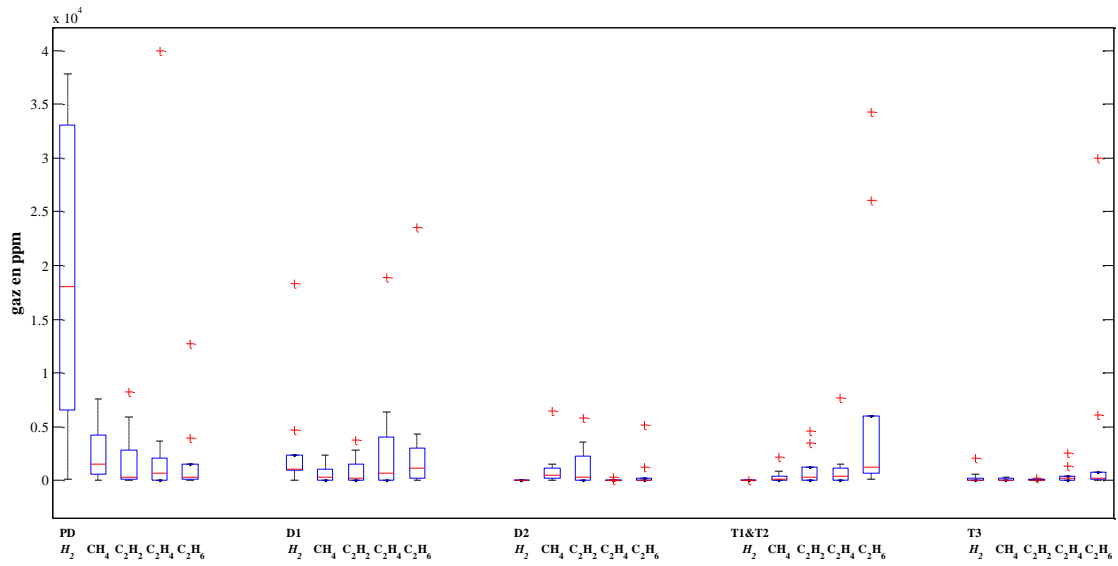


Figure IV.6. Représentation d'une boîte à moustaches

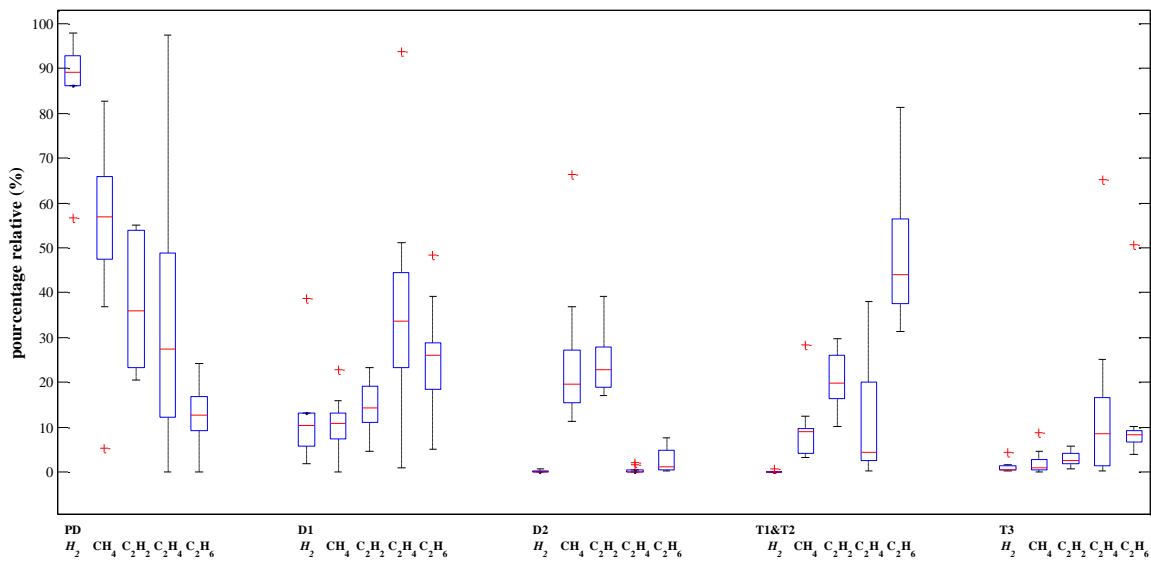
50 % de la base de données est inclus dans la boîte. La partie correspondante est connue sous le nom d'écart interquartile (IQR). La médiane des valeurs est représentée sous forme de ligne qui fractionne la boîte en deux. La partie de IQR est très importante, car si le IQR petit indique que la plupart des valeurs se trouvent proches du centre. Si IQR est très grand, les valeurs qui sont non uniformément éparpillées dans la boîte. Deux moustaches se présentent. La première est située entre le minimum et le premier quartile. Quant à la deuxième, elle est localisée entre le troisième quartile et le maximum. La différence entre le minimum et le maximum correspond à l'étendue.

Les algorithmes élaborés ont été testés pour cinq types de vecteur d'entrée. Il s'agit des gaz en ppm (H_2 , CH_4 , C_2H_2 , C_2H_4 et C_2H_6), des gaz en pourcentage ($\%H_2$, $\%CH_4$, $\%C_2H_2$, $\%C_2H_4$ et $\%C_2H_6$), des ratios de Rogers (C_2H_2/C_2H_4 , CH_4/H_2 et C_2H_4/C_2H_6), des ratios de Dornenburg (CH_4/H_2 , C_2H_2/CH_4 , C_2H_4/C_2H_6 et C_2H_2/C_2H_4) et finalement des ratios du triangle de Duval ($\%CH_4$, $\%C_2H_2$ et $\%C_2H_4$).

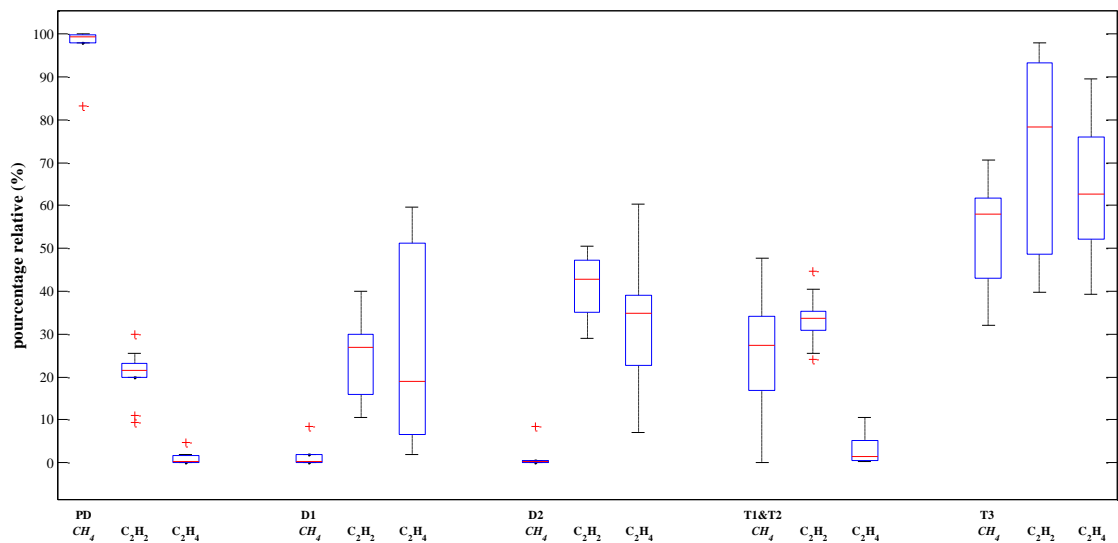
La figure IV.7 illustre la distribution de la base de données de la phase d'apprentissage pour chaque type de vecteur d'entrée utilisant la boîte à moustaches. Cette distribution varie en fonction du type de vecteur d'entrée. Une analyse des résultats de la figure IV.7 s'avère indispensable, parce que nous remarquerons dans la suite que cette distribution a un effet sur le taux de réussite.



(a)



(b)



(c)

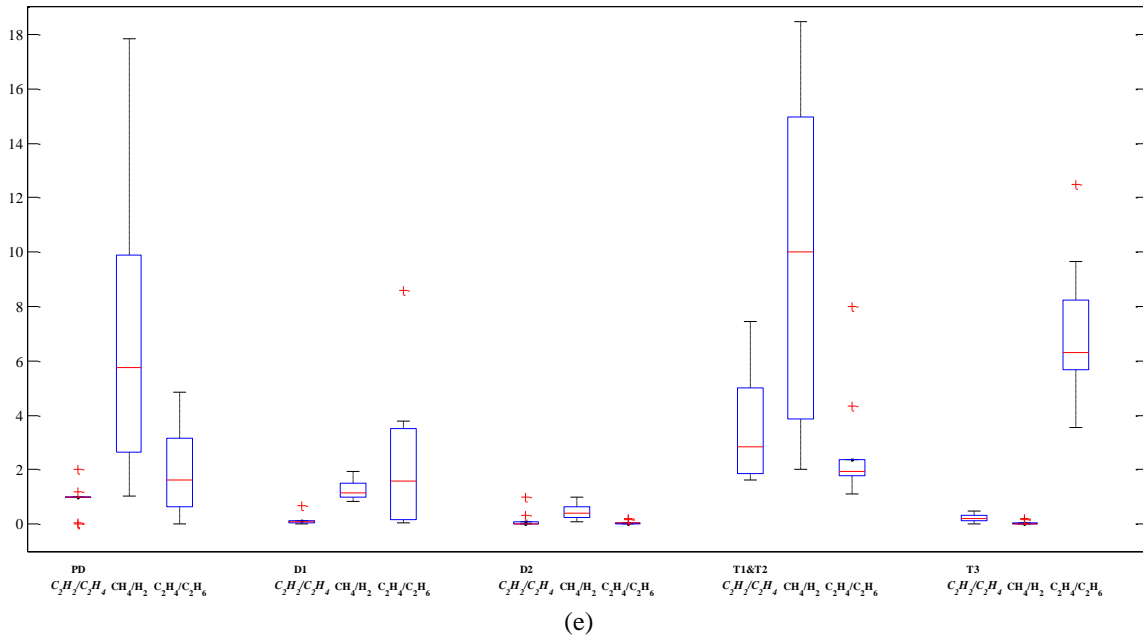


Figure IV.7. Boîte à moustaches (a) gaz en ppm, (b) gaz en pourcentage, (c) ratios de Triangle de Duval et (e) ratios de Rogers

La figure IV.8 illustre la distribution pour les ratios de Dornenburg utilisant la boîte à moustaches. Le rapport R_1 (CH_4/H_2) de défaut T1&T2 a une étendue (différences entre les valeurs maximales et minimales) importante, les boîtes à moustache des défauts DP, D1, D2 et T3 deviennent invisibles. De ce fait, la représentation en boîte à moustaches de chaque défaut a été tracée à part (figure IV.9). Rappelons que : $R_1 = CH_4/H_2$, $R_2 = C_2H_2/C_2H_4$, $R_3 = C_2H_2/CH_4$ et $R_4 = C_2H_6/C_2H_4$,

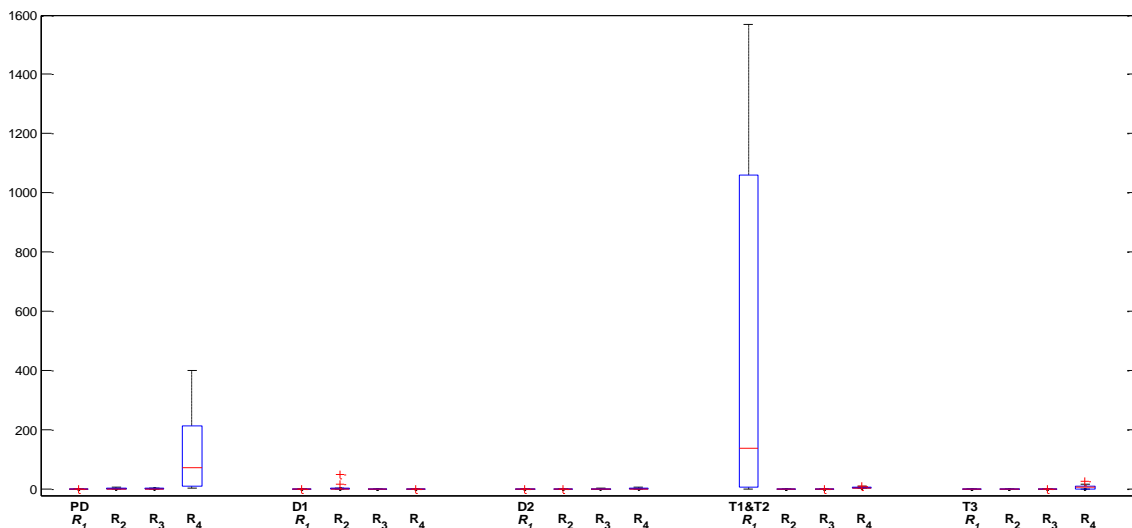


Figure IV.8. Boîte à moustaches des ratios de Dornenburg (pour tous les défauts)

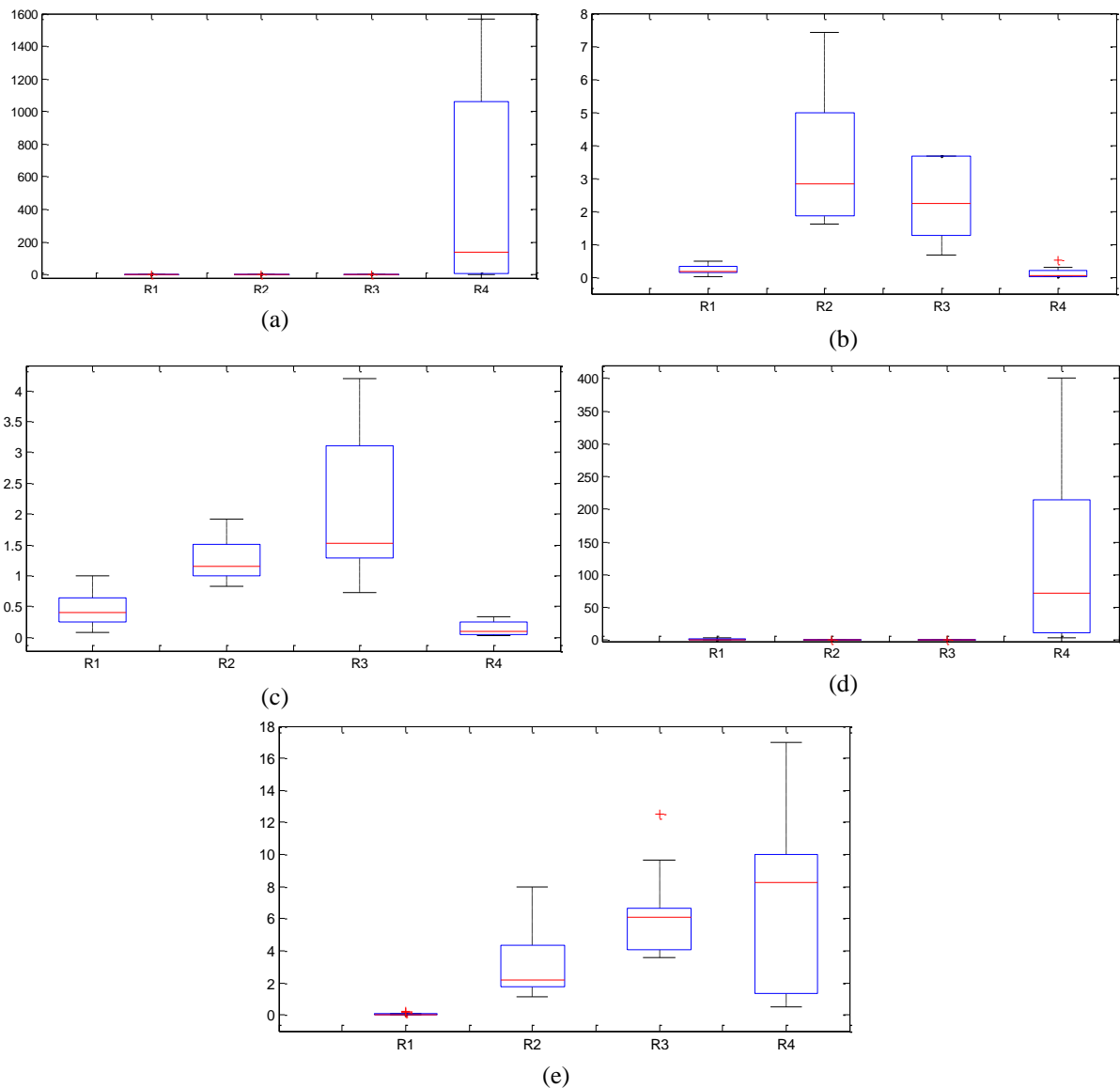


Figure IV.9. Boîte à moustaches des ratios de Dornenburg par défaut : (a) défaut DP, (b) défaut D1, (c) défaut D2, (d) défaut T1&T2 et (e) défaut T3

IV.6.1 Algorithme KNN

Dans le but d'atteindre le maximum en terme du taux de réussite de l'ensemble des défauts (ou global), le nombre de voisins a été varié de 1 à 105. Les résultats ainsi obtenus pour les différents vecteurs d'entrée et type de distance sont présentés sur la figure IV.9.

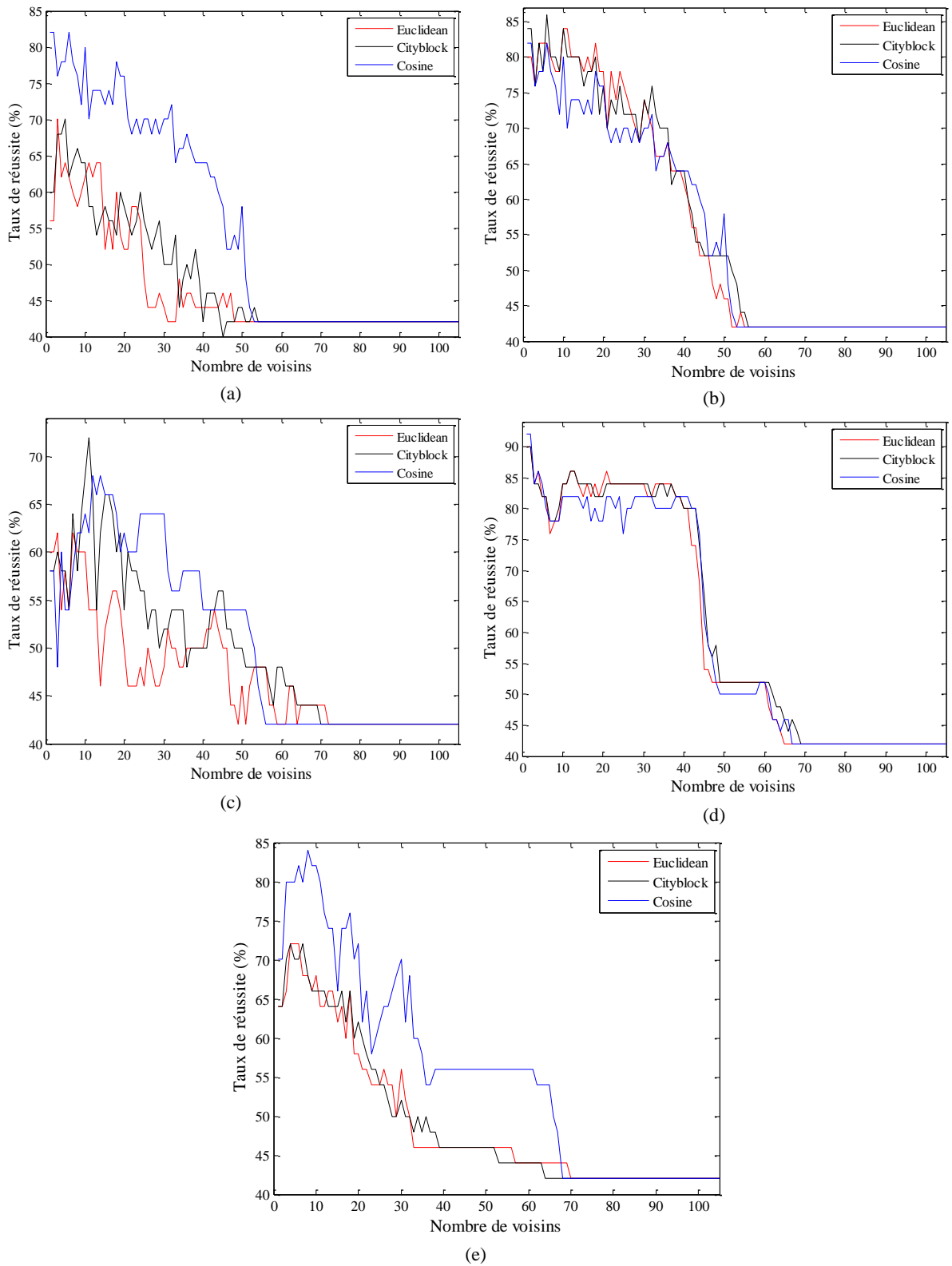


Figure IV.9. Impact du nombre de voisins et le type de vecteur d'entrée sur la classification : (a) gaz en ppm, (b) gaz en pourcentage, (c) ratios de triangle de Duval, (d) ratios de Rogers, (e) ratios de Dornenburg

Le meilleur taux de réussite de 92% a été obtenu lors de l'introduction des rapports du triangle de Duval comme vecteur d'entrée avec le type de distance " Cosine" (nombre de voisins étant K=1 ou 2). Ce résultat est suivi par celui présentant un taux de réussite de 86% (avec

K=6). Ce dernier a été obtenu en injectant les gaz en pourcentage utilisant la distance "Cityblock". En troisième position, viennent les ratios de Dornenburg avec un taux de réussite de 84% (obtenu pour K=10) en employant la distance "Cosine". Le vecteur d'entrée utilisant les gaz en ppm donne un taux de réussite de 82% (trouvé pour K=1, 2 ou 6) avec la distance "Cosine". Finalement, un taux de réussite de 70% (avec K=11) a été obtenu lors de l'injection des ratios de Rogers avec l'utilisation de la distance "Cityblock".

Tableau IV.3. Taux de réussite pour chaque vecteur d'entrée du classifieur KNN

	Gaz en ppm	Gaz en pourcentage	Ratios de Rogers	Ratios de Duval	Ratios de Dornenburg
KNN	82%	86%	72%	92%	84%
Nombres de voisins	1, 2, 6	6	11	1, 2	8
Type de distance	Cosine	Cityblock	Cityblock	Cosine	Cosine

La figure IV.10 illustre le taux de réussite du classifieur KNN par type de défaut pour les cinq types de vecteur d'entrées. Cette représentation permet d'avoir une meilleure vision sur les résultats obtenus à partir de chaque type de vecteur d'entrées et de prendre les décisions adéquates pour parer aux défaillances engendrées par chaque défaut.

Comme nous pouvons le constater, il est favorable d'utiliser le vecteur des ratios du triangle de Duval pour les défauts DP, D1, T1&T2 et T3, et d'utiliser le vecteur d'entrée de Rogers ou bien les gaz en pourcentage pour le défaut D2.

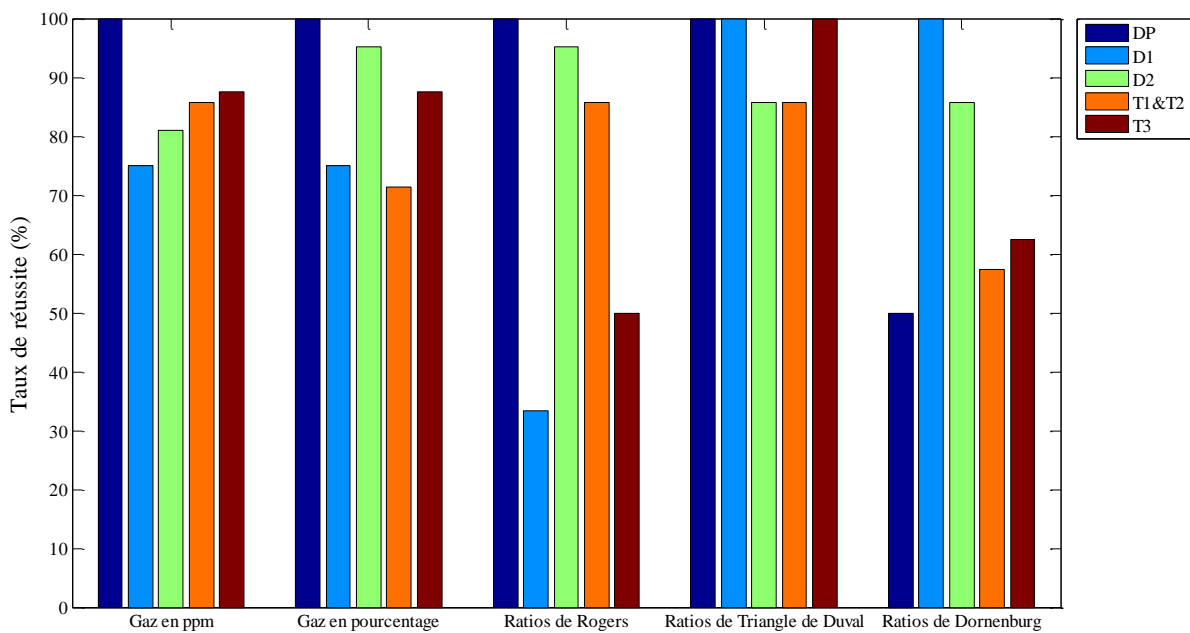


Figure IV.10. Taux de réussite par type de défaut du classifieur KNN

IV.6.2 Algorithme naïve Bayes

Le classifieur naïve Bayes a été appliqué pour tous les types du vecteur d'entrée précédemment cités. Il convient de calculer, pour chaque type, l'écart type et la moyenne de chaque défaut conformément aux expressions suivantes :

$$\mu = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 \quad (4.4)$$

$$m = \frac{1}{n} \sum_{i=1}^n x_i \quad (4.5)$$

μ et m sont des matrices dans lesquelles le nombre de lignes est équivalent au nombre de défauts (5, dans notre cas), et le nombre de colonnes correspond au nombre d'élément du vecteur d'entrée. A titre d'exemple, pour le premier vecteur d'entrée (les gaz en ppm) $V_{in} = [H_2 \ CH_4 \ C_2H_2 \ C_2H_4 \ C_2H_6]$, les matrices μ et m relatives à l'écart type et à la moyenne peuvent se mettre sous la forme :

$$\begin{pmatrix} \mu_{PD} \\ \mu_{D1} \\ \mu_{D2} \\ \mu_{T1\&T2} \\ \mu_{T3} \end{pmatrix} = \begin{pmatrix} \mu_{PDH_2} & \mu_{PDCH_4} & \mu_{PDC_2H_2} & \mu_{PDC_2H_4} & \mu_{PDC_2H_6} \\ \mu_{D1H_2} & \mu_{D1CH_4} & \mu_{D1C_2H_2} & \mu_{D1C_2H_4} & \mu_{D1C_2H_6} \\ \mu_{D2H_2} & \mu_{D2CH_4} & \mu_{D2C_2H_2} & \mu_{D2C_2H_4} & \mu_{D2C_2H_6} \\ \mu_{T\&T2H_2} & \mu_{T1\&T2CH_4} & \mu_{T1\&T2C_2H_2} & \mu_{T1\&T2C_2H_4} & \mu_{T1\&T2C_2H_6} \\ \mu_{T3H_2} & \mu_{T3CH_4} & \mu_{T3C_2H_2} & \mu_{T3C_2H_4} & \mu_{T3C_2H_6} \end{pmatrix} \quad (4.6)$$

$$\begin{pmatrix} m_{PD} \\ m_{D1} \\ m_{D2} \\ m_{T1\&T2} \\ m_{T3} \end{pmatrix} = \begin{pmatrix} m_{PDH_2} & m_{PDCH_4} & m_{PDC_2H_2} & m_{PDC_2H_4} & m_{PDC_2H_6} \\ m_{D1H_2} & m_{D1CH_4} & m_{D1C_2H_2} & m_{D1C_2H_4} & m_{D1C_2H_6} \\ m_{D2H_2} & m_{D2CH_4} & m_{D2C_2H_2} & m_{D2C_2H_4} & m_{D2C_2H_6} \\ m_{T\&T2H_2} & m_{T1\&T2CH_4} & m_{T1\&T2C_2H_2} & m_{T1\&T2C_2H_4} & m_{T1\&T2C_2H_6} \\ m_{T3H_2} & m_{T3CH_4} & m_{T3C_2H_2} & m_{T3C_2H_4} & m_{T3C_2H_6} \end{pmatrix} \quad (4.7)$$

Pour les rations du triangle de Duval, les matrices μ et m deviennent :

$$\begin{pmatrix} \mu_{PD} \\ \mu_{D1} \\ \mu_{D2} \\ \mu_{T1\&T2} \\ \mu_{T3} \end{pmatrix} = \begin{pmatrix} \mu_{PD_%CH_4} & \mu_{PD_%C_2H_2} & \mu_{PD_%C_2H_4} \\ \mu_{D1_%CH_4} & \mu_{D1_%C_2H_2} & \mu_{D1_%C_2H_4} \\ \mu_{D2_%CH_4} & \mu_{D2_%C_2H_2} & \mu_{D2_%C_2H_4} \\ \mu_{T1\&T2_%CH_4} & \mu_{T1\&T2_%C_2H_2} & \mu_{T1\&T2_%C_2H_4} \\ \mu_{T3_%CH_4} & \mu_{T3_%C_2H_2} & \mu_{T3_%C_2H_4} \end{pmatrix} \quad (4.8)$$

$$\begin{pmatrix} m_{PD} \\ m_{D1} \\ m_{D2} \\ m_{T1\&T2} \\ m_{T3} \end{pmatrix} = \begin{pmatrix} m_{PD_%CH_4} & m_{PD_%C_2H_2} & m_{PD_%C_2H_4} \\ m_{D1_%CH_4} & m_{D1_%C_2H_2} & m_{D1_%C_2H_4} \\ m_{D2_%CH_4} & m_{D2_%C_2H_2} & m_{D2_%C_2H_4} \\ m_{T1\&T2_%CH_4} & m_{T1\&T2_%C_2H_2} & m_{T1\&T2_%C_2H_4} \\ m_{T3_%CH_4} & m_{T3_%C_2H_2} & m_{T3_%C_2H_4} \end{pmatrix} \quad (4.9)$$

La fonction Gaussienne a été utilisée pour calculer la probabilité conditionnelle $P(x|w_k)$. La probabilité postérieure est calculée en employant l'équation suivante :

$$P(x_i|w_k) = \frac{1}{\sqrt{2\pi}\mu_k} e^{-\frac{1}{2} \frac{(x_i - m_k)^2}{\mu_k^2}} \quad (4.10)$$

La sortie (contenant 5 valeurs) du classifieur est comme suit : $[P(x|w_1) P(x|w_2) P(x|w_3) P(x|w_4) P(x|w_5)]$, avec $[w_1 w_2 w_3 w_4 w_5] = [DP D1 D2 T1\&T2 T3]$.

Pour identifier le type de défaut (la classe) de l'échantillon x_i , nous avons appliqué la règle du maximum à posteriori ($\max P(x_i|w_k)$).

Le taux de réussite a été calculé pour 50 échantillons de test. Les résultats ainsi obtenus sont présentés en Annexe 3.

A titre d'exemple nous présentons une application de l'analyse des gaz dissous dans trois échantillons d'huile différents (Tableau IV.4).

Tableau IV.4. AGD de trois échantillons

Echantillon	H ₂ (ppm)	CH ₄ (ppm)	C ₂ H ₂ (ppm)	C ₂ H ₄ (ppm)	C ₂ H ₆ (ppm)
1	1820	405	634	365	35
2	35	6	482	26	3
3	6790	10500	750	17700	1400

Nous donnons pour exemple les probabilités obtenues pour le cas de l'échantillon 1. Nous avons $P(w_1|x) = P(w_5|x) = 0$, $P(w_2|x) = 0,318$, $P(w_3|x) = 0,680$ et $P(w_4|x) = 0,002$. Ainsi, $P(w_3|x) > P(w_2|x) > P(w_4|x) > P(w_1|x) > P(w_5|x)$, donc le défaut correspondant de la classe w_3 , d'une part est le défaut D2.

Rappelons que l'équation 4.11 reste toujours valable :

$$\sum_{k=1}^5 P(w_k|x_i) = 1 \quad (4.11)$$

Les probabilités d'occurrence des défauts pour les 3 cas d'échantillon sont données dans le tableau IV.5.

Tableau IV.5. Sortie de classifieur naïve Bayes

Cas	$(P(w_k x))$	Défaut estimé
1	[0 0.318 0.680 0.002 0]	D2
2	[0 1 0 0 0]	D1
3	[0 0 0 0.001 0.999]	T3

La procédure précédemment présentée nous a permis de calculer le taux de réussite de chaque type de vecteur d'entrée. Les résultats sont dans le tableau IV.6. Le meilleur taux de réussite de 86% a été obtenu lors l'injection des gaz en pourcentage comme vecteur d'entrée. En seconde position viennent les ratios du triangle de Duval avec un taux de réussite de 84%, suivie par celle des ratios de Dornenburg en ppm avec un taux de 74%.

Tableau IV.6. Taux de réussite pour chaque vecteur d'entrée du classifieur naïve de Bayes

	Gaz en ppm	Gaz en pourcentage	Ratios de Rogers	Ratios de Duval	Ratios de Dornenburg
Classifieur naïve Bayes	32%	86%	66%	84%	74%

La figure IV.11 illustre le taux de réussite du classifieur naïve Bayes par type de défaut pour les cinq types de vecteur d'entrée. Nous constatons, qu'il est favorable d'employer le vecteur des ratios du triangle de Duval pour les défauts DP et T3, et celle du vecteur d'entrée des ratios de Dornenburg pour le défaut D1. Pour le défaut D2, à part le vecteur des gaz en ppm, les autres vecteurs donnent le même taux de réussite. Concernant le défaut T1&T2, il est préférable d'utiliser le vecteur d'entrée des gaz en pourcentage.

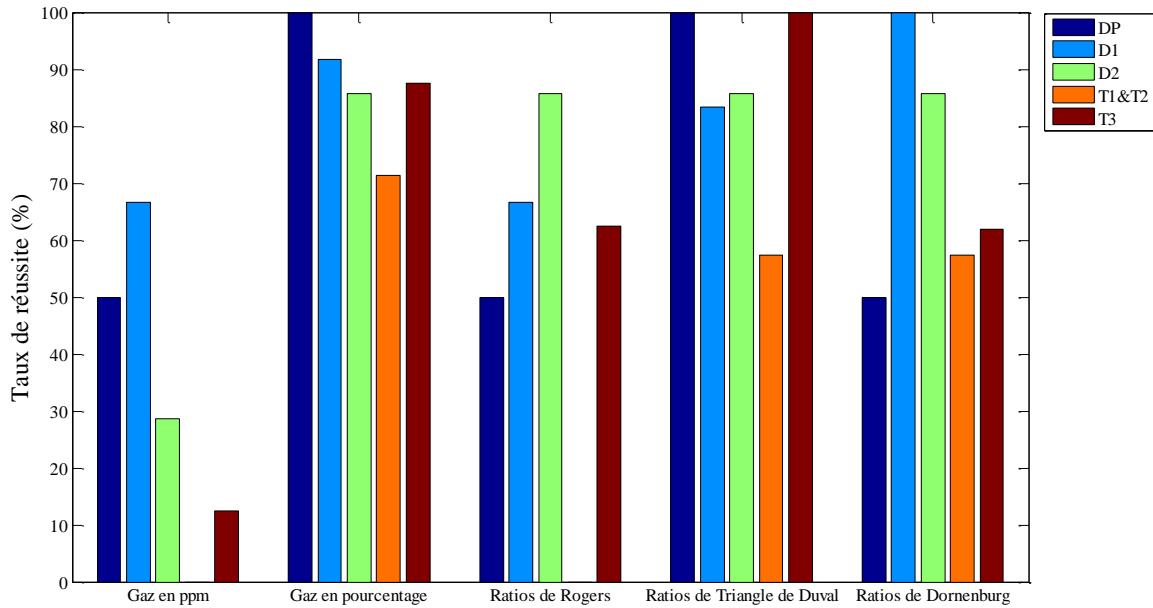


Figure IV.11. Taux de réussite par type de défaut du classifieur naïve Bayes

IV.6.3 Algorithme SVM-PSO

L'optimisation des paramètres de marge et de conditionnement (C, λ) a été effectuée par le biais de l'algorithme PSO. L'influence du degré du polynôme est prise en considération. La figure IV.12 présente l'évolution de la performance du classifieur SVM-PSO en fonction du nombre d'itérations. Les paramètres de l'algorithme SVM ajustés par PSO obtenus pour chaque type de vecteur d'entrée sont présentés dans le tableau IV.7.

Tableau. 4.7. Paramètres ajustés donnant le meilleur taux de réussite pour chaque vecteur d'entrée

Vecteur d'entrée	Paramètres optimisés	Degré de polynôme
Gaz en ppm	$(C, \lambda) = (9.0475, 0.0325)$	1
Gaz en pourcentage	$(C, \lambda) = (0.3241, 0.0102)$	1
Ratios de Rogers	$(C, \lambda) = (34.2007, 0.0267)$	1
Ratios de Duval	$(C, \lambda) = (4.1630, 0.194)$	2
Ratios de Dornenburg	$(C, \lambda) = (5.1824, 0.01314)$	1

De même que précédemment, l'injection des rapports du triangle de Duval présente le taux le plus élevé de réussite égal à 98%. Ce taux a été trouvé en utilisant la fonction noyau "polynomiale d'ordre 2" et la stratégie d'apprentissage "un contre un". Pour les autres vecteurs d'entrées, le meilleur taux de réussite le degré du polynôme du fonction de noyau est d'ordre 1, avec la stratégie d'apprentissage "un contre un".

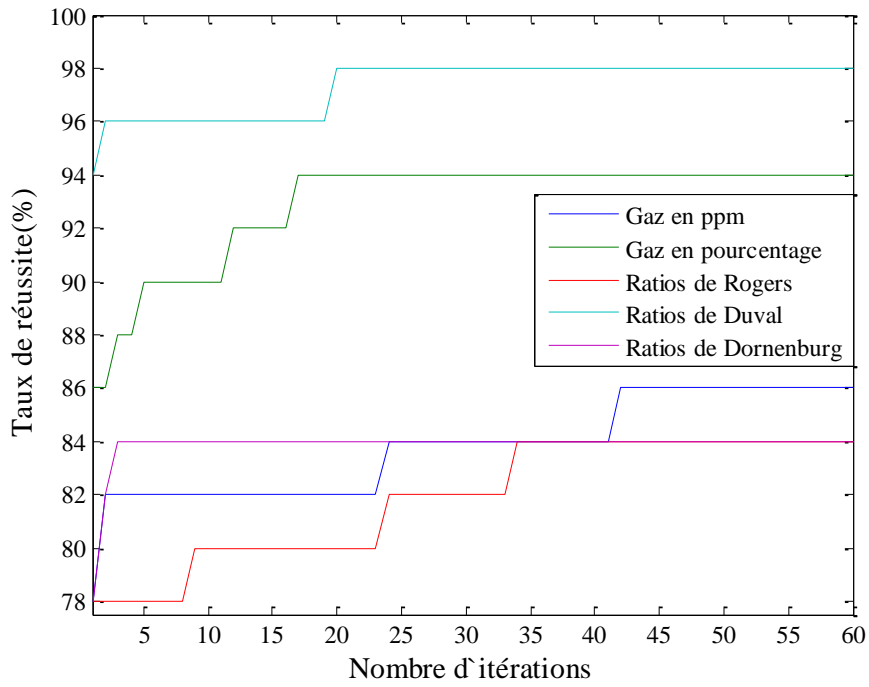


Figure IV.12. Performance du SVM-PSO en fonction du nombre d'itérations pour différents types du vecteur d'entrée

La figure IV.13 illustre le taux de réussite par type de défaut du classifieur SVM-PSO pour les cinq types de vecteur d'entrée. De ce fait, il est favorable d'utiliser le classifieur PSO-SVM avec les ratios du triangle de Duval pour les défauts DP, D1, D2, T1&T2 et T3.

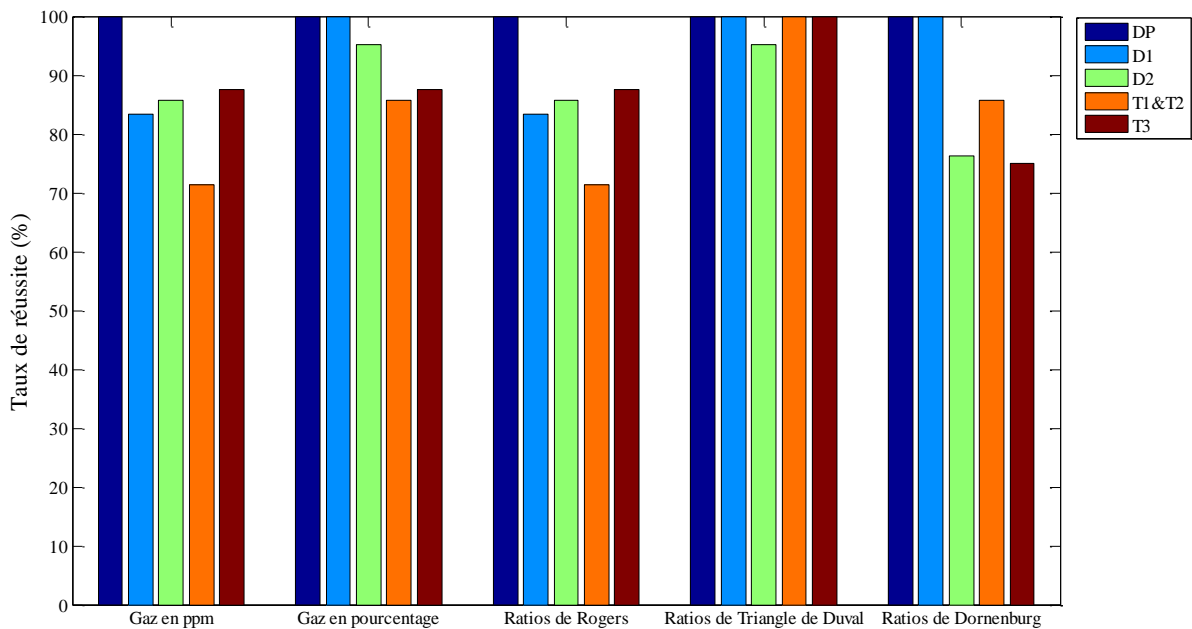


Figure IV.13. Taux de réussite par type de défaut du classifieur SVM-PSO

Le tableau IV.8 récapitule les différents résultats ainsi obtenus lors de l'application des algorithmes SVM-PSO, KNN et naïve Bayes pour différents types de vecteur d'entrée.

Tableau IV.8. Récapitulatif des résultats

	Gaz en ppm	Gaz en pourcentage	Ratios de Rogers	Ratios de Duval	Ratios de Dornenburg
PSO-SVM	86%	94%	84%	98%	84%
KNN	82%	86%	72%	92%	84%
naïve Bayes	32%	86%	66%	84%	74%

D'après le tableau IV.8, l'algorithme SVM-PSO offre le meilleur taux de réussite quant au diagnostic de l'huile de transformateur de puissance. En outre, SVM a l'aptitude de traiter les problèmes aussi bien linéaires, que non linéaires. Par ailleurs, l'algorithme KNN, qui nécessite l'optimisation du nombre de voisins, présente de très bons résultats également. Quant à l'algorithme naïve Bayes, il donne des résultats acceptables.

IV.7 Conclusion

Les algorithmes hybrides à savoir Pentagone de Duval-SVM-PSO et Pentagone de Duval-KNN ont été développés pour améliorer le taux de classification du pentagone de Duval (estimé à 80%). Les différents résultats obtenus sont :

- Lors de l'élaboration de l'algorithme Pentagone de Duval-KNN, trois types de distance à savoir "Euclidienne", "Cityblock" et "Cosine" ont été testées. La distance "Euclidienne" a donné le meilleur taux de réussite de 82% avec un nombre de voisins $K=7$ ou 10, soit une amélioration du pentagone de Duval de 2%.
- La fonction « noyau » linéaire et les stratégies d'apprentissage "tous ensemble", "un contre ensemble" et "un contre un" ont été testées par l'algorithme pentagone de Duval-SVM-PSO. Les paramètres (C, λ) ont été optimisés en utilisant la technique d'optimisation par essaim particulière (PSO). L'algorithme hybride Pentagone de Duval-SVM offre un taux de réussite de 88%, soit une amélioration de 8%. Ce scénario a été obtenu en utilisant la fonction « noyau » polynomial (d'ordre 1), la stratégie d'apprentissage "un contre un" et le couple $(C, \lambda) = (27,922021, 0,00047)$.

Les performances des algorithmes SVM-PSO, KNN et naïve Bayes ont été finalement testés en variant le type du vecteur d'entrée. Pour cela, cinq types ont été adoptés à savoir, gaz en ppm, gaz en pourcentage, ratios de Rogers, ratios de Dornenburg et ratios du triangle de Duval.

Un taux de réussite de 98% a été obtenu lors de l'utilisation de l'algorithme PSO-SVM en utilisant une fonction « noyau » polynomial d'ordre 2, la stratégie d'apprentissage "un contre un", $C= 4,1630$ et $\lambda=0,194$, et en introduisant les ratios du triangle de Duval comme vecteur d'entrée. Dans le cas de l'algorithme KNN, nous avons obtenu un taux de réussite de 92%. Ce scénario a été obtenu en utilisant une distance " Cosine " et un nombre de voisins $K=1$ ou 2 . En dernier lieu vient l'algorithme naïve Bayes avec un taux de réussite de 86% lors l'injection des concentrations de gaz en pourcentage comme vecteur d'entrée.



Conclusion Générale

Notre travail a consisté à l'application de certaines méthodes intelligentes à savoir SVM-PSO, KNN et naïve Bayes pour le diagnostic de l'huile de transformateur par l'analyse des gaz dissous. Les travaux de recherche que nous avons effectués ont deux principaux objectifs.

Le premier objectif a été d'améliorer le taux de réussite (de précision) du pentagone de Duval en se basant sur les pourcentages des cinq principaux gaz : %H₂, %CH₄, % C₂H₂, %C₂H₄ et %C₂H₆. Inspirés de la CEI et d'IEEE, nous avons considéré les cinq défauts suivants : DP, D1, D2, T1&T2, et T3. Les bases de données que nous avons utilisées sont celles de CEI TC 10 et celle de Duval contenant ensemble 155 échantillons dont 105 (68% de données) ont été réservés pour l'apprentissage et 50 (32% de données) pour le test.

Nous avons établi un algorithme principal et utilisé deux classifieurs, obtenant ainsi deux algorithmes combinés qui nous ont permis de reproduire exactement le diagnostic du pentagone de Duval. Les deux classifieurs que nous avons élaborés sont le SVM-PSO et le KNN. Les centres de masse, distribués de manière non-linéaire, ont été introduits comme entrées.

L'algorithme pentagone de Duval-SVM-PSO a été testé pour plusieurs fonctions noyau et stratégies d'apprentissage. Les paramètres de la marge et de conditionnement de l'algorithme SVM ont été optimisés par la technique d'optimisation par essaim particulière PSO. De même, l'algorithme KNN a été testé pour divers types de distances et nombres de voisins.

Nous avons obtenu une amélioration du taux de réussite de 8% en utilisant l'algorithme SVM-PSO et une amélioration de 2% du taux de réussite dans le cas de l'algorithme KNN.

Le deuxième objectif a consisté à étudier l'influence du type de vecteur d'entrée sur les performances des algorithmes SVM-PSO, KNN et naïve Bayes. Les cinq types de vecteurs qui ont été adoptés sont : les concentrations de gaz en ppm, celles des gaz en pourcentage, les ratios de Rogers, les ratios de Dornenburg et les ratios du triangle de Duval. Une étude comparative a été effectuée et il en ressort les principales conclusions suivantes :

- Le choix du nombre de voisins et du type de distance influent sur le classifieur KNN, quel que soit le type de vecteur d'entrée. Nous avons obtenu un taux de réussite de 92% en utilisant les ratios du triangle de Duval comme vecteur d'entrée. Ce taux coïncide avec la distance "Cosine" et le nombre de voisins $K = 1$ ou 2 .
- Le type et les paramètres de la fonction noyau et la stratégie d'apprentissage ont une grande influence sur le classifieur SVM-PSO dont les paramètres de marge et de conditionnement ont été sélectionnés par l'algorithme d'optimisation PSO. En utilisant le noyau polynomial d'ordre 2 et la stratégie d'apprentissage « un contre un » pour le vecteur d'entrée des ratios du triangle de Duval, nous avons obtenu un taux de réussite de 98%.

- La simplicité du classifieur naïve Bayes ne nécessite aucun ajustement. Il faut juste choisir le vecteur d'entrée le plus compatible avec ce type de classifieur. Un taux de réussite de 86% a été obtenu avec le vecteur d'entrée des concentrations de gaz en pourcentage.

Comme perspective de nos travaux de recherche, nous pouvons citer les études suivantes :

- Appliquer les algorithmes élaborés avec des bases de données plus importantes.
- Utiliser d'autres algorithmes d'intelligence artificielle (réseaux de neurones, logique floue, essaim d'abeille, colonies de fourmis, etc.) dans les méthodes conventionnelles que nous avons utilisées.
- Développer des méthodes de monitoring basées sur l'intelligence artificielle.



Bibliographie

-
- [1] CIGRE Technical Brochure #296, “Recent Developments in DGA Interpretation,” 2006.
- [2] Yishan Liang, Zhenyuan Wang and Yilu Liu, "Power transformer DGA data processing and alarming tool for on-line monitors," *2009 IEEE/PES Power Systems Conference and Exposition*, Seattle, WA, 2009, pp. 1-8.
- [3] M. Duval, “Dissolved Gas Analysis: It Can Save Your Transformer”, *IEEE Elec. Insul. Mag.*, Vol. 5, No. 6, pp. 22-27, 1989.
- [4] N. Abu Bakar, A. Abu-Saida and S. Islam, “A Review of Dissolved Gas Analysis Measurement and Interpretation Techniques “, *IEEE Elec. Insul. Mag.*, vol. 30, No. 3, pp. 39-49, 2014.
- [5] Mineral Oil-Impregnated Electrical Equipment in Service—Guide to the Interpretation of Dissolved and Free Gases Analysis, IEC Publication 60599, 2007.
- [6] IEEE Guide for the Detection and Determination of Generated Gases in Oil-Immersed Transformers and Their Relation to the Serviceability of the Equipment, ANSI/IEEE Std C57.104, 1978.
- [7] Rogers, R.R.: “IEEE and IEC codes to interpret incipient faults in transformers using gas in oil analysis “, *IEEE Trans. Electr. Insul.*, Vol. 13, No.5, pp.348-354, 1978.
- [8] E. Dornenburg and W. Strittmatter, "Monitoring oil-cooled transformers by gas analysis" *Brown Boveri Review*1974, 61:238–247.
- [9] M. Duval and L. Lamarre, “The Duval Pentagon- Anew Complementary Tool for the Interpretation of Dissolved Gas Analysis in Transformers”, *IEEE Elec. Insul. Mag.*, Vol. 30, No. 6, pp. 9-12, 2014.
- [10] Mineral Oil-Impregnated Electrical Equipment in Service—Guide to the Interpretation of Dissolved and Free Gases Analysis, IEC Publication 60599, 2015.
- [11] W.H. Tang, J. Y. Goulermas, Q.H. Wu, Z. J. Richardson, and J. Fitch,“A Probabilistic Classifier for Transformer Dissolved Gas Analysis With a Particle Swarm Optimizer”, *IEEE Trans. Power Del.*, Vol.23, No. 2, pp. 751–759, April. 2008.
- [12] Q. Su, C. Mi, L.L. Lai and P. Austin, “A Fuzzy Dissolved Gas Analysis Method for the Diagnosis of Multiple Incipient Faults in a Transformer”, *IEEE Trans. Power Del.*, vol. 15, No. 2, pp. 593–798, 2000.
- [13] W. H. Tang, Z. Lu and Q. H. Wu, "A Bayesian network approach to power system asset management for transformer dissolved gas analysis", *Conf on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, China 2008, pp. 1460-1466.
- [14] A. Shintemirov, W.H. Tang and Q.H. Wu, “Power Transformer Fault Classification Based on Dissolved Gas Analysis by Implementing Bootstrap and Genetic Programming”, *IEEE Trans. Power Del.*, Vol. 39, No. 1, pp. 69–79, January. 2009.
- [15] W. Chen, C. Pan, Y. Yun and Y. Liu, “Wavelet Network in Power Transformers Diagnosis Using Dissolved Gas Analysis”, *IEEE Trans. Power Del.*, Vol.24, No. 1, pp. 187–194, 2009.

- [16] L. Mokhnache, A. Boubakeur "Comparison of different Back-propagation Algorithms in the Diagnosis of Transformer Oil", 2002 IEEE CEIDP, paper 2C-17, Cancun, Mexique, October 2002.
- [17] L. Mokhnache, A. Boubakeur "Fuzzy-Neuro System for Oil Insulated Transformer Diagnosis and Maintenance", IEE and COMADEM International, 17th International Congress and Exhibition on Condition Monitoring and Diagnosis Engineering Management, COMADEM 2004, Cambridge, UK, August 2004.
- [18] A. Smara, M. Bouktit, A. Boubakeur, Swarm Intelligence Metaheuristics Application in the Diagnosis of Transformer Oil, Book Chapter, Advances in Intelligent Systems and Computing, Springer, Volume 286, pp 413 422.
- [19] **Y. Benmahamed, M. Tegar and A. Boubakeur, "Application of SVM and KNN to Duval Pentagon 1 for transformer oil diagnosis," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 6, pp. 3443-3451, Dec. 2017. doi: 10.1109/TDEI.2017.006841.**
- [20] **Y. Benmahamed, Y. Kemari, M. Tegar and A. Boubakeur, "Diagnosis of Power Transformer Oil Using KNN and Naive Bayes Classifiers," 2018 *IEEE 2nd International Conference on Dielectrics (ICD)*, Budapest, 2018, pp. 1-4. doi: 10.1109/ICD.2018.8468532.**
- [21] **Y. Benmahamed, M. Tegar and A. Boubakeur, "Diagnosis of Power Transformer Oil Using PSO-SVM and KNN Classifiers," 2018 *International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)*, Algiers, 2018, pp. 1-4. doi: 10.1109/CISTEM.2018.8613548.**
- [22] IEC, "IEC 60076-7 First edition 2005-1 2 Power transformers Part 7: Loading guide for oil immersed power transformers," ed, 2005.
- [23] <http://www.tsv-transfo.com/fr/controler-element-transformateur/>
- [24] Y. Du, "Measurements and modeling of moisture diffusion processes in transformer insulation using interdigital dielectrometry sensors," PhD Thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1999.
- [25] T. K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," *Dielectrics and Electrical Insulation*, IEEE Transactions on, vol. 10, pp. 903-917, 2003.
- [26] K. S. Kassi, "Etude de l'isolation hybride en vue de son application dans les transformateurs de puissance" in *Mémoire de Maîtrise UQAC*, Chicoutimi : Université du Québec, Décembre 2013.
- [27] P.J. Vuarchex, "Huiles et liquides isolants", *Technique d'ingénieur D230, D231, D232*, 1986.
- [28] S. Abdi ; « Influence de vieillissement thermique sur les propriétés de l'huile de transformateur », Thèse de Doctorat d'Etat, Ecole Nationale Polytechnique d'Alger, Février 2012.

- [29] N. Berger ; « Liquides isolants en électrotechnique – Caractéristiques des produits », Techniques de l'Ingénieur, Traité de génie électrique, D2471, Novembre 2002.
- [30] Norme CEI 61099 ; « Spécifications pour esters organiques de synthèse neufs à usage électrique », Mai 1992.
- [31] Norme CEI 61203 ; « Esters organiques de synthèse à usage électrique – Guide de maintenance des ester pour transformateurs dans les matériels », Décembre 1992.
- [32] M. Rafi, Y.Z. Lv, Y. Zhou, K.B. Ma, W. Wang, C.R. Li and Q. Wang; «Use of vegetable oils as transformer oils – a review», Renewable and Sustainable Energy Reviews 52 (2015) 308–324.
- [33] Norme CEI 60296 ; « Fluides pour applications électrotechniques – Huiles minérales isolantes neuves pour transformateurs et appareils de connexion », Novembre 2003.
- [34] Norme CEI 60422 ; « Lignes directives de la maintenance et de la surveillance des huiles minérales isolantes en service dans les matériels électriques », Octobre 2005.
- [35] « Transformer oil handbook », Copyright Nynas Naphtenics AB, Sweden, 2004.
- [36] S. E. Chitour : « Raffinage du pétrole », Tome 1, Université des Sciences et de la Technologie Houari Boumediene, Alger, OPU 1983.
- [37] B. P. P. Wiklund, "Properties of Mineral Insulating Oils in Service," CIGRE, Issue 1, pp. 65- 73, 2010.
- [38] Norme CEI 60156 ; « Liquides isolants – Mesure de la tension de claquage à fréquence industrielle - Méthode d'essai », Août 1995.
- [39] R. Tobazeon ; « Préclaquage et claquage des liquides diélectriques », Techniques de l'ingénieur, Traité de génie électrique, D2450, Novembre 1997.
- [40] R. Musil, M. Baur and W. Pfister; « Testing practice for the AC breakdown voltage testing of insulating liquids », IEEE Electrical Insulating Magazine, Vol. 11, n°1, pp. 21-26, January/February 1995.
- [41] R. Benamar ; « Influence de nombre de claquage sur les propriétés de l'huile de transformateur », Mémoire de Magister, Ecole Nationale Polytechnique d'Alger, Juin 2008.
- [42] Norme ISO 3016 ; « Produits pétroliers - Détermination du Point d'écoulement », Août 1994.
- [43] Norme ISO 2719 ; « Détermination du point d'éclair - Méthode Pensky-Martens en vase clos », Novembre 2002.
- [44] A. Nacer, « Caractérisation des matériaux diélectriques, Application à l'huile minérale de transformateur », Thèse de Doctorat, Département d'électrotechnique, USTHB, 2008.
- [45] I. Fofana ; « Isolants Liquides en Haute Tension », Atelier CNHT 2016.
- [46] L. Mokhnache ; « Application des réseaux de neurones dans le diagnostic et la prédiction des isolations de haute tension », Thèse de Doctorat d'Etat, Ecole Nationale Polytechnique d'Alger, Janvier 2004.
- [47] ASTM D 1903-96. Standard test method for coefficient of thermal expansion of electrical insulating liquids of petroleum origin, and askarels.

- [48] Norme CEI 60814 ; « Isolants liquides — Cartons et papiers imprégnés d’huile — Détermination de la teneur en eau par titrage coulométrique de Karl Fischer automatique », August 1997.
- [49] Norme ISO 6618 ; « Produits pétroliers et lubrifiants – Détermination de l’indice d’acide ou de l’indice de base – Méthode par titrage en présence d’un indicateur coloré », Février 1997.
- [50] CEI 62021-1 «Insulating liquids - Determination of acidity -Part 1: Automatic potentiometric titration», September 2003.
- [51] J. S. N’CHO, "Développement de nouvelles méthodes de diagnostic et de régénération des huiles pour transformateurs de puissance," thèse de Doctorat, Ecole centrale de Lyon (France), 2011.
- [52] Y. Cui;" Investigation of data centric diagnostic techniques for transformer condition assessment", P.hD thesis ,University of Queensland, 2016.
- [53] M. Duval, A. dePablo, "Interpretation of Gas-In-Oil Analysis Using New IEC Publication 60599 and IEC TC 10 Databases", IEEE Electr. Insul. Mag., Vol. 17, No.2, pp.31-41, 2001M.
- [54] Duval, "A Review of Faults Detectable by Gas in Oil Analysis transformers ", IEEE Electr. Insul. Mag., Vol.18, No.3, pp. 8-17, 2002.
- [55] I. B. M. Taha, S. S. M. Ghoneim and A. S. A. Duaywah, "Refining DGA methods of IEC Code and Rogers four ratios for transformer fault diagnosis," *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Boston, MA, 2016, pp. 1-5.
- [56] I. B. M. Taha, S. S. M. Ghoneim and H. G. Zaini, "Improvement of Rogers four ratios and IEC Code methods for transformer fault diagnosis based on Dissolved Gas Analysis," *2015 North American Power Symposium (NAPS)*, Charlotte, NC, 2015, pp. 1-5
- [57] Pugh, D. R., "Advances in fault diagnosis by combustible gas analysis," *Doble Conference Index of Minutes*, Sec. 10-1201, 1974.
- [58] M. Duval, "New techniques for dissolved gas-in-oil analysis," in *IEEE Electrical Insulation Magazine*, vol. 19, no. 2, pp. 6-15, March-April 2003.
- [59] A. BEN ISHAK " sélection de variables par les machines à vecteurs supports pour la discrimination binaire et multi-classe en grande dimension ", thèse de doctorat, Institut Supérieur de Gestion de Tunis, 2007.
- [60] C. Cortes and V. Vapnik, "Support-vector network," *Mach. Learn.*, Vol. 20, pp. 273–297, 1995.
- [61] M. H. Bendiabdallah ; "L’utilisation des systèmes intelligents pour le diagnostic médical", Thèse de Doctorat, Université d'Oran 1 Ahmed Ben Bella.
- [62] J. Lehaire ;" Détection et caractérisation du cancer de la prostate par images IRM 1.5T multiparamétriques", thèse de doctorat, l’université de Lyon, 2016.
- [63] Z. Zidelmal ; "Reconnaissance d’arythmies cardiaques par Support Vector Machines (SVMs)", thèse de doctorat, université Mouloud Mammeri Tizi-Ouzou, 2012.

- [64] F. Barigou ; "Contribution à la catégorisation de textes et à l'extraction d'information", Thèse de Doctorat, Université d'Oran 1 Ahmed Ben Bella, 2012
- [65] C.-W. Hsu and C.-J. Lin, "A Comparison of Methods for Multiclass Support Vector Machines," *IEEE Trans. Neural Networks*, Vol. 13, pp. 415-425, 2002.
- [66] C.-C. Chang, C.-W. Hsu and C.-J. Lin, "The analysis of decomposition methods for support vector machines," *IEEE Trans. Neural Networks*, Vol. 11, No. 4, pp. 1003-1008, 2000.
- [67] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 2nd ed., Academic Press, Elsevier, Ch.3, 2003.
- [68] Boser, A., I. Guyon and V.N. Vapnik, 1992. A training algorithm for optimal margin classifiers. In *Fifth Annual Workshop on Computational Learning Theory*, ACM, pages 144-152, Pittsburgh.
- [69] Aizerman A., Braverman E.M. and Rozoner L.I., "Theoretical foundations of the potential function method in pattern recognition learning", *Automation and Remote Control*, Vol. 25, pp. 821-837, 1964.
- [70] J. Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Philos. Trans. Roy. Soc. London*, 209:415–446,, 1909.
- [71] A. Rakotomamonjy and F. Sudar ; "SVM Variable selection with application to pedestrian detection", *Congrès de Reconnaissance de Formes et d'Intelligence Artificielle 4*, Toulouse, France, 2004.
- [72] E. Mayoraz and E. Alpaydin. Support vector machines for multi-class classification. Technical Report 98-06, IDIAP, 1998.
- [73] S. Knerr, L. Personnaz, and G. Dreyfus. Single-layer learning revisited: A stepwise procedure for building and training a neural network. In F. Fogelman-Soulié and J. Héroult, editors, *Neurocomputing: Algorithms, Architectures and Applications*, volume F68 of NATO ASI Series, pages 4150. Springer-Verlag, 1990.
- [74] T.G. Dietterich and G. Bakiri. Solving multiclass learning problems via error-correcting output codes. *Journal of Artificial Intelligence Research*, 2 :263-286, 1995.
- [75] J. Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization, *Advances in Kernel Methods-Support Vector Learning*", MIT Press, pp. 185– 208, 1999.
- [76] R. E. Fan, P. H. Chen and C. J. Lin, "Working Set Selection Using the Second Order Information for Training SVM", *JMLR*, Vol. 6, pp. 1889-1918, 2005.
- [77] K.B. Duan¹, S. Keerthi, "Which is the best multiclass SVM method: an empirical study", Springer-Verlag, Berlin Heidelberg, pp. 278–285, 2005.
- [78] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification", *IEEE Trans. Inf. Theory*, Vol. 13, No. 1, pp. 21–27, 1967.
- [79] Eve Mathieu-Dupas, "Algorithme des K plus proches voisins pondérés et Application en diagnostic", nria-00494814, pp. 1-24, Jun 2010.

- [80] Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, Wiley-interscience, 2001 (ISBN 0-471-05669-3).
- [81] Glover F., « Future paths for integer programming and links to artificial intelligence». *Computers and Operations Research*, Vol. 13, pp. 533–549, 1986.
- [82] J. Kennedy and R. Eberhart, "Particle swarm optimization," *IEEE Int'l. Conf. Neural Networks*, Perth, pp. 1942-1948, 1995.
- [83] Q. Bai, "Analysis of particle swarm optimization algorithm," *Comput. Inf. Sci.*, Vol. 3, No. 1, pp. 180–184, Feb. 2010.
- [84] Y. Cooren, "Perfectionnement d'un algorithme adaptatif d'Optimisation par Essaim Particulaire, Applications en génie médical et en électronique", thèse de Doctorat, Université Paris-Est Créteil, 2008.
- [85] R. Mendes, J. Kennedy, and J. Neves. "The Fully Informed Particle Swarm: Simpler, Maybe Better". *IEEE Trans. Evolutionary Computation*, Vol. 8, No. 3, pp. 204–210, 2004.
- [86] M. Clerc and J. Kennedy. "The particle swarm: explosion, stability, and convergence in multi-dimensional complex space". *IEEE Transactions on Evolutionary Computation*, Vol. 6, pp. 58–73, 2002.
- [87] Y. Shi, R. Eberhart "A modified particle swarm optimizer" *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence, The 1998 IEEE International Conference*, Anchorage, AK, pp. 69 - 73 4-9 May 1998.
- [88] Eberhart R. C. & Shi Y., «Particle Swarm Optimization: Developments, Applications and Resources », *IEEE Xplore Digital Library*, 2001.
- [89] X.-S. Yang *Engineering Optimization: An Introduction with Metaheuristic Applications*, New Jersey, Wiley, July 2010, p. 173-179.
- [90] Abbas El Dor : " Perfectionnement des algorithmes d'optimisation par essaim particulaire : applications en segmentation d'images et en électronique ", thèse de doctorat, Université Paris-Est, 2012.
- [91] K.E. Parsopoulos and M.N. Vrahatis, "Particle Swarm Optimization and Intelligence: Advances and Applications", New York, Information Science Reference; 2010.
- [92] McGill, Robert; Tukey, John W.; Larsen, Wayne A. (February 1978). "Variations of Box Plots". *The American Statistician*. 32 (1): 12–16. doi:10.2307/2683468.



Annexe

Annexe 1

Cas	H2 (ppm)	CH4 (ppm)	C2H2 (ppm)	C2H4 (ppm)	C2H6 (ppm)	Inspection	Triangle de Duval	Pentagone de Duval	Rogers	CEI 60599
1	16000	3600	0	14	670	DP	DP	S	NF	UD
2	6600	1000	19	2	38	DP	D1	PD	UD	UD
3	78	20	28	13	11	D1	D1	D1	UD	D1
4	305	100	541	161	33	D1	D1	D1	UD	D1
5	35	6	482	26	3	D1	D1	D1	UD	D1
6	543	120	1880	411	41	D1	D1	D1	UD	D1
7	1230	163	692	233	27	D1	D1	D1	D2	D1
8	6870	1028	5500	900	79	D1	D1	D1	UD	D1
9	10092	5399	37565	6500	530	D1	D1	D1	UD	UD
10	650	81	270	51	170	D1	D1	D1	UD	UD
11	210	22	7	6	6	D1	D1	PD	UD	UD
12	385	60	159	53	8	D1	D1	D1	D2	D1
13	60	10	4	4	4	D1	D1	S	UD	UD
14	95	10	39	11	1	D1	D1	D1	UD	D1
15	8200	3790	5830	4620	250	D2	D2	D2	D2	D1
16	210	43	187	102	12	D2	D2	D2	D2	D1
17	2850	1115	3675	1987	138	D2	D2	D2	D2	D1
18	7020	1850	4410	2960	0,1	D2	D2	D2	D2	D1
19	545	130	239	153	16	D2	D2	D1	D2	D1
20	7150	1440	1760	1210	97	D2	D2	D2	D2	D2
21	620	325	244	181	38	D2	D2	D2	D2	D1
22	120	31	94	66	0,1	D2	D2	D2	D2	D1
23	755	229	460	404	32	D2	D2	D2	D2	D2
24	5100	1430	1010	1140	0,1	D2	D2	D2	D2	D2
25	13500	6110	4040	4510	212	D2	D2	D2	D2	D2
26	1570	1110	1830	1780	175	D2	D2	D2	D2	D2
27	10000	6730	10400	7330	345	D2	D2	D2	D2	D1
28	1820	405	634	365	35	D2	D2	D2	D2	D1
29	535	160	680	305	16	D2	D2	D2	UD	D1
30	13	3	6	3	1	D2	D2	D2	D2	D1
31	137	67	104	53	7	D2	D2	D2	D2	D1
32	1084	188	769	166	8	D2	D1	D1	UD	D1
33	34	21	56	49	4	D2	D2	D2	D2	D2
34	7940	2000	5390	3120	355	D2	D2	D2	D2	D1
35	150	130	30	55	9	D2	D2	T1&T2	D2	UD
36	1270	3450	8	1390	520	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
37	3420	7870	33	6990	1500	T1&T2	T1&T2	T3	T3	T3
38	360	610	9	260	259	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
39	1	27	1	4	49	T1&T2	T1&T2	T1&T2	UD	T1&T2
40	3675	6392	5	7691	2500	T1&T2	T3	T3	T3	T1&T2
41	48	610	0,1	10	29	T1&T2	PD	T1&T2	UD	T1&T2
42	12	18	0,1	4	4	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
43	8800	64064	0,1	95650	72128	T3	T3	T1&T2	T1&T2	T1&T2
44	6709	10500	750	17700	1400	T3	T3	T3	T3	T3
45	1100	1600	26	2010	221	T3	T3	T3	T3	T3
46	290	966	57	1810	299	T3	T3	T3	T3	T3
47	2500	10500	6	13500	4790	T3	T3	T3	T1&T2	T1&T2
48	1860	4980	1600	10700	0,1	T3	D1	T3	UD	T3
49	860	1670	40	2050	30	T3	T3	T3	T3	T3
50	150	22	11	60	9	T3	T3	D1	D2	UD

Annexe 2

Cas	H2 (ppm)	CH4 (ppm)	C2H2 (ppm)	C2H4 (ppm)	C2H6 (ppm)	Inspection	Pentagone de Duval I	Pentagone de Duval-SVM-PSO	Pentagone de Duval-KNN
1	16000	3600	0	14	670	DP	S	DP	DP
2	6600	1000	19	2	38	DP	DP	DP	DP
3	78	20	28	13	11	D1	D1	D2	D2
4	305	100	541	161	33	D1	D1	D1	D2
5	35	6	482	26	3	D1	D1	D1	D1
6	543	120	1880	411	41	D1	D1	D1	D1
7	1230	163	692	233	27	D1	D1	D1	D1
8	6870	1028	5500	900	79	D1	D1	D1	D1
9	10092	5399	37565	6500	530	D1	D1	D1	D1
10	650	81	270	51	170	D1	D1	D1	D1
11	210	22	7	6	6	D1	DP	D1	PD
12	385	60	159	53	8	D1	D1	D1	D1
13	60	10	4	4	4	D1	S	D1	D1
14	95	10	39	11	1	D1	D1	D1	D1
15	8200	3790	5830	4620	250	D2	D2	D2	D2
16	210	43	187	102	12	D2	D2	D2	D2
17	2850	1115	3675	1987	138	D2	D2	D2	D2
18	7020	1850	4410	2960	0,1	D2	D2	D2	D2
19	545	130	239	153	16	D2	D2	D2	D2
20	7150	1440	1760	1210	97	D2	D1	D2	D2
21	620	325	244	181	38	D2	D2	D2	D2
22	120	31	94	66	0,1	D2	D2	D2	D2
23	755	229	460	404	32	D2	D2	D2	D2
24	5100	1430	1010	1140	0,1	D2	D2	D2	D2
25	13500	6110	4040	4510	212	D2	D2	D2	D2
26	1570	1110	1830	1780	175	D2	D2	D2	D2
27	10000	6730	10400	7330	345	D2	D2	D2	D2
28	1820	405	634	365	35	D2	D2	D2	D2
29	535	160	680	305	16	D2	D2	D2	D2
30	13	3	6	3	1	D2	D2	D2	D2
31	137	67	104	53	7	D2	D2	D2	D2
32	1084	188	769	166	8	D2	D1	D1	D1
33	34	21	56	49	4	D2	D2	D2	D2
34	7940	2000	5390	3120	355	D2	D2	D2	D2
35	150	130	30	55	9	D2	T1&T2	D2	T1&T2
36	1270	3450	8	1390	520	T1&T2	T1&T2	T1&T2	T1&T2
37	3420	7870	33	6990	1500	T1&T2	T3	T3	T3
38	360	610	9	260	259	T1&T2	T1&T2	T1&T2	T1&T2
39	1	27	1	4	49	T1&T2	T1&T2	T1&T2	T1&T2
40	3675	6392	5	7691	2500	T1&T2	T3	T3	T3
41	48	610	0,1	10	29	T1&T2	T1&T2	T1&T2	T1&T2
42	12	18	0,1	4	4	T1&T2	T1&T2	T1&T2	T1&T2
43	8800	64064	0,1	95650	72128	T3	T1&T2	T1&T2	T1&T2
44	6709	10500	750	17700	1400	T3	T3	T3	T3
45	1100	1600	26	2010	221	T3	T3	T3	T3
46	290	966	57	1810	299	T3	T3	T3	T3
47	2500	10500	6	13500	4790	T3	T3	T3	T3
48	1860	4980	1600	10700	0,1	T3	T3	T3	T3
49	860	1670	40	2050	30	T3	T3	T3	T3
50	150	22	11	60	9	T3	D1	D2	D2

Annexe 3

Cas	Inspection	DGA en ppm	DGA en %	Ratios de Rogers	Ratios de Duval	Ratios de Dornenburg
1	DP	DP	DP	DP	DP	DP
2	DP	DP	DP	DP	DP	T1&T2
3	D1	D2	D2	DP	D1	D1
4	D1	D2	D2	D1	D1	D1
5	D1	D1	D1	D1	D1	D1
6	D1	D1	D1	D2	D1	D1
7	D1	D1	D1	D1	D1	D1
8	D1	D1	D1	D2	D1	D1
9	D1	D1	D1	D2	D1	D1
10	D1	D2	D1	DP	D1	D1
11	D1	D1	DP	DP	D1	D1
12	D1	D1	D1	D1	D1	D1
13	D1	D1	D1	DP	D1	D1
14	D1	D1	D1	D2	D1	D1
15	D2	D2	D2	D2	D2	D2
16	D2	D2	D2	D2	D2	D2
17	D2	D2	D2	D2	D2	D1
18	D2	D2	D2	D2	D2	D1
19	D2	D2	D2	D2	D1	D2
20	D2	D1	D2	D2	D2	D2
21	D2	D2	D2	D2	D2	D2
22	D2	D2	D2	D2	D2	D2
23	D2	D2	D2	D2	D2	D2
24	D2	D1	D2	D2	D2	D2
25	D2	D2	D2	D2	D2	D2
26	D2	D2	D2	D2	D2	D2
27	D2	D2	D2	D2	D2	D2
28	D2	D2	D2	D2	D2	D2
29	D2	D2	D2	D2	D2	D2
30	D2	D2	D2	D2	D2	D2
31	D2	D2	D2	D2	D2	D2
32	D2	D1	D1	D2	D1	D1
33	D2	D2	D2	D2	D2	D2
34	D2	D2	D2	D2	D2	D1
35	D2	T1&T2	D2	T3	T1&T2	D2
36	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
37	T1&T2	T3	T3	T3	T3	T1&T2
38	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
39	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T3
40	T1&T2	T1&T2	T3	T1&T2	T1&T2	T1&T2
41	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
42	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
43	T3	T3	T3	T1&T2	T3	DP
44	T3	T3	T3	D2	T3	T3
45	T3	T3	T3	T3	T3	T3
46	T3	T3	T3	T3	T3	T3
47	T3	T3	T3	T3	T3	T1&T2
48	T3	T3	T3	D2	T3	T3
49	T3	T3	T3	D2	T3	T3
50	T3	D2	D2	T3	T3	T3

Annexe 4

Cas	Inspection	DGA en ppm	DGA en %	Ratios de Rogers	Ratios de Duval	Ratios de Dornenburg
1	PD	DP	DP	DP	DP	DP
2	PD	D1	DP	D1	DP	D1
3	D1	D1	D1	D2	D1	D2
4	D1	D1	D1	D1	D1	D1
5	D1	D1	D1	D1	D1	D1
6	D1	D1	D1	D1	D1	D1
7	D1	D1	D1	D1	D1	D2
8	D1	D2	D1	D1	D1	D1
9	D1	D2	D1	D1	D1	D1
10	D1	D1	T1&T2	D1	D1	D1
11	D1	DP	D1	DP	T1&T2	D1
12	D1	D1	D1	D1	D1	D2
13	D1	DP	D1	DP	T1&T2	D1
14	D1	D1	D1	D1	D1	D1
15	D2	D1	D2	D2	D2	D2
16	D2	D2	D2	D2	D2	D2
17	D2	D2	D2	D1	D2	D2
18	D2	D1	D2	D2	D2	D2
19	D2	D2	D1	D2	D2	D2
20	D2	D1	D2	D2	D2	D2
21	D2	D1	D2	D2	D2	D2
22	D2	D1	D2	D2	D2	D2
23	D2	D1	D2	D1	D2	D2
24	D2	D2	D2	D2	D2	D2
25	D2	D1	D2	D2	D2	D2
26	D2	D2	D2	D2	D2	D2
27	D2	D1	D1	D2	D2	D2
28	D2	D1	D2	D2	D1	D2
29	D2	DP	D2	D2	D2	D2
30	D2	DP	D2	D2	D2	D2
31	D2	D1	D2	D2	D2	D2
32	D2	D1	D1	D1	D1	D1
33	D2	D1	D2	D2	D2	D2
34	D2	D2	D2	D2	D2	D2
35	D2	D1	D2	D2	T1&T2	D2
36	T1&T2	D2	T1&T2	T3	T1&T2	T1&T2
37	T1&T2	T3	T3	T3	T3	T1&T2
38	T1&T2	D1	T1&T2	T3	T1&T2	T1&T2
39	T1&T2	DP	T1&T2	T3	T1&T2	T1&T2
40	T1&T2	T3	T3	T3	T3	T1&T2
41	T1&T2	DP	T1&T2	T3	DP	T1&T2
42	T1&T2	DP	T1&T2	T3	T1&T2	T1&T2
43	T3	T1&T2	T3	T3	T3	T1&T2
44	T3	T1&T2	T3	T3	T3	T1&T2
45	T3	D2	T3	T3	T3	T1&T2
46	T3	D2	T3	T3	T3	T1&T2
47	T3	T3	T3	T3	T3	T1&T2
48	T3	D2	T3	D2	T3	D2
49	T3	D2	T3	D2	T3	D2
50	T3	D1	D2	D2	T3	D1

Annexe 5

Cas	Inspection	DGA en ppm	DGA en %	Ratios de Rogers	Ratios de Duval	Ratios de Dornenburg
1	PD	DP	DP	DP	DP	DP
2	PD	DP	DP	D1	DP	DP
3	D1	D1	D1	D1	D1	D1
4	D1	D1	D1	D1	D1	D1
5	D1	D1	D1	D1	D1	D1
6	D1	D1	D1	D1	D1	D1
7	D1	D1	D1	D1	D1	D1
8	D1	D1	D1	D1	D1	D1
9	D1	D1	D1	D1	D1	D1
10	D1	D1	D1	D1	D1	D1
11	D1	T1&T2	D1	DP	D1	D1
12	D1	D1	D1	D1	D1	D1
13	D1	D2	D1	DP	D1	D1
14	D1	D1	D1	D1	D1	D1
15	D2	D2	D2	D2	D2	D2
16	D2	D2	D2	D2	D2	D2
17	D2	D2	D2	D2	D2	D1
18	D2	D2	D2	D2	D2	D2
19	D2	D1	D2	D2	D2	D2
20	D2	D2	D2	D2	D2	D2
21	D2	D2	D2	D2	D2	D1
22	D2	D2	D2	D2	D2	D2
23	D2	D2	D2	D2	D2	D1
24	D2	D2	D2	D2	D2	D2
25	D2	D2	D2	D2	D2	D2
26	D2	D2	D2	D2	D2	D2
27	D2	D2	D2	D2	D2	D2
28	D2	D2	D2	D2	D2	D2
29	D2	D2	D2	D2	D2	D2
30	D2	D2	D2	D2	D2	D2
31	D2	D2	D2	D2	D2	D1
32	D2	D1	D1	D1	D1	D1
33	D2	D2	D2	D2	D2	D2
34	D2	D2	D2	D2	D2	D2
35	D2	T1&T2	D2	D2	D2	D2
36	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
37	T1&T2	T3	T3	T3	T1&T2	T1&T2
38	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
39	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
40	T1&T2	D2	T1&T2	T1&T2	T1&T2	T3
41	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
42	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2	T1&T2
43	T3	T3	T3	T1&T2	T3	T1&T2
44	T3	T3	T3	T3	T3	T3
45	T3	T3	T3	T3	T3	T3
46	T3	T3	T3	T3	T3	T3
47	T3	T3	T3	T3	T3	T3
48	T3	T3	T3	D2	T3	D1
49	T3	T3	T3	T3	T3	D2
50	T3	D2	D2	D2	T3	T3

Travaux effectués dans le cadre de cette thèse de Doctorat D/LMD

Publication internationale

- Youcef Benmahamed, Madjid Tegar and Ahmed Boubakeur, "**Application of SVM and KNN to Duval Pentagon 1 for transformer oil diagnosis**," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 6, pp. 3443-3451, Dec. 2017. doi: **10.1109/TDEI.2017.006841**.

Communications internationales

- Youcef Benmahamed, Youcef. Kemari, Madjid Tegar and Ahmed Boubakeur, "**Diagnosis of Power Transformer Oil Using KNN and Nave Bayes Classifiers**," *2018 IEEE 2nd International Conference on Dielectrics (ICD)*, Budapest, 2018, pp. 1-4. doi: **10.1109/ICD.2018.8468532**

- Y. Benmahamed, M. Tegar and A. Boubakeur, "**Diagnosis of Power Transformer Oil Using PSO-SVM and KNN Classifiers**," *2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)*, Algiers, 2018, pp. 1-4. doi: **10.1109/CISTEM.2018.8613548**