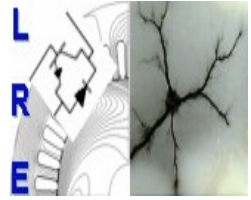




المدرسة الوطنية المتعددة التخصصات
Ecole Nationale Polytechnique

Ecole Nationale Polytechnique
Département d'Electrotechnique
Laboratoire de Recherche en Electrotechnique



Mémoire de Master en Electrotechnique

Présenté par :

Ayyoub ZOUAGHI

Intitulé

Classification de la sévérité de pollution des isolateurs haute tension par la méthode support vecteur machine

Directeur(s) du mémoire :

A.MEKHALDI	Professeur	ENP
B.MOULA	M.A.B	UMMTO

Membres de jury:

R.IBTIOUEN	Professeur	ENP
O.TOUHAMI	Professeur	ENP
S.MEKHTOUB	Professeur	ENP

ENP 2013

(Arial 8)

ملخص :

في العمل الحالي، يتم إجراء تصنيف شدة تلوث عوازل التوتر العالي بناءً على قواعد و نصوص IEC 71، ويكون ذلك باستغلال قواعد بيانات جمعت في مختبر التوتر العالي في المدرسة الوطنية المتعددة التقنيات على نماذج مسطحة. أولاً، سنطبق DWT على إشارات التيار المتسرب لاستخراج خصائص التفاصيل الخاصة بهم من أجل دراسة العلاقة بين إشارات التيار المتسرب وشدة التلوث. ثانياً، سنقوم بتصنيف شدة التلوث باستعمال طريقة SVM مع اثنين من الاستراتيجيات و باستخدام وظائف نواة مختلفة في حالة تلوث منتظم و غير منتظم.

الكلمات الرئيسية:

شدة تلوث، تلوث منتظم ، تلوث غير منتظم، التيار المتسرب ، التنبؤ المتحكم فيه ، التصنيف، تقنية الموجات المستمرة، تقنية الموجات المتقطعة.

Résumé :

Dans le présent travail, la classification de la sévérité de pollution des isolateurs de haute tension est effectuée, suivant les recommandations de la CEI 71, en exploitant des bases de données établies au niveau du laboratoire de haute tension de l'Ecole Nationale Polytechnique sur des modèles plans. Dans un premier lieu, nous appliquons la DWT sur les signaux du courant de fuite pour extraire les caractéristiques de leurs détails afin d'étudier la corrélation entre les signaux du courant de fuite et la sévérité de pollution. En deuxième lieu, nous effectuons une classification de la sévérité de pollution par la méthode SVM avec deux stratégies en utilisant différentes fonctions noyaux pour les deux cas de pollution uniforme et non uniforme.

Mots clés :

Sévérité de pollution, pollution uniforme, pollution non uniforme, courant de fuite, apprentissage supervisé, classification, support vecteur machine, transformée en ondelettes discrète.

Abstract:

In this work, the classification of the pollution severity of polluted insulators is carried out, according to IEC 60-1 recommendations, by using databases set at high voltage laboratory of Ecole Nationale Polytechnique on plan models. Firstly, we apply the DWT of the leakage current signals to extract the characteristics from their details in order to study the correlation between the signals of the leakage current and the severity of pollution. Then, we perform a classification of the pollution severity using SVM method with two strategies and various kernel functions for both uniform and non-uniform pollution.

Key words:

Pollution severity, uniform pollution, not uniform pollution, leakage current, supervised training, classification, Support Vector Machine, Discrete Wavelet Transform.

SOMMAIRE

Introduction générale.....	1
----------------------------	---

Chapitre I : Analyse du signal du courant de fuite par la technique des ondelettes

Introduction	3
I.1. La transformée en ondelettes	3
I.2. L'analyse multi-résolution (AMR)	4
I.3. La transformée en ondelettes discrètes (DWT)	4
I.3.1. Niveau maximal de décomposition	5
I.3.2. Choix de l'ondelette analysante	5
I.4. Décomposition du signal du courant de fuite par DWT	6
I.5. Calcul des caractéristiques des coefficients de détails d'ondelette	7
I.5.1. Etude des écart types des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin	8
I.5.2. Etude des maximums des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin	10
I.5.3. Etude des minimums des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin	12
I.5.4. Etude des moyennes des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin	14
Conclusion	16

Chapitre II : Classification de la sévérité de pollution des isolateurs par la méthode SVM

Introduction	17
II.1. Théorie de l'apprentissage	17
II.2. Les machines à vecteurs de support	19
II.3. Principe de la technique SVM	19
II.4. Les SVM binaires	20
II.4.1. Notion de base	20

II.4.1.1. Hyperplan	20
II.4.1.2. Vecteur de support	20
II.4.1.3. Marge	21
II.4.2. Construction de l'hyperplan optimal	21
II.4.2.1. Cas des données linéairement séparables	21
II.4.2.2. Cas des données non-linéairement séparables.....	24
II.4.3. L'astuce du noyau (kernel trick).....	26
II.4.3.1. Exemples de noyaux	27
II.5. SVM pour le cas multi-classes	28
II.5.1. La stratégie un contre tous.....	28
II.5.2. La stratégie un contre un.....	29
II.6. Avantages et inconvénients des SVM	31
II.7. Les Classes de pollution	32
II.8. Application des SVM pour la classification de sévérité de pollution des isolateurs	33
II.8.1. Composition du vecteur d'état	33
II.8.2. Entraînement du modèle	35
II.8.2.1. Base de données pour le cas de la pollution uniforme	35
II.8.2.2. Base de données pour le cas de la pollution non uniforme	35
II.8.3. Les différentes fonctions noyaux appliquées	35
II.9. Validation et résultat de classification	36
II.9.1. Cas de pollution uniforme	36
II.9.2. Cas de pollution non uniforme	38
II.10. Interprétation des résultats	40
II.10.1. Pollution uniforme	40
II.10.2. Pollution non uniforme	41
Conclusion	41
Conclusion générale	43
Références bibliographiques	

Introduction générale

Les lignes aériennes et les postes des réseaux de transport d'énergie électrique sont exposés à diverses contraintes. Parmi celles-ci, la pollution des isolateurs qui constitue l'un des facteurs de première importance dans la qualité et la fiabilité du transport d'énergie. En effet par temps de pluie ou de brouillard, les dépôts polluants se fixant sur les surfaces isolantes réduisent considérablement la résistivité superficielle et le contournement peut alors survenir [1].

Le phénomène de pollution des isolateurs constitue un sujet d'étude d'une importance primordiale. Cela provient du fait que la pollution des isolateurs constitue un sérieux problème pour l'exploitation des réseaux de haute tension. En effet, et suite à l'interaction entre l'air transportant des grains de poussières et l'isolateur, une couche étrangère se dépose sur les surfaces de celui-ci [2]. Une fois cette couche humidifiée, la tension de tenue peut diminuer considérablement, provoquant des incidents très graves.

Le choix des isolateurs à installer dans le réseau ne peut se faire que si la sévérité de pollution des différents sites concernés est connue. En effet, la mesure de cette sévérité est indispensable afin de dimensionner convenablement les isolateurs susceptibles d'assurer un service sans défaillance dans un site donné [3].

La mesure de la sévérité est généralement basée sur la détermination expérimentale de la conductance superficielle de la couche polluante ou la détermination au laboratoire du courant de fuite d'un isolateur pollué artificiellement [4].

L'apprentissage supervisé a émergé ces dernières années comme un outil efficace pour la classification dans de nombreux domaines [5]. Les Machines à Vecteurs de Supports (SVM) sont une de ces techniques de classification, elles ont été appliquées à de très nombreux domaines (bioinformatique, recherche d'information, vision par ordinateur, finance etc...) [6], et récemment pour le diagnostic des isolateurs en haute tension.

Le présent travail consiste à utiliser le signal du courant de fuite pour extraire des critères de jugement sur l'état de surface de l'isolateur pollué et servir d'entrées pour un classifieur SVM dont les sorties représentent la classe de pollution assignée à un isolateur pollué donné.

Dans le premier chapitre, nous donnons des rappels sur la transformée en ondelettes, plus particulièrement la transformée en ondelette discrète (DWT). Nous appliquons par la suite la DWT sur les signaux du courant de fuite pour extraire les caractéristiques de ces

détails afin de les utilisées pour la construction des vecteurs d'état de notre modèle SVM par la suite.

Dans le deuxième chapitre, dans une première partie nous présentons dans un premier temps la théorie d'apprentissage ainsi que la théorie de la méthode des Machines à Vecteurs de Supports qui représente un type d'apprentissage supervisé. Ensuite, nous nous intéressons à la classification basée sur les SVM dans le cas binaire ou on a deux classes et les différentes stratégies de classification dans le cas multi-classe. En outre, nous parlons aussi sur les fonctions noyaux (kernel function) afin de les exploiter dans notre modèle SVM par la suite.

Dans une deuxième partie, Nous présentons par la suite les différentes classes associées à chaque niveau de pollution selon la CEI 71 ainsi que les résultats de classification obtenus par les modèles SVM construits et nous comparons les résultats obtenus par les différentes stratégies et avec les différentes fonctions noyaux pour obtenir le meilleur classifieur dans les deux cas de pollution uniforme et non uniforme.

En dernier lieu, nous terminons par une conclusion générale qui représente une synthèse globale des principales observations au cours de l'application des différentes méthodes, ainsi que les perspectives qui permettront de réaliser une étude plus approfondie.

Chapitre I

Analyse du courant
de fuite par la
technique des
ondelettes

Introduction :

La transformée en ondelettes a l'avantage de décrire le signal dans le domaine temps / fréquence et donc de palier l'inconvénient de la transformée de Fourier qui a le désavantage de la perte de l'information temporelle. Le pavage du plan temps-fréquence induit par cette transformée a pour particularité de permettre une résolution temporelle fine aux hautes fréquences et une résolution fréquentielle fine aux basses fréquences. Cette propriété permet souvent une analyse intéressante du signal mais reste rigide. Il existe deux types de la transformée en ondelettes : la transformée en ondelettes continue (CWT pour Continuous Wavelet Transform) et la transformée en ondelettes discrète (DWT pour Discrete Wavelet Transform). Chaque transformation a ses avantages et ses inconvénients [7].

Dans le présent chapitre, nous abordons d'une manière explicite la transformée en ondelettes discrète (DWT). Nous passons ensuite à l'analyse du signal du courant de fuite par la DWT en utilisant l'ondelette de Daubechies 4 comme ondelette analysante afin d'extraire les critères nécessaires à l'étude de l'état de surface des isolateurs pour les deux types de pollution : uniforme et non uniforme.

I.1. La transformée en ondelettes :

La théorie des ondelettes est apparue au début des années 1990. Elle touche de nombreux domaines des mathématiques, notamment le traitement du signal et des images. Le traitement du signal par ondelettes permet d'élargir les intervalles de temps lorsque l'on recherche plus de précision sur les composantes de basses fréquences, elle permet également de rétrécir les zones pour l'extraction d'informations sur les hautes fréquences cela démontre la capacité des ondelettes à garder une bonne résolution temporelle et fréquentielle. L'analyse par la méthode des ondelettes nous permet aussi de partitionner l'énergie du signal en différentes bandes de fréquences pour séparer toutes ses composantes [8, 9].

La transformée en ondelette permet d'appliquer une analyse multi-résolution (AMR) sur le signal étudié. Cette analyse qu'il conviendrait d'appeler temps-échelle, utilise une gamme très étendue d'échelles pour analyser le signal. L'analyse multi-résolution de la transformée en ondelettes est équivalente à une décomposition atomique temps-échelle. Chacun des atomes peut s'interpréter comme étant une projection locale du signal analysé et il est obtenu à partir d'une ondelette $\Psi(t)$ unique par une translation en temps et une dilatation.

Partant d'une fonction bien localisée; dans le plan temps-échelle, on lui associe la famille d'ondelettes $\Psi_{(b,a)}(t)$ générées par des translations et des dilatations de $\Psi(t)$:

$$\Psi_{(b,a)}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad , b \in \mathbb{R}, a > 0 \quad (\text{I.1})$$

Les ondelettes sont de formes d'ondes standards, proportionnelles au paramètre de dilatation a (variable d'échelle). La transformée en ondelettes est aussi interprétée comme étant un processus de filtrage du signal analysé par un filtre passe-bande de bande passante variable. C'est le paramètre a qui fixe la valeur de cette bande [10].

I.2. L'analyse multi-résolution (AMR):

L'analyse multi résolution consiste à décomposer successivement le signal suivant un terme d'approximation a_i et un terme de détail d_i aux différentes échelles. Cette décomposition est obtenue en appliquant un banc de filtres orthogonaux construit de façon à permettre une reconstruction parfaite du signal initial. Ces filtres correspondent en fait à une paire d'ondelettes analysantes dites fonction d'échelle et fonction ondelette. L'implémentation de l'analyse multi résolution est facilitée par l'emploi de filtres grâce à un schéma de type pyramidal. Le signal s peut alors s'exprimer de la manière suivante : [11]

$$s \cong d_1 + d_2 + d_3 \dots + d_i + a_i$$

I.3. La transformée en ondelettes discrètes (DWT) :

La transformée en ondelette discrète (Discrete Wavelet Transform ou DWT) est considérée comme un processus de décomposition du signal. Le signal d'origine $s(t)$ traverse deux filtres complémentaires : passe-haut et passe-bas, et émerge en tant que deux signaux: respectivement le signal d'approximations A et le signal de détails D comme le montre la Figure suivante [12].

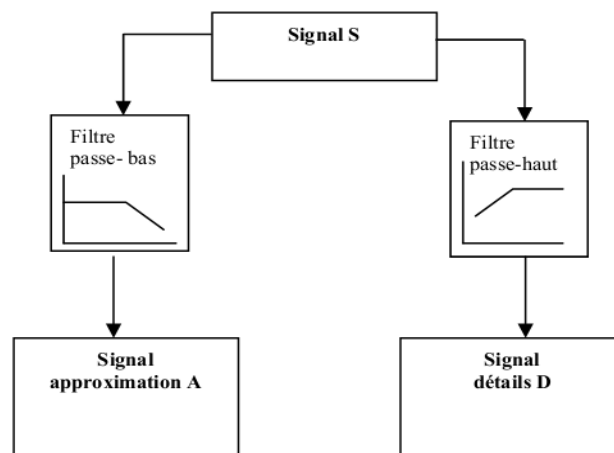


Figure I.1: décomposition du signal s en approximations et détails.

Pour beaucoup de signaux, le contenu de basse fréquence est la pièce la plus importante. Il donne au signal son identité. Le contenu à haute fréquence, d'autre part, nous donne la saveur ou la nuance. Dans la transformée en ondelette, on dit souvent l'approximation et le détail [12].

I.3.1. Niveau maximal de décomposition :

Avant de décomposer un signal en approximations et détails, il y a lieu de considérer un niveau maximal de décomposition. Pour la DWT, le nombre maximal pour lequel un signal est décomposé est obtenu par la relation [13]:

$$J_{ful} = \text{fix}(\log_2(N)) \quad (\text{I.2})$$

N est la longueur du signal échantillonné et la fonction fix donne le nombre entier le plus proche de la valeur obtenue après calcul du logarithme à base 2 de N .

I.3.2. Choix de l'ondelette analysante :

Le choix de l'ondelette analysante joue un rôle primordial pour la compréhension des phénomènes transitoires comme les impulsions des décharges électriques observées sur les signaux du courant de fuite pour un état de pollution d'isolateurs sous des conditions humides. Il est cependant connu que les ondelettes Daubechies 4 'db4' et Daubechies 6 'db6' sont plus appropriées pour la détection des perturbations de courtes durées ; alors que les ondelettes db8 et db10 sont adéquates pour l'étude des perturbations de relatives longues durées [14].

L'expression mathématique de l'ondelette mère complexe de Daubechies 4 est :

$$\Psi(t) = e^{-\frac{t^2}{2}} * (e^{-j\omega t} - e^{-\frac{\omega^2}{2}}) \quad (\text{I.3})$$

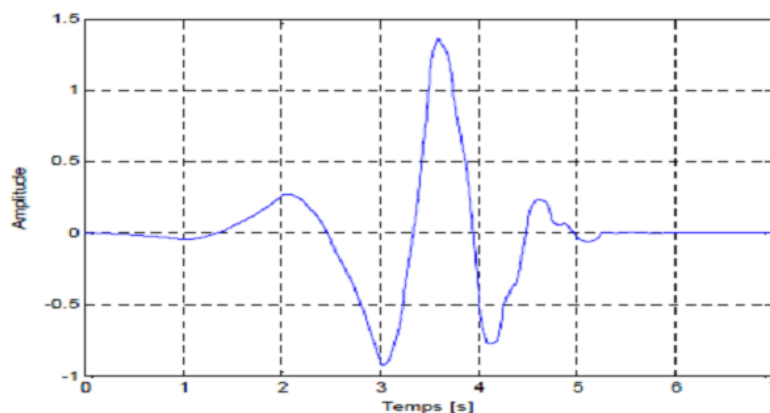


Figure I.2: Forme d'onde de l'ondelette de Daubechies 4.

L'ondelette de Daubechies 4 est choisie pour appliquer la DWT sur le signal du courant de fuite qui affiche des pics de hautes fréquences à partir d'un certain niveau de tension. Cette ondelette semble adéquate pour la détermination des bandes fréquentielles qui caractérisent les perturbations de hautes fréquences via l'analyse multi-résolution. A cet effet, la détermination du niveau de décomposition est importante pour récupérer de bonnes informations sans trop de calculs.

I.4. Décomposition du signal du courant de fuite par DWT :

Dans le présent travail, nous étudions la corrélation qui pourrait exister entre le signal courant de fuite (en termes de forme et d'amplitude) et la sévérité de la pollution caractérisant l'état de surface de l'isolateur. Pour cela nous allons appliquer la DWT sur les différents signaux du courant de fuite. On peut décomposer ces signaux par la DWT en utilisant différentes ondelettes mère (Haar, Daubechies 4, Coiflet, Daubechies 8). L'ondelette la plus adéquate pour notre étude est l'ondelette de Daubechies 4.

Le choix du niveau de décomposition est lié à la taille de notre signal, il est calculé à l'aide de l'équation (I.2). Pour notre étude, la taille des signaux analysés est de 10000 points donc la décomposition par DWT doit être réalisée jusqu'au niveau 10 afin de récupérer toutes les informations fréquentielles et temporelles contenues dans les signaux du courant de fuite. La figure suivante représente un signal du courant de fuite et ses 10 détails.

La base de données utilisée dans notre étude a été réalisée par Mr B.MOULA [15], les essais ont été effectués au laboratoire de haute tension de l'Ecole Nationale Polytechnique. Pour Cinq conductivités volumiques étudiées sous pollution uniforme à savoir (1.2; 3.5; 6; 10 et 20mS/cm) et pour plusieurs paliers de tension appliquée, il prélève cinq fois le signal du courant de fuite via un oscilloscope numérique TEKTRONIX. La forme d'onde du courant de fuite est très sensible à l'état de surface de l'isolateur pollué. Par ailleurs, l'amplitude du courant de fuite augmente avec l'augmentation de la tension appliquée et/ou de la conductivité volumique de la solution polluante.

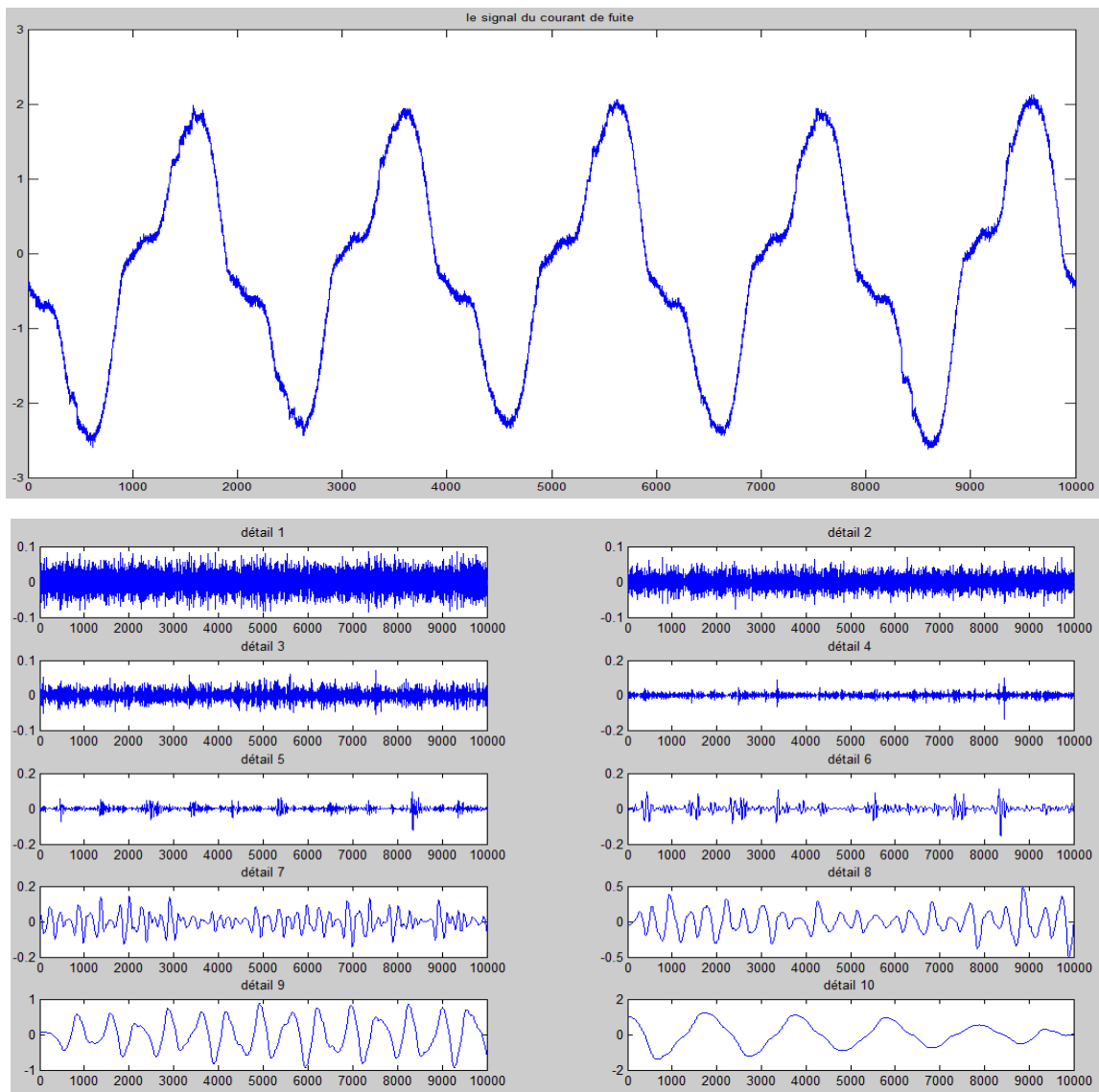


Figure I.3: Représentation de la forme d'onde du signal du courant de fuite et ses dix détails.

I.5. Calcul des caractéristiques des coefficients de détails d'ondelette :

Après la décomposition du signal du courant de fuite, nous avons calculé pour chaque détail les valeurs maximum, minimum, moyenne ainsi que l'écart type. Le choix de ces quatre paramètres semble être suffisant pour bien identifier notre signal.

Le maximum et le minimum nous donnent une information sur les valeurs crêtes des détails à chaque niveau de décomposition. L'écart type est calculée pour les détails afin d'identifier les transitions d'énergie présentes dans le signal du courant de fuite [16]. Elle est donnée par l'expression :

$$\delta = \sqrt{\frac{\sum_{i=1}^N (d_n - d_m)^2}{N-1}} \quad (\text{I.4})$$

δ : l'écart type du signal d .

d_n : correspond à un échantillon du détail d .

d_m : la moyenne du signal d .

N : correspond à la taille du signal d .

Par la suite nous allons étudier la variation de ces paramètres en fonction de la conductivité volumique pour différents paliers de tension et pour les deux types de pollution.

I.5.1. Etude de l'écart type des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :

Afin de réaliser cette étude, nous avons calculé l'écart type pour les dix détails en utilisant la formule (I.4). Après avoir obtenu les valeurs de l'écart type, Nous avons établi une représentation graphique de ces dernières en fonction de la conductivité volumique de l'agent polluant pour différents paliers de tension appliquée et pour les deux types de pollution comme le montre les figures suivantes.

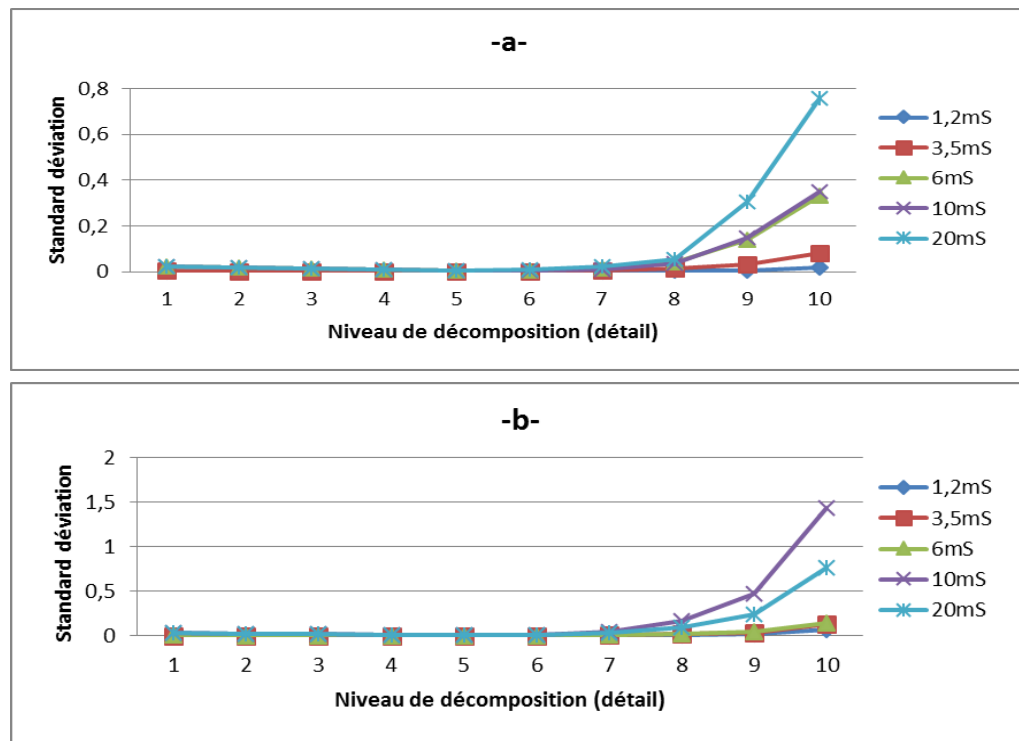


Figure I.4: Evolution de l'écart type des détails du signal du courant de fuite pour une tension appliquée de 5 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

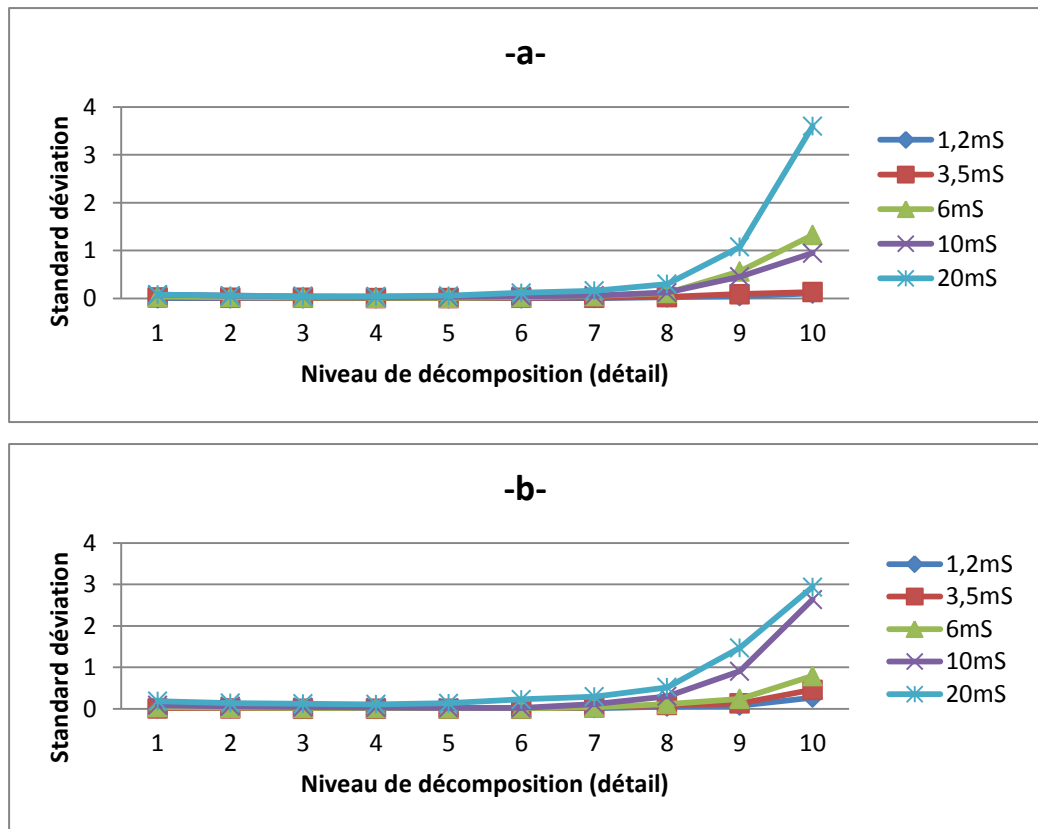


Figure 1.5: Evolution de l'écart type des détails du signal du courant de fuite pour une tension appliquée de 15 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

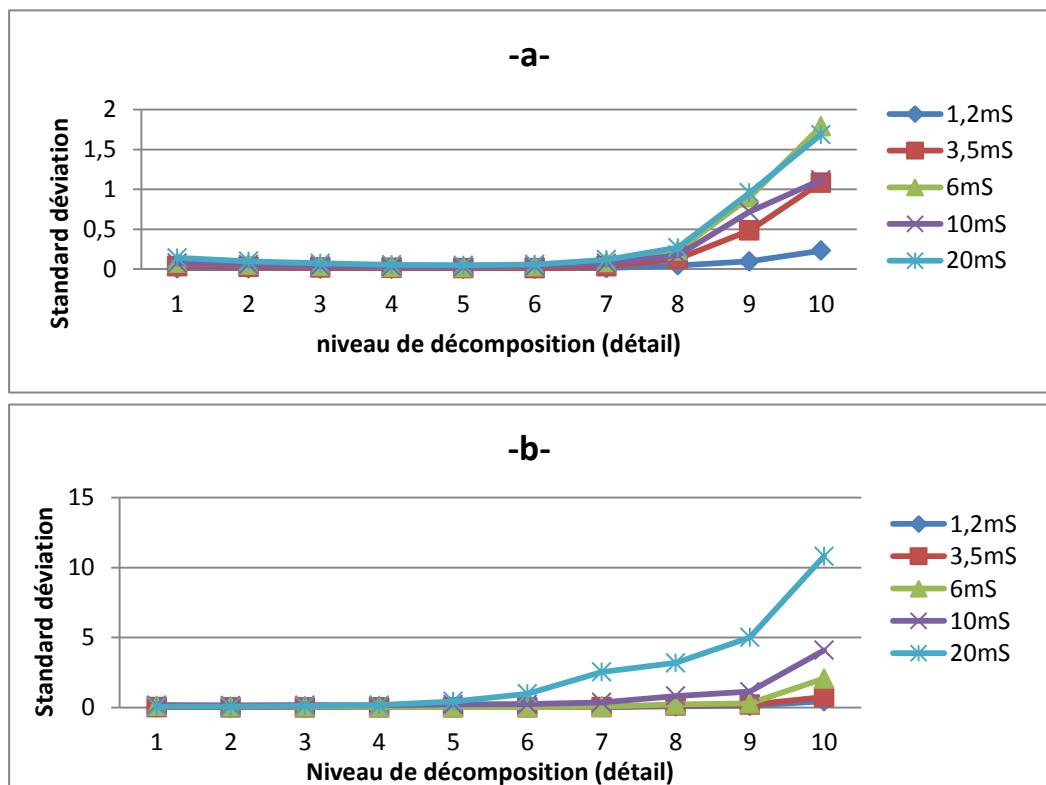


Figure 1.6: Evolution de l'écart type des détails du signal du courant de fuite pour une tension appliquée de 25 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

Interprétation :

Nous constatons que les valeurs de l'écart type augmentent avec le niveau de décomposition. Cette augmentation est vérifiée pour tous les niveaux de tension appliqués ainsi pour les deux types de pollution. L'écart type est d'autant plus important que la conductivité volumique de la solution polluante est élevée. Nous observons que l'augmentation de l'écart type est remarquable à partir du huitième niveau de décomposition pour le cas de pollution uniforme et à partir du septième niveau de décomposition pour le cas de pollution non uniforme.

Suite à cette étude nous constatons que les valeurs de l'écart type peuvent être considérées comme information importante sur le degré de sévérité de pollution car leurs valeurs sont proportionnelles à la conductivité volumique de l'agent polluant.

I.5.2. Etude des maxima des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :

Afin de réaliser cette étude, nous avons calculé les valeurs des maxima pour les dix détails. Une représentation graphique de ces dernières en fonction de la conductivité volumique de la solution polluante a été établie pour différents paliers de tension appliquée comme le montre les figures suivantes.

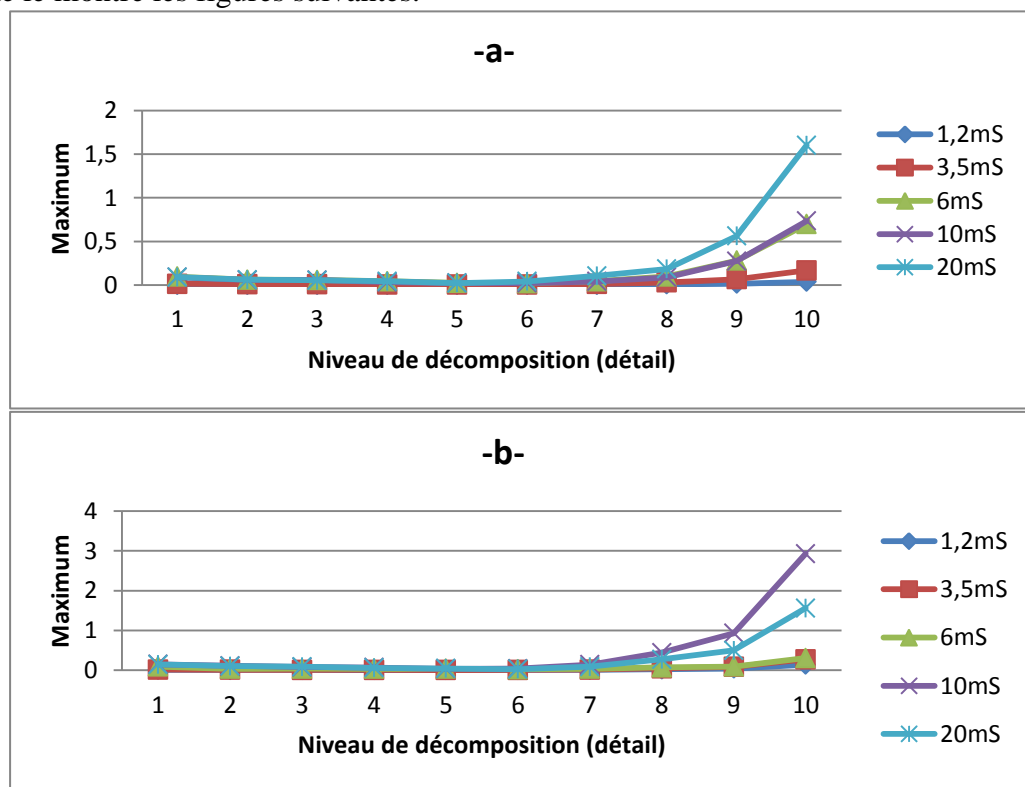


Figure I.7: Evolution des valeurs des maxima des détails du signal du courant de fuite pour une tension appliquée de 5 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

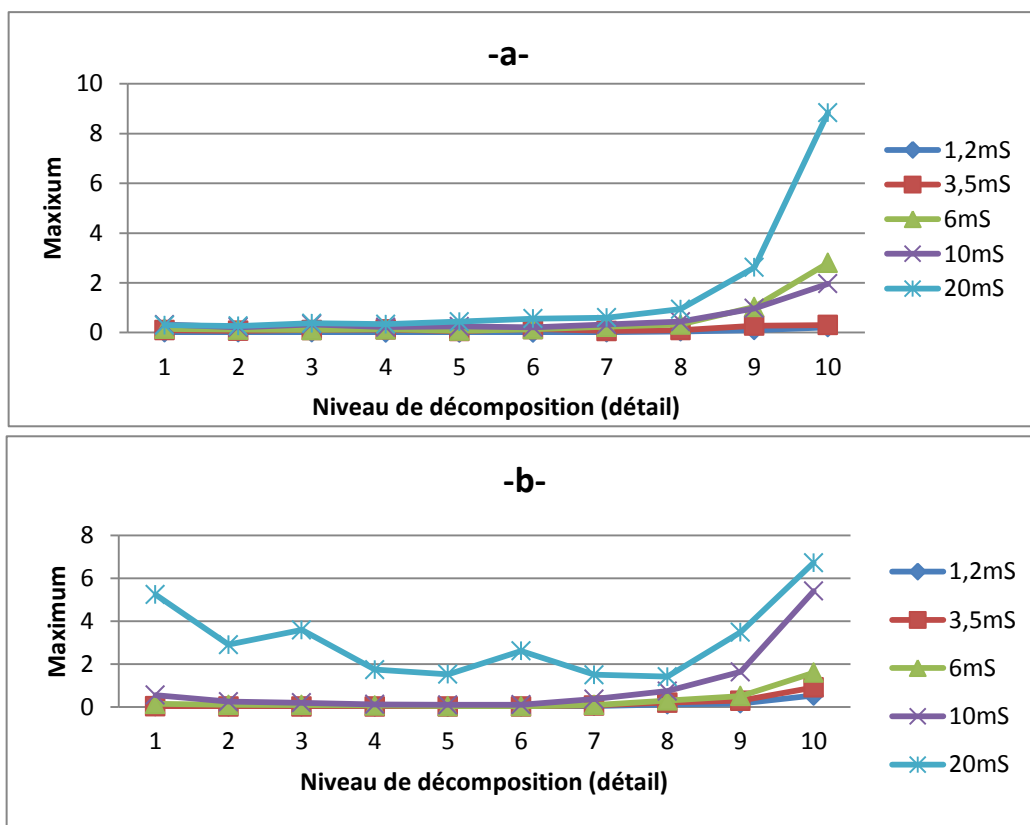


Figure I.8: Evolution des valeurs des maxima des détails du signal du courant de fuite pour une tension appliquée de 15 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

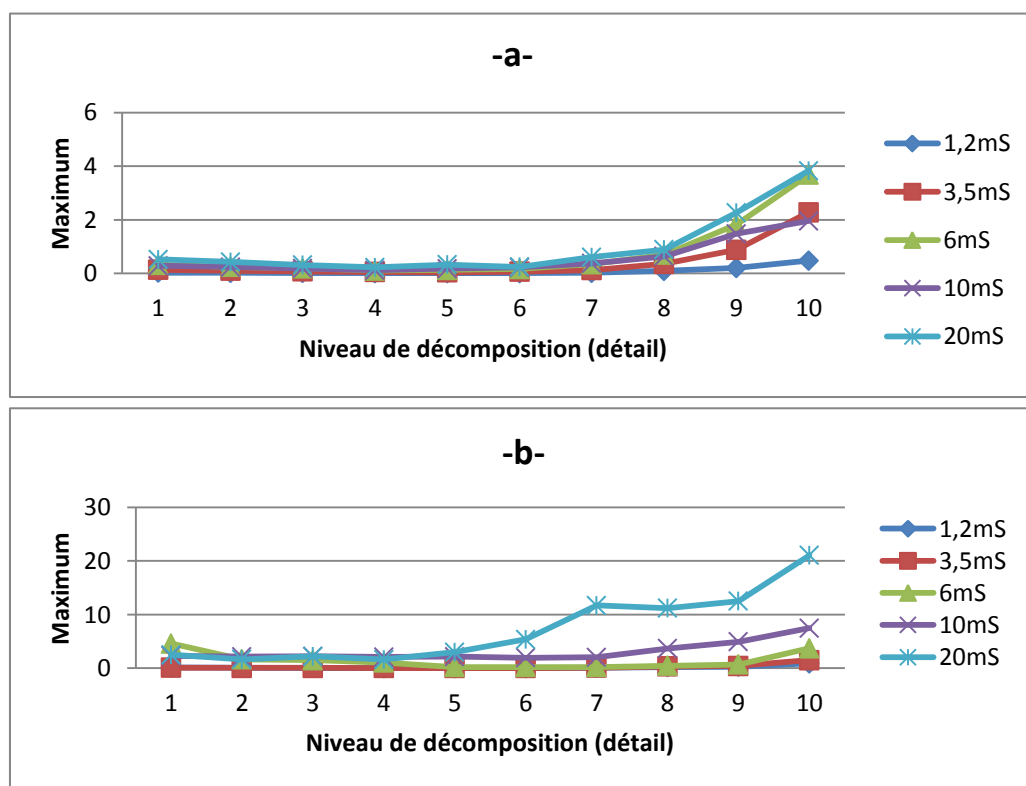


Figure I.9: Evolution des valeurs des maxima des détails du signal du courant de fuite pour une tension appliquée de 25 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

Interprétation :

Quelques soit les niveaux de tension appliquée, les valeurs des maxima augmentent avec le niveau de décomposition. Cette augmentation est remarquable à partir du huitième niveau de décomposition pour le cas de pollution uniforme et à partir du septième niveau pour le cas de pollution non uniforme. Les valeurs des maxima sont d'autant plus importantes que la conductivité volumique de l'agent polluant est élevée.

Cette étude nous permet de dire que les valeurs des maxima peuvent être considérées comme une deuxième information importante pour la prédiction de l'état de surface des isolateurs car leurs valeurs sont aussi proportionnelles à la conductivité.

I.5.3. Etude des minima des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :

Afin de réaliser cette étude, nous avons calculé les valeurs des minima pour les dix détails du signal du courant de fuite. Nous avons établi une représentation graphique de ces derniers en fonction de la conductivité volumique de la solution polluante pour trois paliers de tension comme le montre les figures suivantes.

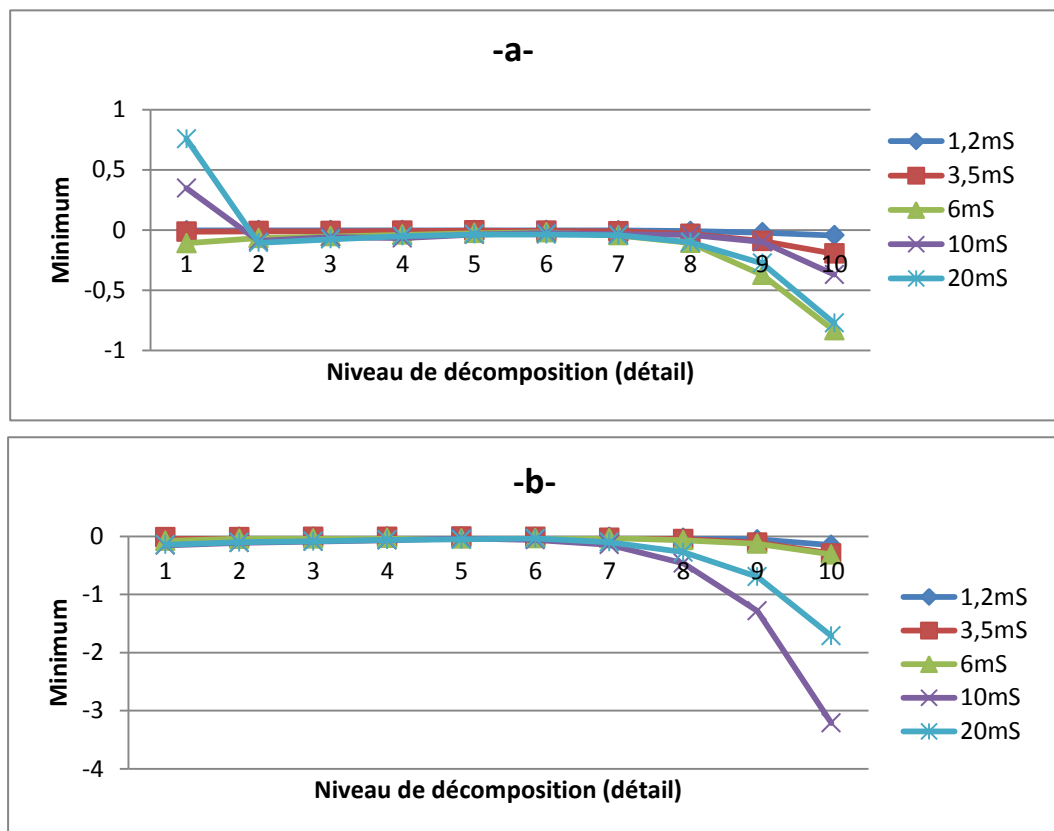


Figure I.10: Evolution des valeurs des minima des détails du signal du courant de fuite pour une tension appliquée de 5 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

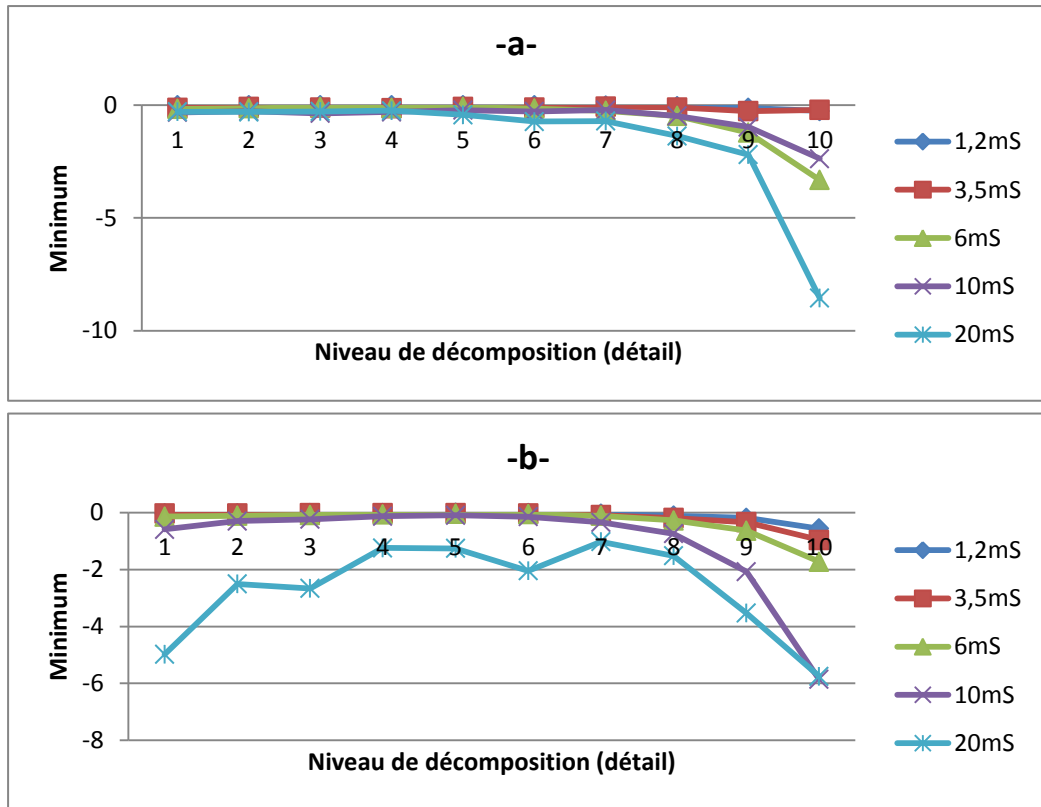


Figure I.11: Evolution des valeurs des minima des détails du signal du courant de fuite pour une tension appliquée de 15 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

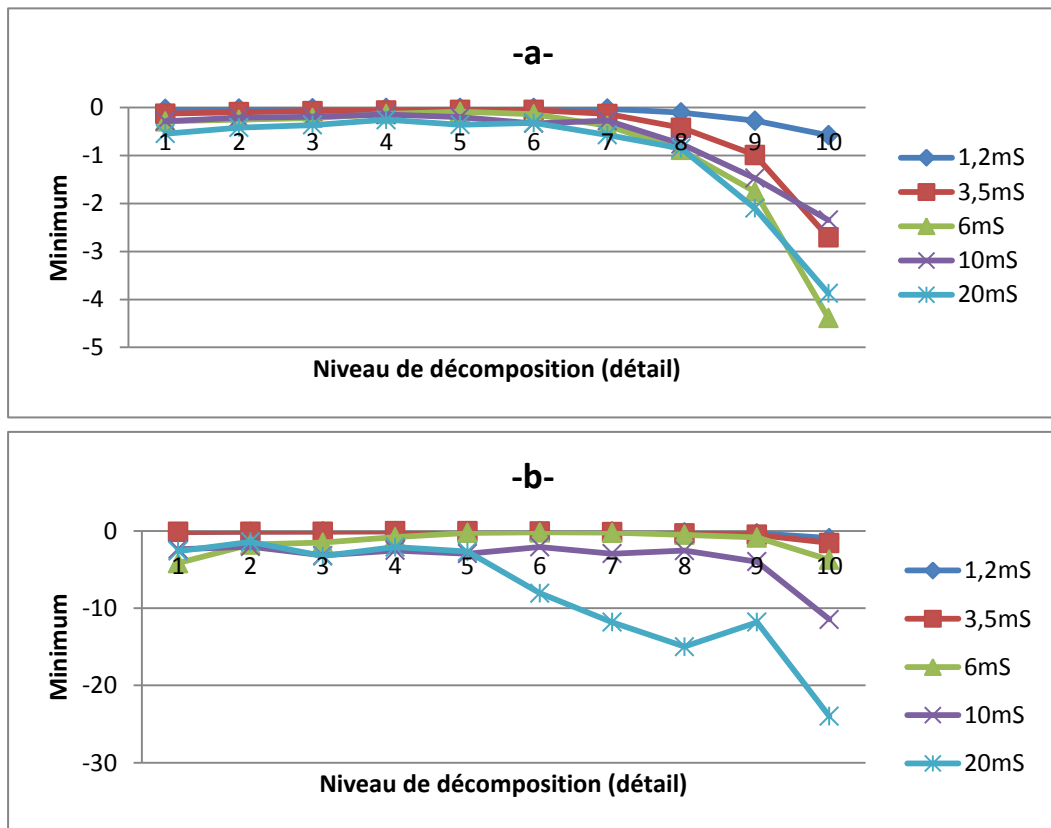


Figure I.12: Evolution des valeurs des minima des détails du signal du courant de fuite pour une tension appliquée de 25 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

Interprétation :

Les valeurs des minima diminuent avec le niveau de décomposition pour les différents niveaux de tensions et ils sont d'autant plus faibles que la conductivité de l'agent polluant est élevée.

L'étude des valeurs des minima des détails des signaux du courant de fuite nous permet de constater que les valeurs des minima peuvent être considérées comme une troisième information importante pour la prédiction de l'état de surface des isolateurs.

I.5.4. Etude des moyennes des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :

Afin de réaliser cette étude, nous avons calculé les valeurs moyennes pour les dix détails. Nous avons établi une représentation graphique de ces derniers en fonction de la conductivité volumique de la solution polluante pour trois paliers de tension appliquée comme le montre les figures suivantes.

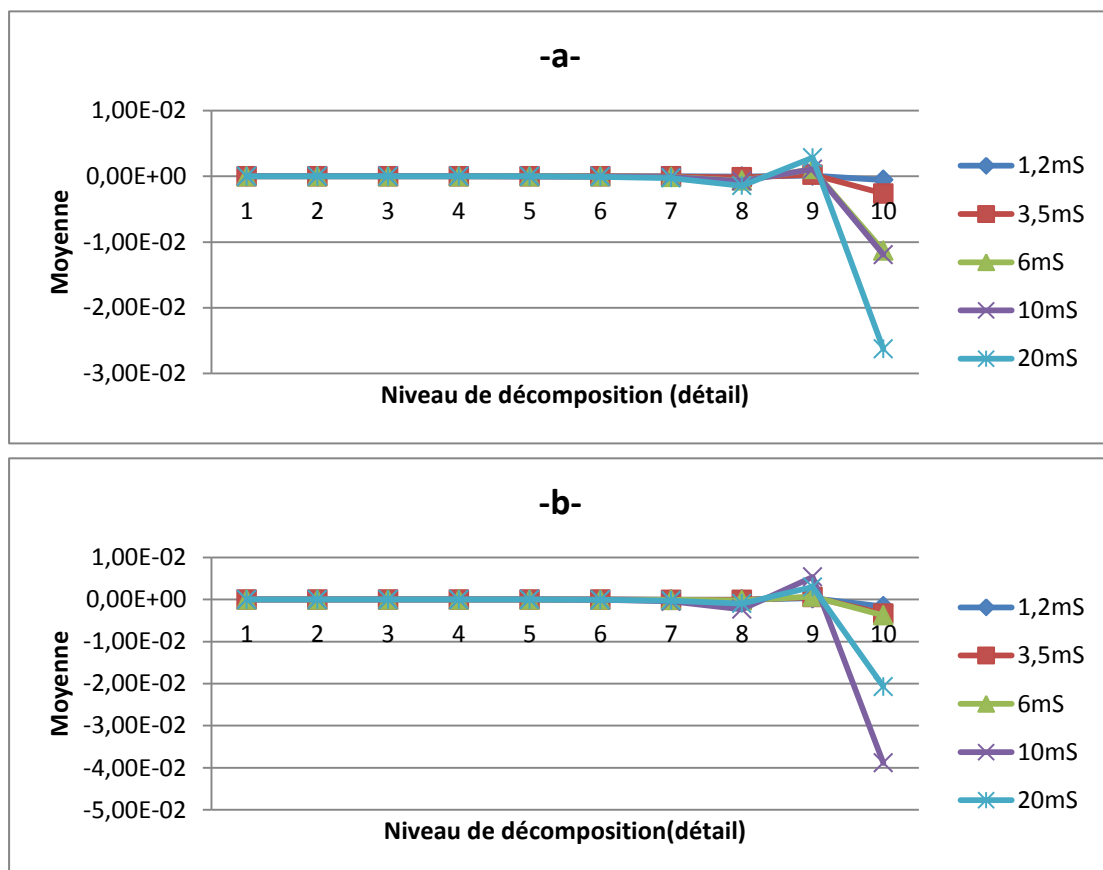


Figure I.13: Evolution des valeurs des moyennes des détails du signal du courant de fuite pour une tension appliquée de 5 kVeff. a) pollution uniforme. b) pollution non uniforme.

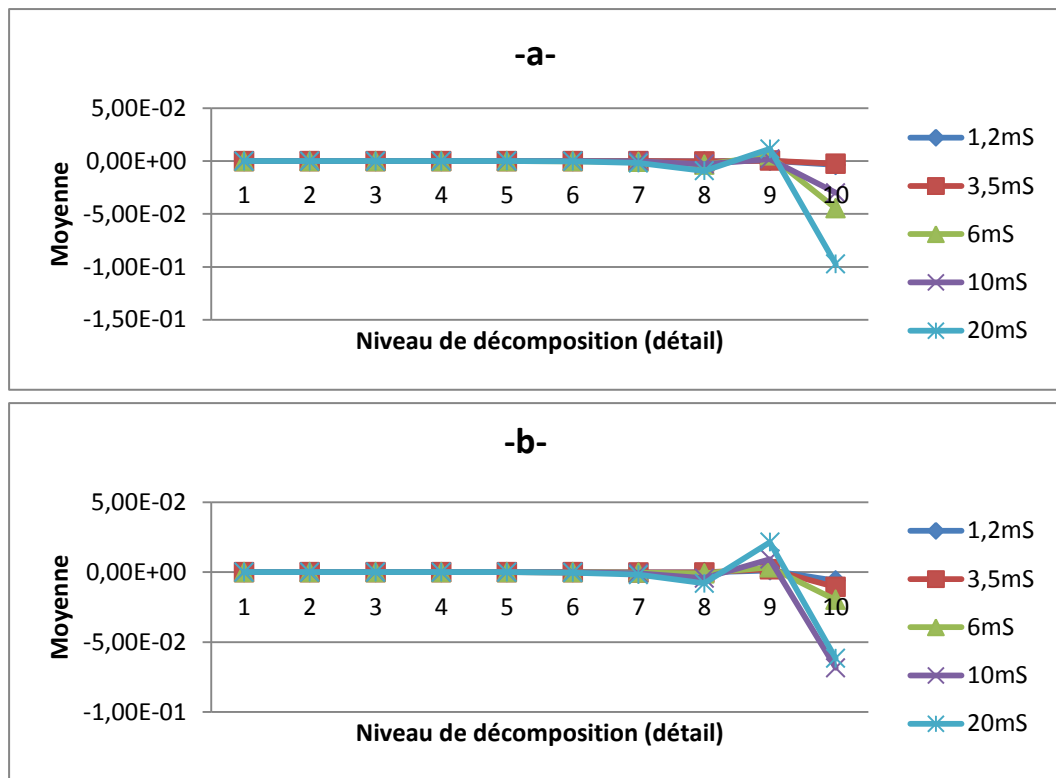


Figure I.14: Evolution des valeurs des moyennes des détails du signal du courant de fuite pour une tension appliquée de 15 kVeff. a) pollution uniforme. b) pollution non uniforme.

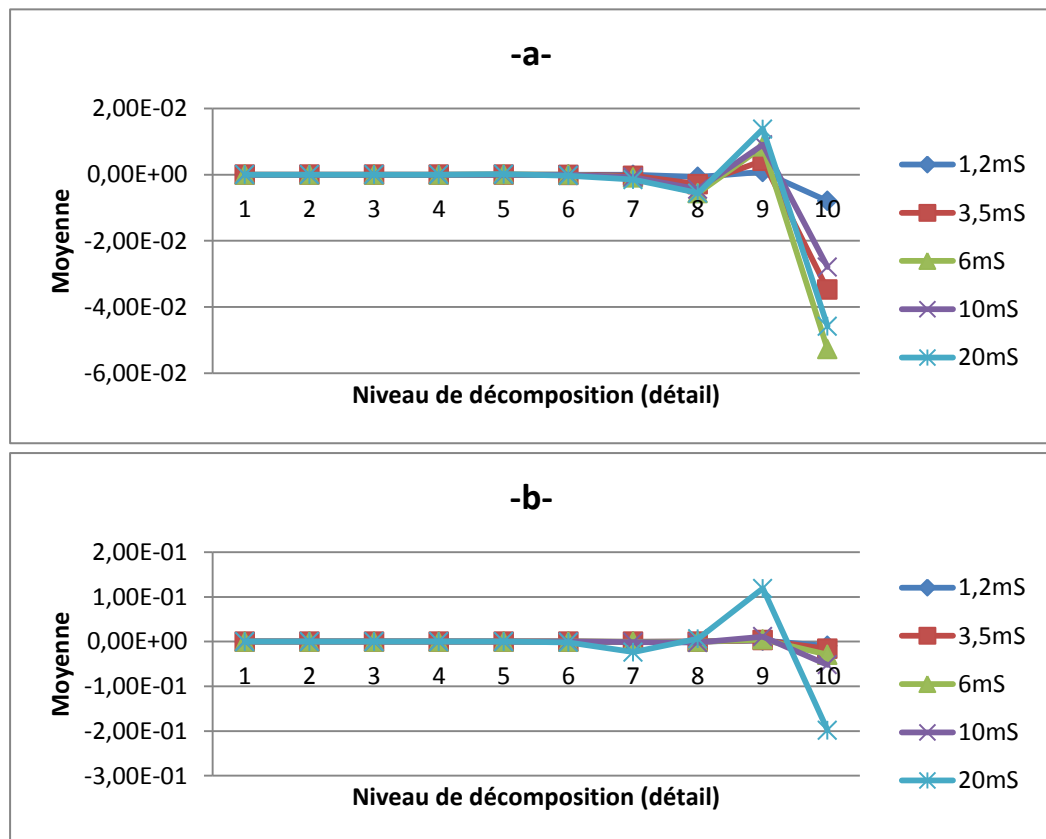


Figure I.15: Evolution des valeurs des moyennes des détails du signal du courant de fuite pour une tension appliquée de 25 kVeff. a) pollution uniforme. b) pollution non uniforme.

Interprétation :

Pour les trois niveaux de tension appliquée, nous constatons une diminution légère des valeurs des moyennes entre le premier et le huitième niveau de décomposition. Par ailleurs, nous observant un pic au niveau du neuvième détail. Les valeurs des moyennes diminuent de nouveau dans le dixième détail.

Les résultats de l'étude des valeurs des moyennes des détails des signaux du courant de fuite nous pousse à l'utilisé comme une quatrième information pour la prédiction de l'état de surface des isolateurs.

Conclusion :

Dans ce chapitre, nous avons étudié la corrélation qui pourrait exister entre le signal du courant de fuite et la sévérité de la pollution caractérisant l'état de surface de l'isolateur. La technique de décomposition par la DWT est une technique très intéressante qui nous a permis d'étudier les détails en tirant les valeurs de l'écart type, des maxima, des moyennes et des minima pour les dix niveaux de décomposition.

Les critères obtenus à l'issu de cette étude ont servis comme base de données pour l'entraînement des modèles SVM que nous avons construit.

Chapitre II

Classification de la
sévérité de pollution
des isolateurs par la
méthode SVM

Introduction:

La classification est une opération de structuration qui vise à organiser un ensemble d'observations en groupes homogènes et contrastés afin de faciliter l'analyse des informations et d'effectuer des prédictions. Le but principal des différentes méthodes de classification est l'identification des classes auxquelles appartiennent des objets à partir de certains paramètres descriptifs. Elles s'appliquent à un grand nombre d'activités humaines et conviennent en particulier au problème de la prise de décision automatisée. La procédure de classification sera extraite automatiquement à partir d'un ensemble d'exemples. Un exemple consiste en la description d'un cas avec la classification correspondante. Un système d'apprentissage doit alors, à partir de cet ensemble d'exemples, extraire une procédure de classification, il s'agit en effet d'extraire une règle générale à partir des données observées. La procédure générée devra classer correctement les exemples de l'échantillon et avoir un bon pouvoir prédictif pour classer correctement de nouvelles descriptions [17, 18].

Les méthodes utilisées pour la classification sont nombreuses: la méthode des machines à vecteurs de support (Séparateurs à Vastes Marges (SVM)), les réseaux de Neurones, etc.

Nous présentons dans la première partie de ce chapitre une étude détaillée sur la méthode SVM. Cette méthode a montré son efficacité dans plusieurs domaines d'applications tels que le traitement d'image, la catégorisation de textes, le diagnostic médicale et comme dans notre cas, dans le domaine de la Haute Tension.

Dans la seconde partie, nous allons d'abord présenter les différentes classes de pollution associées à chaque concentration en kaolin. Ensuite, nous appliquerons les deux stratégies de la méthode SVM pour le cas multi-classe (one versus one et one versus all) afin de classer la sévérité de pollution avec différentes fonctions noyaux pour les deux types de pollution en se basant sur les caractéristiques des signaux du courant de fuite.

II.1. Théorie de l'apprentissage :

L'apprentissage est défini comme l'acquisition de connaissances et compétences permettant la synthèse d'information. Un algorithme d'apprentissage va permettre de passer d'un espace d'exemples X à un espace dit des hypothèses H .

L'objectif de l'apprentissage à partir d'exemples étiquetés appelé aussi apprentissage supervisé est de construire une fonction qui permet d'approcher au mieux une fonction inconnue qui génère des données aléatoires, indépendantes et identiquement

distribuées et dont nous ne disposons que de quelques exemples [17, 19].

Un système d'apprentissage à partir d'exemples est composé de trois modules principaux :

- Un générateur qui génère des données aléatoires appelées les vecteurs d'entrée. Ces vecteurs sont indépendants et identiquement distribués suivant une distribution de probabilité inconnue $P(x)$.
- Un superviseur qui associe pour chaque vecteur d'entrée x une sortie y (la classe) suivant une distribution de probabilité également inconnue $P(x, y)$.
- Une machine d'apprentissage qui permet d'implémenter une famille de fonctions $f_{\alpha}(x)$; $\alpha \in A$, ou A est un ensemble de paramètres. Ces fonctions doivent produire pour chaque vecteur d'entrée x une sortie \tilde{y} la plus proche possible de la sortie y du superviseur [17, 19].

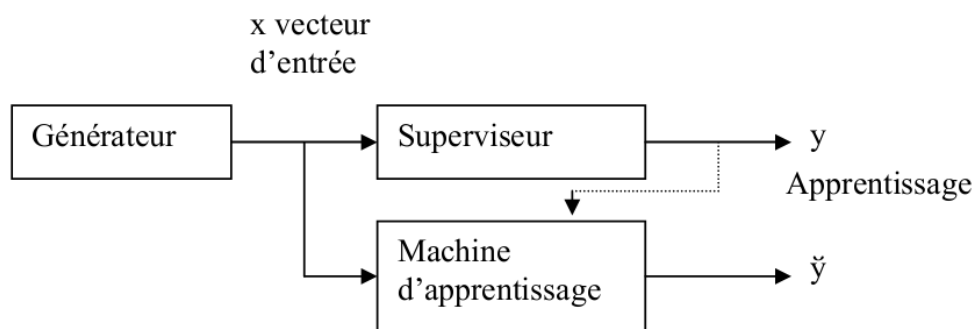


Figure II.1: les modules d'un système d'apprentissage.

Le problème d'apprentissage à partir d'exemples apparaît dans plusieurs domaines comme la prédiction, la classification, la reconnaissance de formes, la reconnaissance d'objets, etc.

Il y a deux types d'apprentissage :

- **L'apprentissage supervisé** : on dispose d'un nombre fini d'exemples d'une tâche à réaliser, sous forme de paires (entrée, sortie désirée), et on souhaite obtenir, d'une manière automatique, un système capable de trouver de façon relativement fiable la sortie correspondant à toute nouvelle entrée qui pourrait lui être présentée.
- **L'apprentissage non-supervisé** : Dans l'apprentissage non-supervisé il n'y a pas de notion de sortie désirée, on dispose seulement d'un nombre fini de données d'apprentissage, constituées "d'entrées", sans qu'aucun label n'y soit rattaché.

II.2. Les machines à vecteurs de support:

Les machines à vecteur de Support ou séparateurs à vaste marge (SVM) sont de nouvelles techniques discriminantes dans la théorie de l'apprentissage statistique supervisé. Elles ont été proposées en 1995 par V. Vapnik dans son livre « The nature of statistical learning theory » [20]. Elles permettent d'aborder plusieurs problèmes divers et variés comme la régression, la classification, la prévision... etc. Depuis leur introduction dans le domaine de la reconnaissance de formes, plusieurs travaux ont pu montrer l'efficacité de cette technique.

Les SVM sont des méthodes d'apprentissage automatique, c'est-à-dire à partir d'un échantillon d'apprentissage, on va créer un modèle applicable sur des échantillons à tester, par contre elles ne sont pas adaptées au cas d'une discrimination à plus de deux classes. Par rapport aux techniques classiques d'apprentissage, les SVM ne dépendent pas de la dimension de l'espace de représentation des données. Grâce à l'usage d'une fonction noyau, elles permettent une classification non linéaire comme nous le verrons par la suite [21, 18].

L'idée des hyperplans à marge maximale a été explorée dès 1963 par V.Vapnik et A. Lerner, et en 1973 par Richard Duda et Peter Hart dans leur livre « Pattern Classification ». Les fondations théoriques des SVM ont été explorées par V.Vapnik et ses collègues dans les années 70 avec le développement de la Théorie de Vapnik-Chervonenkis, et par Valiant [20].

Ce n'est toutefois qu'en 1992 que ces idées seront bien comprises et rassemblées par Boser, Guyon et Vapnik dans un article, qui est l'article fondateur des séparateurs à vaste marge.

II.3. Principe de la technique SVM:

Le principe des séparateurs à vaste marge (SVM) consiste à projeter les données de l'espace d'entrée (appartenant à deux classes différentes) non-linéairement séparables dans un espace de plus grande dimension appelé espace de caractéristiques de façon à ce que les données deviennent linéairement séparables. Dans cet espace, on construit un hyperplan optimal séparant les classes tel que :

-Les vecteurs appartenant aux différentes classes se trouvent de différents côtés de l'hyperplan.

-La plus petite distance entre les vecteurs et l'hyperplan (la marge) soit maximale.

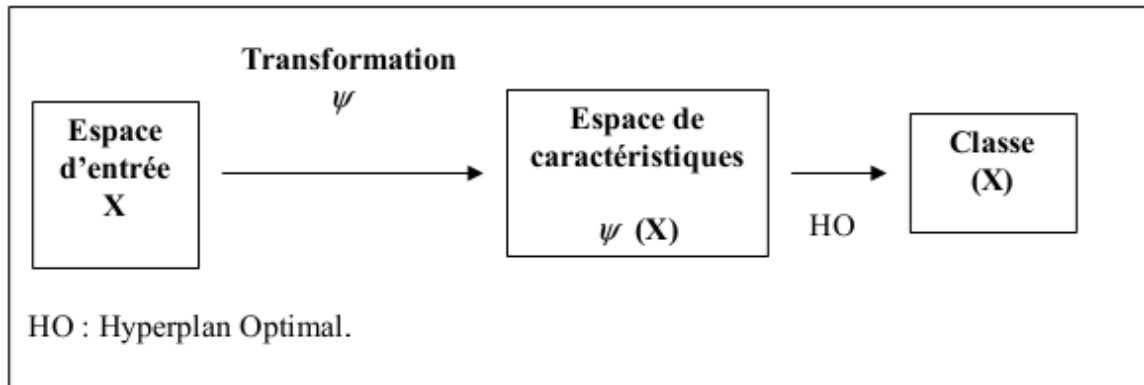


Figure II.2: principe de la technique SVM.

La méthode cherche alors l'hyperplan qui sépare les deux classes, en garantissant que la marge entre les classes soit maximale. Cela garantit une généralisation du principe car de nouveaux exemples pourront ne pas être trop similaires à ceux utilisés pour trouver l'hyperplan mais être situés d'un côté ou l'autre de la frontière. L'intérêt de cette méthode est la sélection de vecteurs supports qui représentent les vecteurs discriminant grâce auxquels est déterminé l'hyperplan. Les exemples utilisés lors de la recherche de l'hyperplan ne sont alors plus utiles et seuls ces vecteurs supports sont utilisés pour classer un nouveau cas, ce qui peut être considéré comme un avantage pour cette méthode.

II.4. Les SVM binaires:

II.4.1. Notion de base :

II.4.1.1. Hyperplan :

Plaçons-nous dans le cas d'une classification binaire (i.e. les exemples à classifier réparties en 2 classes). On appelle hyperplan séparateur un hyperplan qui sépare les deux classes figure II.6, en particulier il sépare leurs points d'apprentissage. Comme il n'est en générale pas possible d'en trouver un, on se contentera donc de chercher un hyperplan discriminant qui est une approximation au sens d'un critère à fixer (maximiser la distance entre ces deux classes) [22].

II.4.1.2. Vecteur de support :

Une tâche de détermination de l'hyperplan séparable des SVM est d'utiliser seulement les points les plus proches (i.e. les points de la frontière entre les deux classes des données) parmi l'ensemble total d'apprentissage, ces points sont appelés vecteurs de support

(Figure II.3) [22].

II.4.1.3. Marge :

Il existe une infinité d'hyperplans capables de séparer parfaitement les deux classes d'exemples. Le principe des SVM est de choisir celui qui va maximiser la distance minimale entre l'hyperplan et les exemples d'apprentissage (i.e. la distance entre l'hyperplan et les vecteurs de support), cette distance est appelée la marge (Figure II.3) [22].

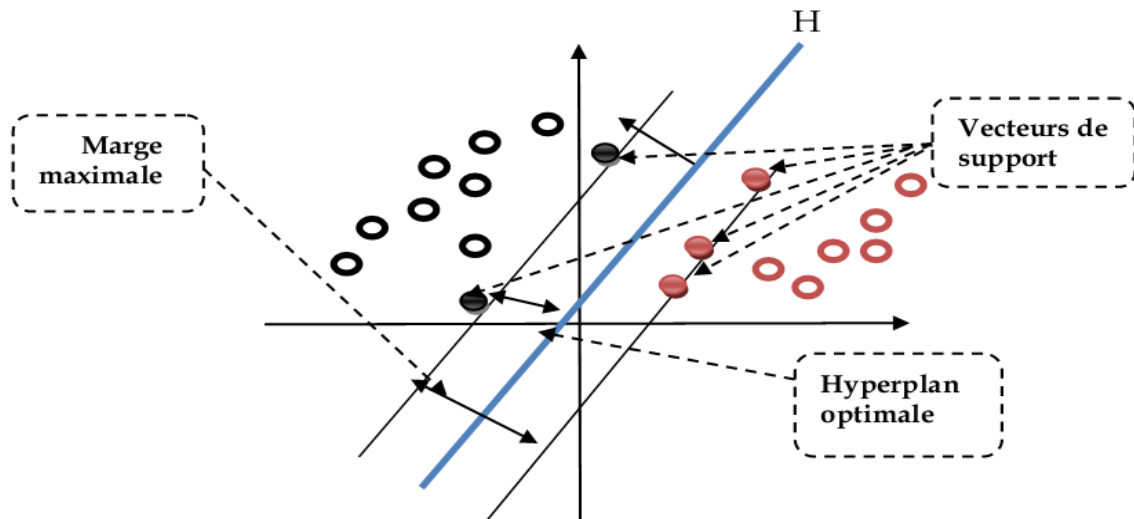


Figure II.3: hyperplan optimal, vecteurs de support et marge maximale.

II.4.2. Construction de l'hyperplan optimal:

Les machines à vecteurs de support forment une classe d'algorithmes d'apprentissage supervisés. Nous nous intéressons à une fonction notée f qui à toute entrée x fait correspondre une sortie $y = f(x)$. Le but est d'essayer d'apprendre f à partir d'un ensemble de couple $(x_i; y_i)$. Dans ce problème les machines à vecteurs support vont être utilisées pour classifier une nouvelle observation x en se limitant à deux classes $y \in \{-1, 1\}$.

II.4.2.1. Cas des données linéairement séparables:

On considère une base d'apprentissage B de m objets à deux classes linéairement séparables. Chaque objet est alors décrit par le couple $\{x_i, y_i\}$, $x_i \in \mathbb{R}^n$, $y_i \in \{-1, 1\}$ où x_i est le vecteur décrivant chaque objet dans l'espace de description de dimension n et y_i l'étiquette de la classe de cet objet [6].

Un classifieur est dit linéaire lorsqu'il est possible d'exprimer sa fonction de décision par une fonction linéaire. On peut exprimer une telle fonction par $w \in \mathbb{R}^n$ et $b \in \mathbb{R}$ tel que :

$$h(x) = \langle w, x \rangle + b = \sum_{i=1}^n w_i x_i + b \quad i \in [1 \dots m]. \quad (\text{II.1})$$

Dans cette partie, nous présentons la méthode générale de construction de l'Hyperplan Optimal (HO) qui sépare des données appartenant à deux classes différentes linéairement séparables. La figure II.9 donne une représentation visuelle de l'HO dans le cas des données linéairement séparables.

On suppose qu'il existe un hyperplan séparant les points de chaque classe, l'équation d'un tel hyperplan s'écrit alors : $\langle w, x \rangle + b = 0$ (II.2)

Où w est une normale à l'hyperplan, x est la donnée à classer, $b \in \mathbb{R}$.

Il existe une infinité d'hyperplans respectant cette équation, il faut trouver le meilleur. La notion de meilleur hyperplan est introduite en considérant deux hyperplans parallèles (H_1 et H_2) passant chacun par les points les plus proches de la frontière des classes qu'ils limitent et n'ont aucun point de la base d'apprentissage B entre eux (figure II.9). Leurs équations respectives sont :

$$H_1 : \langle w, x \rangle + b = 1 \quad \text{et} \quad H_2 : \langle w, x \rangle + b = -1 \quad (\text{II.3})$$

Pour garantir qu'il n'y a aucun point de B entre eux, tout point de B doit satisfaire l'inégalité suivante :

$$y_i \cdot (\langle w, x_i \rangle + b) \geq 1 \quad i \in [1 \dots m]. \quad (\text{II.4})$$

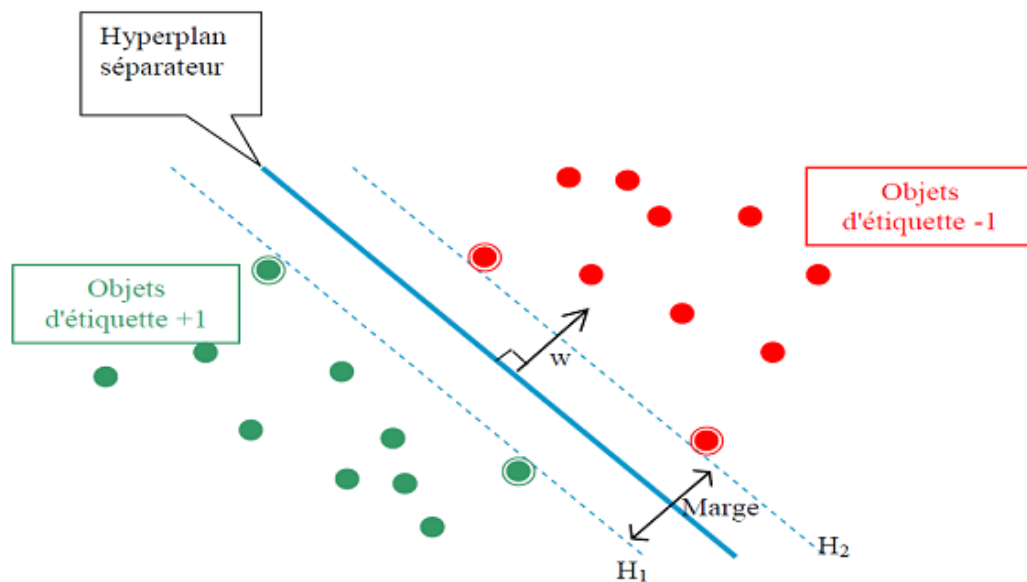


Figure II.4: les deux hyperplans qui séparent les deux classes des éléments.

Comme nous l'avons déjà mentionné, un HO est un hyperplan qui maximise la marge M qui représente la plus petite distance entre les différentes données des deux classes et

l'hyperplan. Maximiser la marge M est équivalent à maximiser la somme des distances des deux classes par rapport à l'hyperplan. Ainsi, la marge a l'expression mathématique suivante :

$$\begin{aligned} M &= \min_{x_i|y_i=1} \frac{\omega \cdot x + b}{\|\omega\|} - \max_{x_i|y_i=-1} \frac{\omega \cdot x + b}{\|\omega\|} \\ &= \frac{1}{\|\omega\|} - \frac{-1}{\|\omega\|} = \frac{2}{\|\omega\|} \end{aligned} \quad (\text{II.5})$$

Trouver l'hyperplan optimal revient donc à maximiser $\frac{2}{\|\omega\|}$. Ce qui est équivalent à minimiser $\|\omega\|$. Ceci est un problème de minimisation d'une fonction objective quadratique avec contraintes linéaires. Pour cela plusieurs méthodes ont été proposées :

Principe de Kuhn-Tucker (1951) [17]: Avec des fonctions g et C convexes, il est toujours possible de trouver un point-selle (z_0, α^*) qui vérifie :

$$\min_z L(z, \alpha^*) = L(z_0, \alpha^*) = \max_{\alpha \geq 0} L(z_0, \alpha) \quad (\text{II.6})$$

En appliquant le principe de Kuhn-Tucker, on est amené à rechercher un point-selle (w_0, b_0, α^0) . Le lagrangien correspondant à notre problème est :

$$L(w, b, \alpha) = \frac{1}{2} w \cdot w - \sum_{i=1}^m \alpha_i [y_i [(x_i \cdot w) + b] - 1] \quad (\text{II.7})$$

Le lagrangien doit être minimal par rapport à w et b et maximal par rapport à $\alpha \geq 0$.

- $L(w, b, \alpha)$ est minimal par rapport à b :

$$\frac{\partial L(w, b, \alpha)}{\partial b} = 0 \Leftrightarrow \sum_{i=1}^m \alpha_i y_i = 0 \quad (\text{II.8})$$

- $L(w, b, \alpha)$ est minimal par rapport à w :

$$\frac{\partial L(w, b, \alpha)}{\partial w} = 0 \Leftrightarrow w - \sum_{i=1}^m \alpha_i x_i y_i = 0 \quad (\text{II.9})$$

On remplace (II.16) et (II.17) dans (II.15), on aura :

$$L(w, b, \alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (\text{II.10})$$

$L(w, b, \alpha)$ est alors maximisé selon les $\alpha_i \geq 0$ en respectant les contraintes précédentes.

D'après le théorème de Kuhn-Tucker une condition nécessaire et suffisante pour que la solution α^0 soit optimal est: $\alpha_i^0 [y_i [(x_i \cdot w) + b] - 1] = 0$ pour $i = 1, \dots, m$ (II.11)

Ainsi on peut calculer w_0 et b_0 :

$$w_0 = \sum_{VS} \alpha_i^0 y_i x_i \quad (\text{II.12})$$

$$b_0 = \frac{1}{2} [(w_0 \cdot x^*(1))] + [(w_0 \cdot x^*(-1))] \quad (\text{II.13})$$

Tel que VS représente tout vecteur support x_i tel que $y_i[(x_i \cdot w) + b] = 1$

$$\text{Donc : } VS = \{x_i | \alpha_i > 0\} \quad (\text{II.14})$$

La fonction de classement **class** (x) est défini par:

$$\begin{cases} \text{class}(x) = \text{sign}[(w_0 \cdot x) + b_0] \\ = \text{sign}[\sum_{x \in VS} \alpha_i^0 y_i (x_i \cdot x) + b_0] \end{cases} \quad (\text{II.15})$$

Si **class**(x) est négative, x est de classe -1, si elle est positive, x est de classe 1 [17, 23].

II.4.2.2. Cas des données non-linéairement séparables:

Les machines à vecteurs supports présentées ci-dessus ne concernent que les modèles linéaires. Il serait complètement illusoire de se dire que l'on pourrait séparer efficacement n'importe quel jeu de données par un simple hyperplan. Si par exemple les données des deux classes se chevauchent sévèrement comme dans la figure II.5 de gauche, aucun hyperplan séparateur ne sera satisfaisant [6].

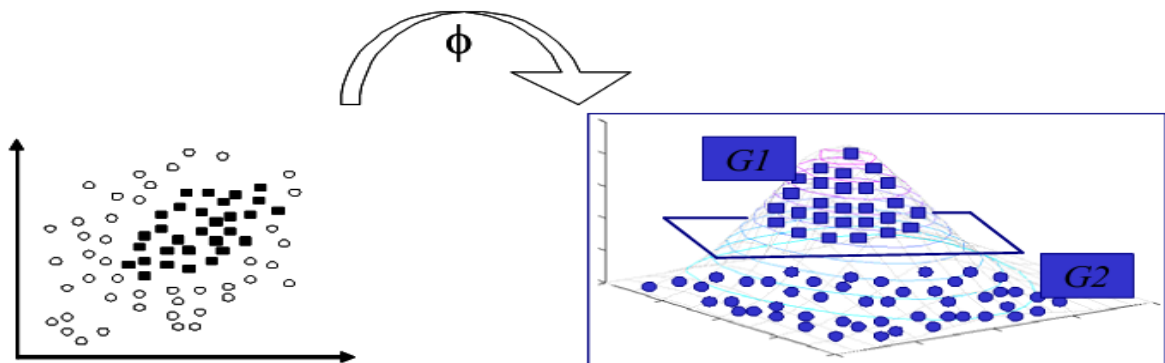


Figure II.5 : Exemple de plongement non-linéaire.

Dans ce cas où les données sont non-linéairement séparables, l'hyperplan optimal est celui qui satisfait les conditions suivantes :

- La distance entre les vecteurs bien classés et l'hyperplan optimal doit être maximale.
- La distance entre les vecteurs mal classés et l'hyperplan optimal doit être minimale.

Pour formaliser tout cela, on introduit la notion de marge souple (soft margin), qui autorise un certain nombre d'éléments d'une classe à se trouver du mauvais côté de la frontière lors de l'apprentissage (Figure II.6). Ces éléments sont appelés des «outliers» [17, 23].

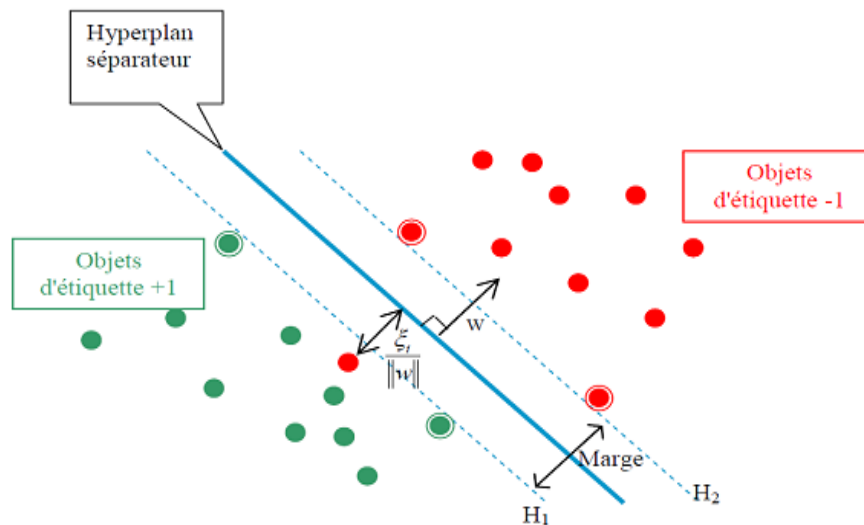


Figure II.6 : Principe de "l'outiler".

Certains éléments sont "autorisés" à se retrouver du mauvais côté de la frontière lors de la phase d'apprentissage.

On applique donc une marge d'erreur ξ_i à l'équation II.11 pour obtenir :

$$y_i \cdot (\langle w, x_i \rangle + b) \geq 1 - \xi_i \quad \text{pour } i \in [1 \dots m] \quad (\text{II.16})$$

Pour maximiser la marge, il ne suffit plus de minimiser $\|w\|$ mais la fonction (II.17) :

$$\Psi(\omega, \xi) = \frac{1}{2} \omega \cdot \omega + C \sum_{i=1}^m \xi_i \quad (\text{II.17})$$

Où C est un paramètre de régularisation. Elle permet de concéder moins d'importance aux erreurs. Cela mène à un problème dual légèrement différent de celui du cas des données linéairement séparables. Maximiser le lagrangien donné par l'équation (II.18) par rapport à α_i sous les contraintes suivantes :

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad \text{avec } 0 \leq \alpha_i \leq C \quad \text{pour } i = 1 \dots m$$

Le calcul de la fonction de classification $f(x)$ reste exactement le même que pour le cas des données linéairement séparables.

La frontière définie est là aussi linéaire mais décalée en fonction du coût fixé. L'erreur due à ce type de séparation est souvent plus importante que l'erreur générée par une frontière non linéaire, c'est pourquoi une modélisation non linéaire a été intégrée, toujours en vue de

minimiser le risque empirique. Cette méthode dite « astuce du noyau » (kernel trick) a été introduite par Aizerman en 1964 [24].

II.4.3. L'astuce du noyau (kernel trick):

Les équations de base font intervenir les données d'apprentissage (les échantillons et leur classe) sous la forme d'un produit scalaire. Il a été démontré qu'il est possible que les données initiales (de l'espace \mathbb{R}^d) soient linéairement séparables dans un espace euclidien $H = \mathbb{R}^e$ (avec $e > d$ fini ou non), dans lequel on projette l'espace de description \mathbb{R}^d à l'aide d'une fonction ϕ telle que :

$$\phi: \mathbb{R}^d \rightarrow \mathbb{R}^e, \phi(x) = \tilde{x}$$

Où x est un exemple estimé.

En effet, plus la dimension de l'espace de description est grande plus la probabilité de pouvoir trouver un hyperplan séparateur entre les classes est élevée. En transformant l'espace d'entrée en un espace de redescription H de très grande dimension, éventuellement infinie, il devient donc possible d'envisager d'utiliser la méthode de résolution quadratique précédente dans le cas de classes non linéairement séparables [19].

Si ces espaces sont munis d'un produit scalaire, on peut définir sur \mathbb{R}^d une fonction noyau telle que : $K: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}, K(x, y) = \langle \phi(x), \phi(y) \rangle = \langle \tilde{x}, \tilde{y} \rangle$ (II.18)

Mais de telle sorte qu'on n'a jamais besoin de produire explicitement les $\phi(x)$.

L'algorithme des SVM est alors le même que dans le cas linéaire : il suffit de remplacer tous les produits scalaires par une fonction noyau non linéaire. Avec cette astuce, les SVM peuvent définir le meilleur hyperplan séparateur linéaire dans le nouvel espace \mathbb{R}^e en utilisant la fonction noyau appropriée.

En appliquant la technique au noyau à la fonction de Lagrange, on obtient la fonction suivante :

$$L(\omega, b, \alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (\text{II.19})$$

Et la fonction de décision devient donc :

$$y = \text{sign} \left(\sum_{i \in \text{SV}} \alpha_i y_i K(x_i, x_j) + b \right) \quad (\text{II.20})$$

Où SV est l'ensemble des indices des vecteurs de support issus de la phase d'apprentissage.

Toute fonction satisfaisant au critère de Mercer (qui se résument essentiellement par

être définie positive) peut être utilisée comme noyau [23]. Ce critère est le suivant :

Soit K une fonction symétrique, on définit alors la matrice $(K(x_i, x_j))_{i,j}$.

Si cette matrice est définie positive alors il existe une fonction ϕ tels que :

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \quad (\text{II.21})$$

II.4.3.1. Exemples de noyaux :

Une famille de ces fonctions noyaux qui sont très appropriées aux besoins des SVM peut être définie, en voici les plus utilisés [23, 19]:

Noyau polynomial d'ordre d :

$$K(u, v) = (\langle u, v \rangle + 1)^d \quad (\text{II.22})$$

La dimension de l'espace transformé induit par un noyau polynomial est de l'ordre $\frac{(p+d)!}{p!d!}$ où p est la dimension de l'espace de départ.

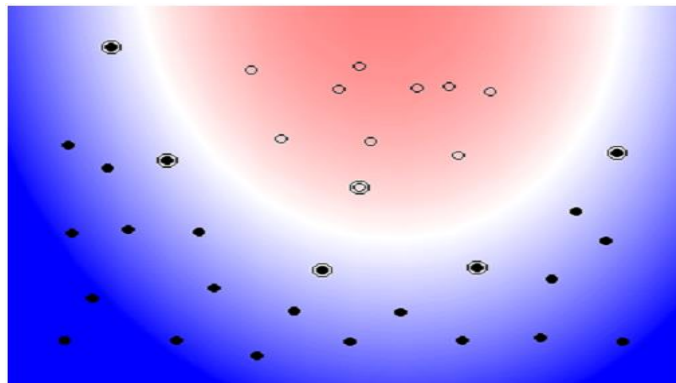


Figure II.7 : un modèle SVM avec fonction noyau polynomial d'ordre 2.

La fonction noyau sigmoïde : $K(u, v) = \tanh(u * v + c)$ (II.23)

La fonction noyau linéaire : $K(u, v) = u \cdot v$ (II.24)

La fonction noyau gaussien à Base Radiale RBF de largeur de bande σ :

$$K(u, v) = \exp\left(-\frac{\|u-v\|^2}{2\sigma^2}\right) \quad (\text{II.25})$$

Le paramètre σ permet de régler la largeur de la gaussienne. En prenant un σ grand, la similarité d'un exemple par rapport à ceux qui l'entourent sera assez élevée, alors qu'on prenant un σ tendant vers 0, l'exemple ne sera similaire à aucun autre.

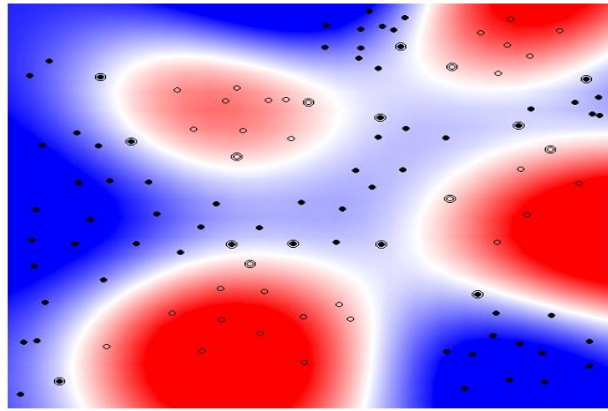


Figure II.8 : un modèle SVM avec la fonction noyau RBF.

II.5. SVM pour le cas multi-classes:

Les SVM multi-classes sont des modèles de l'apprentissage de conception relativement récente, dont l'étude est actuellement en plein essor. Cette situation nouvelle fait naître un besoin, celui de disposer d'une étude synthétique sur les SVM multi-classes, ou plus généralement l'utilisation de SVM pour la discrimination à catégories multiples.

Il existe deux stratégies pour faire la classification multi-classes :

II.5. La stratégie un contre tous:

L'architecture la plus ancienne, utilisée pour les machines à vecteurs supports multi-classes, est probablement la méthode un contre tous. Chaque SVM de la stratégie «un contre tous» étant entraîné à séparer les données d'une classe de celles de toutes les autres classes, chaque sortie est directement liée à la probabilité d'appartenance à l'une des classes [25]. Le problème consiste donc à calibrer les sorties de l'ensemble des SVM. Cette méthode construit m classifieurs binaires à vecteurs de supports où m est le nombre total des classes. L'apprentissage du $k^{\text{ème}}$ classifieur à vecteurs de supports s'effectue en considérant tous les exemples de la $k^{\text{ème}}$ classe dans la région positive et tous les autres exemples dans la région négative. Le $k^{\text{ème}}$ classifieur à vecteurs supports s'obtient en résolvant le problème

$$\text{Minimiser } \frac{\|w_k\|^2}{2} + C \sum_{j=1}^l \xi_j^k$$

$$\begin{aligned} \text{sous } & \langle w_k \cdot \phi(x_j) \rangle + b_k \geq 1 - \xi_j^k & \text{si } y_i = k, \\ & \langle w_k \cdot \phi(x_j) \rangle + b_k \leq -1 + \xi_j^k & \text{à si } y_i \neq k, \\ & \xi_j^k \geq 0 \quad j = 1, 2, \dots, \dots, l. \end{aligned}$$

Où les $\phi(x_s)$ sont les transformées des x_s dans l'espace induit par la fonction ϕ et C est le paramètre de pénalité.

Dans le cas multi-classes ($m > 2$), cette égalité peut être satisfaite pour plus que deux classes. Dans ce cas, l'observation x est dite non-classifiable. Toutes les observations x non-classifiables forment la région d'ambiguïté dite aussi région non-classifiable. Cette région est schématisée dans la Figure II.14.

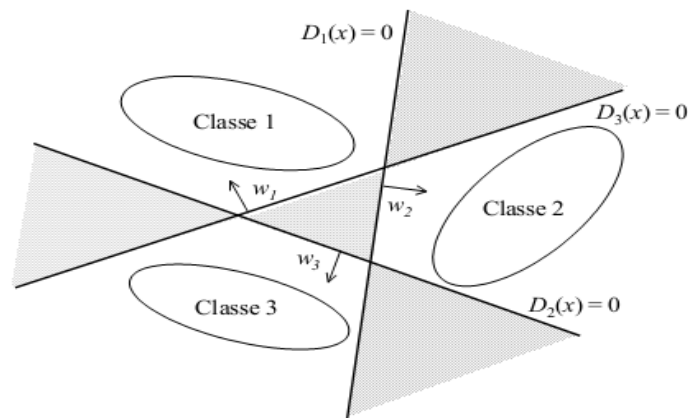


Figure II.9: L'espace hachuré représente la région d'ambiguïté pour l'approche « un contre tous » suite à la prise de décision discrète.

Afin de pouvoir classer une observation x tombant dans la région d'ambiguïté, la règle de décision continue a été utilisée. Cette règle est donnée par :

$$k^* = \text{Arg} \max_{1 \leq k \leq m} f_k(x) \quad (\text{II.26})$$

Géométriquement interprétée, tout nouvel exemple x est affecté à la classe qui correspond à l'hyperplan le plus éloigné. Ainsi, l'espace des variables explicatives X est subdivisé en m régions convexes, chacune correspondant à une classe. La figure II.14 donne un exemple de subdivision de l'espace X [25].

Cette approche est nommée “ le gagnant emporte le tout”. L'inconvénient majeur de cette heuristique est qu'elle ne conserve pas les m frontières de séparation.

II.5.2. La stratégie un contre un:

Cette stratégie de décomposition a été adoptée par S. Knerret et al. [26] et utilisée pour la première fois dans le contexte des machines à vecteurs supports par Krebel [27]. Il consiste à construire $\frac{m(m-1)}{2}$ prédicteurs binaires chacun séparant uniquement deux classes tout en ignorant les autres. Chaque classifieur est donc entraîné à séparer les données d'une classe de celles d'une autre classe. Cette stratégie s'appelle « un contre un » ou en anglais « One

Against One (OAO) .

L'hyperplan séparateur des classes k et s est la solution du problème d'optimisation suivant :

$$\begin{aligned} \text{Minimiser} \quad & \frac{\|w^{ks}\|^2}{2} + C \sum_{j=1}^l \xi_t^{ks} \\ \text{sous} \quad & \langle w^{ks}, \phi(x_t) \rangle + b^{ks} \geq 1 - \xi_t^{ks} \quad \text{si } y_t = k, \\ & \langle w^{ks}, \phi(x_t) \rangle + b^{ks} \leq -1 + \xi_t^{ks} \quad \text{si } y_t = s, \\ & \xi_t^{ks} \geq 0 \quad \forall t \in \{1, \dots, l_{ks}\} \end{aligned}$$

Où l_{ks} est le nombre des observations issues des classes k et s .

Pratiquement, nous résolvons le problème dual correspondant au problème (II.39) ayant l_{ks} variables duales. Si chaque classe contient en moyenne $\frac{l}{m}$ exemples, nous aurons à résoudre dans la phase d'apprentissage $\frac{m(m-1)}{2}$ problèmes quadratiques chacun dépendant à peu près de $\frac{2l}{m}$ variables [25].

L'approche un-contre-un consiste donc à construire un classifieur pour chaque paire de classes (k, s) définissant ainsi des fonctions de décision binaire $h_{ks} : X \subseteq \mathbb{R}^p \rightarrow \{-1, +1\}$

$$h_{ks}(x) = \text{sign}(f_{ks}(x)) = \begin{cases} +1 & \text{si } x \in \text{à la classe } k \\ -1 & \text{si } x \in \text{à la classe } s \end{cases} \quad (\text{II.27})$$

Pour des raisons de symétrie $h_{ks} \equiv -h_{sk}$ et on convient que $h_{kk} \equiv 0$ pour tout $k \in \{1, 2, \dots, m\}$. Sur la base des $\frac{m(m-1)}{2}$ fonctions de décision binaires h_{ks} , nous définissons m autres fonctions de décision de la façon suivante :

$$h_k(x) = \sum_{s=1}^m h_{ks}(x) \quad , k = 1, 2, \dots, m \quad (\text{II.28})$$

Et la règle de classification d'une nouvelle observation x est donnée par :

$$k^* = \text{Arg} \max_{1 \leq k \leq m} h_k(x) \quad (\text{II.41})$$

Cette règle connue sous le nom de vote majoritaire a été appliquée pour la première fois avec les SVM par Krebel [38].

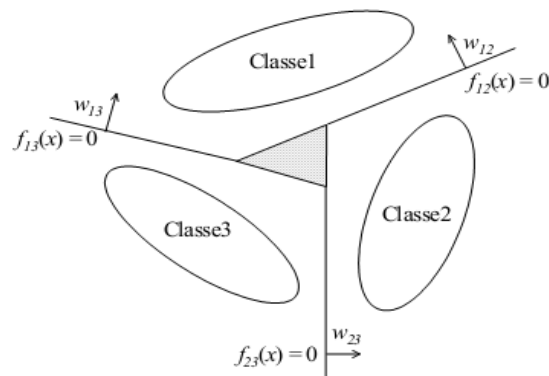


Figure II.10: La région d'ambiguïté hachurée est réduite pour l'approche un-contre-un.

II.6. Avantages et inconvénients des SVM:

L'avantage de la création d'une fonction de décision avec l'algorithme des SVM est que la solution produite correspond à l'optimum d'une fonction convexe. Elle ne possède donc pas plusieurs optima locaux comme pour les réseaux de neurones (dans leur formulation classique), mais un optimum global. Cet optimum correspond à la minimisation du risque structurel et donc à la recherche d'une hypothèse possédant de bonnes capacités de généralisation à partir d'un espace d'hypothèses donné. L'espace d'hypothèses dépend du choix de la fonction noyau, ce qui est à la fois un avantage et un inconvénient. Un avantage car l'algorithme d'apprentissage reste identique. Le problème est que la fonction noyau a souvent des paramètres libres, comme la largeur de bande avec un noyau gaussien, et que la recherche de la valeur optimale de ces paramètres libres ne correspond plus à la recherche du minimum d'une fonction convexe [28].

La recherche des valeurs idéales de ces paramètres, appelée communément sélection de modèles, est critique. Illustrons cet état critique avec l'utilisation d'un noyau gaussien et une largeur de bande σ très faible. La fonction de décision produite, pour ces conditions, sera cohérente avec les données et elle aura autant de vecteurs de support que d'exemples. L'hyperplan sélectionné, bien que de marge géométrique maximum, a une marge faible. Les bornes supérieures vues précédemment sont très élevées et elles ne garantissent plus rien sur le pouvoir de généralisation. Donc la sélection d'une mauvaise fonction noyau (ou de ces paramètres) peut conduire à produire un effet de sur-apprentissage et c'est le principal désavantage des SVM [28].

Un autre désavantage des SVM est la durée importante de la phase d'entraînement, en particulier si une sélection de modèles doit être réalisée, bien que dans les dix dernières

années plusieurs optimisations de l'algorithme initial aient été proposées.

Un dernier désavantage est la complexité de la fonction de décision produite lorsque la base d'apprentissage est de taille importante, même si, à la différence de la règle PPV, l'ensemble des exemples n'est pas utilisé grâce à la notion de vecteur de support [28].

II.7. Les Classes de pollution :

Dans notre étude, la classification de la sévérité de pollution est un point important. Dans la littérature, il est noté que cette sévérité est généralement caractérisée par l'intensité de précipitation ainsi que par la conductivité des dépôts polluants. La sévérité considérée est alors exprimée en termes de salinité équivalente. En effet, on peut faire correspondre à chaque site, soumis indifféremment à une pollution naturelle ou industrielle, une solution saline équivalente. Pour notre cas, nous distinguons quatre classes de sévérité [29], dont on va utiliser trois d'entre elles, à savoir :

Classe1 : pollution faible d'origine naturelle (Sauf marine). Elle concerne généralement des zones non industrielles ayant une faible densité d'habitation, situées loin de la mer. Un isolement normal convient facilement pour ces zones. La salinité équivalente maximale retenue pour cette classe correspond à une pulvérisation d'une solution de chlorure de sodium de concentration 2.5g/l [29].

Classe2 : pollution moyenne d'origine naturelle (Sauf marine) ou faiblement industrielle. Elle correspond à des zones à moyenne densité d'habitation et situées loin de la mer. Ces zones nécessitent un isolement renforcé, la concentration de la salinité équivalente maximale retenue est de 10g/l [29].

Classe3 : pollution forte d'origine marine et industrielle. Les zones considérées possèdent une forte densité d'habitation et située près de la mer. Elles sont exposées à des embruns marins relativement forts. Ces sites nécessitent un isolement important, la concentration de la salinité équivalente maximale est de 80g/l [29].

En tenant compte de la base de données de B. Moula [15] et des normes définies ci-dessus, nous proposons un tableau (Tableau II.1) récapitulatif des classes assignées aux différentes conductivités utilisées expérimentalement.

Concentration en Kaolin [g/l]	Conductivité volumique [mS/cm]	Concentration de sel utilisé [g/l]	Classe
40	1,2 et 3,5	Entre 0.7 et 2.5	1
	6	Entre 3 et 10	2
	10 et 20	Inférieur à 60	3
100	1,2 et 3,5	Entre 0.7 et 2.5	1
	6	Entre 3 et 10	2
	10 et 20	Inférieur à 60	3
200	1,2 et 3,5	Entre 0.7 et 2.5	1
	6	Entre 3 et 10	2
	10 et 20	Inférieur à 60	3

Tableau II.1: Tableau récapitulatif des différentes classes assignées aux différentes conductivités utilisées.

II.8. Application des SVM pour la classification de sévérité de pollution des isolateurs :

II.8.1. Composition du vecteur d'état :

Afin de juger l'état de surface d'un isolateur, nous avons besoin d'un certain nombre de critères le concernant sur lesquels nous allons nous baser. Dans le paragraphe précédent, nous avons cité les différents critères de classification et nous avons établi des liens entre ces derniers et la sévérité de pollution.

Le vecteur d'état qu'on a utilisé pour la formation du modèle est $V (1 \times 40)$, ce dernier est composé des coefficients extraits des coefficients liés aux détails d'ondelettes. On a extrait dix détails pour chaque signal, par la suite on a tiré les écarts types, les minima, les maxima et les moyennes de ces dix détails pour construire le vecteur d'état suivant :

Std1 _dstd10 _d	Min1 _dmin10 _d	Max1 _dmax10 _d	Moy1 _dmoy10 _d
---	---	---	---

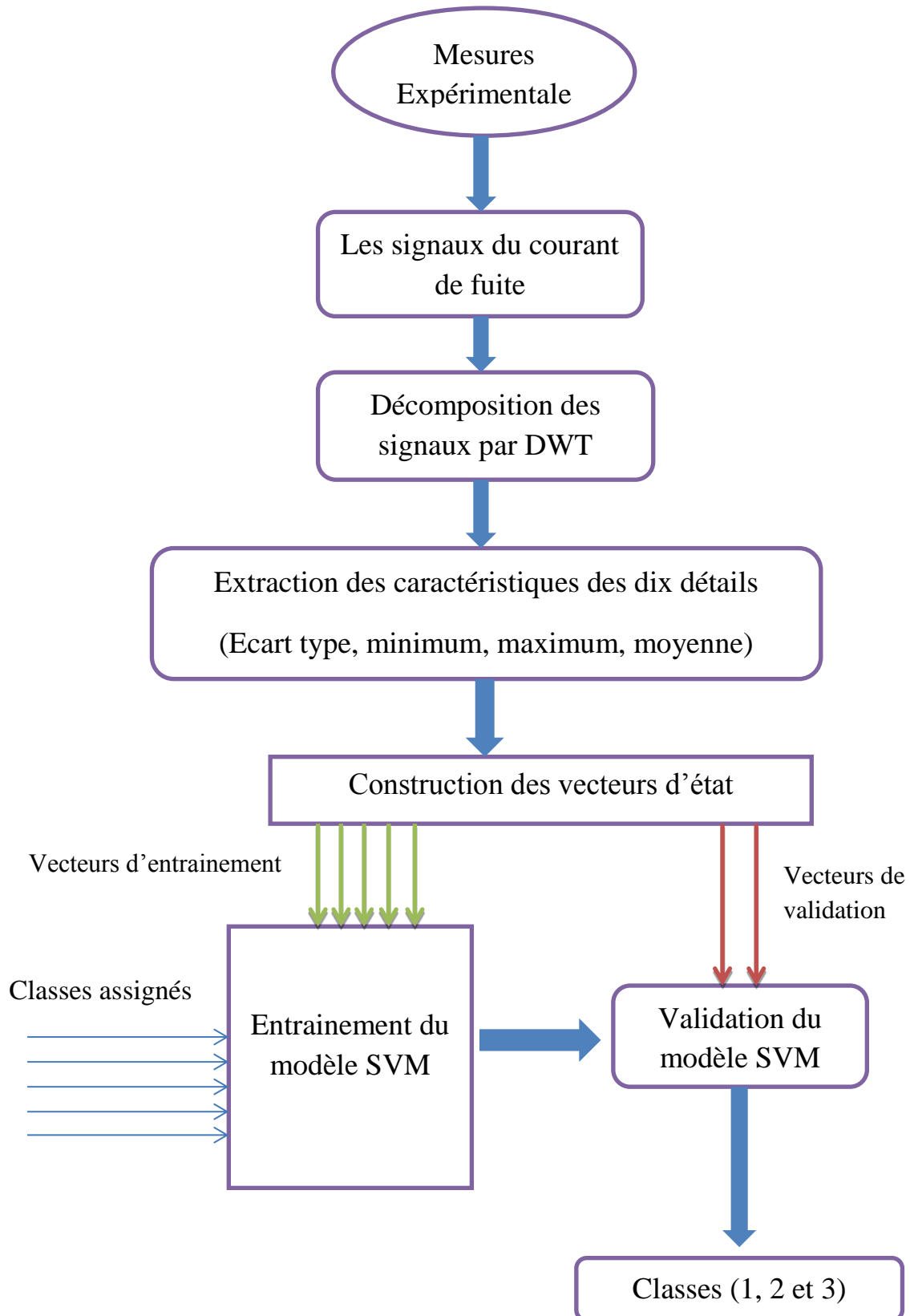


Figure II.11: La procédure utilisée pour la classification.

II.8.2. Entraînement du modèle :

II.8.2.1. Base de données pour le cas de la pollution uniforme :

Après les mesures expérimentales sur les isolateurs avec la pollution uniforme, chaque signal du courant de fuite, est traité par la technique de transformée en ondelettes discrète (DWT) pour extraire différents critères de jugement détaillés précédemment et pour chaque signal du courant de fuite nous faisons correspondre un vecteur V (1×40).

Nous répétons ce processus pour 106 signaux expérimentaux -base de données de B. Moula [58] pour obtenir une matrice de (106×40), dont 95 échantillons servent de base de données d'entraînements pour le modèle à l'aide du logiciel MATLAB®, les 11 échantillons restants servent de base de données de validation du modèle.

II.8.2.2. Base de données pour le cas de la pollution non uniforme :

Après les mesures expérimentales sur les isolateurs avec la pollution non uniforme, chaque signal du courant de fuite est traité de la même manière que pour les signaux du courant de fuite en cas de pollution uniforme.

Parmi les 75 signaux de la base de données, 66 sont utilisés pour l'entraînement du modèle SVM correspondant au type de pollution non uniforme ; et les 9 qui restent sont utilisés pour la validation.

Remarque :

La base de données dispose de plusieurs paliers de tension appliquée, dans notre travail nous avons choisi 20, 21, 25 et 28 kV comme tension des signaux servant dans la construction et la validation du modèle SVM.

La base de données dispose de deux types de signaux, ceux du courant de fuite et ceux de la tension. Dans notre modèle SVM nous avons utilisé les signaux du courant de fuite afin d'avoir de bons résultats.

II.8.3. Les différentes fonctions noyaux appliquées :

La fonction noyau linéaire : $K(u, v) = u \cdot v$

Noyau polynomial d'ordre d : $K(u, v) = (< u, v > + 1)^d$, dans notre cas on a pris deux valeurs : $d=2$ donc l'expression devient : $K(u, v) = (< u, v > + 1)^2$

Et $d=3$ donc : $K(u, v) = (< u, v > + 1)^3$

La fonction à Base Radiale RBF : $K(u, v) = \exp\left(-\frac{\|u-v\|^2}{2\sigma^2}\right)$

Dans notre étude on a pris $\sigma = 1$.

La fonction noyau sigmoïde : $K(u, v) = \tanh(u, v + c)$, dans notre étude on a pris $c = -1$ et $c = -2$ dont les fonctions noyau sont : $K(u, v) = \tanh(u * v - 1)$

Et $K(u, v) = \tanh(u * v - 2)$

II.9. Validation et résultat de classification :

Les résultats de classification sont représentés dans les tableaux suivants :

II.9.1. Cas de pollution uniforme :

Echantillons (n°)	Conductivité (mS/cm)	Classe assignée	Résultats obtenus avec différentes les fonctions noyaux					
			Linéaire	RBF	Polynomial		Sigmoïde	
					d=2	d=3	c=-1	c=-2
1	1.2	1	1	3*	1	1	2*	1
2	1.2	1	1	1	1	1	1	3*
3	3.5	1	1	3*	2*	2*	3*	1
4	3.5	1	3*	3*	3*	1	1	2*
5	6	2	3*	2	2	2	3*	3*
6	6	2	2	3*	2	2	2	3*
7	6	2	2	3*	3*	2	2	3*
8	10	3	1*	3	2*	2*	2*	3
9	10	3	2*	3	2*	3	1*	2*
10	20	3	3	3	2*	3	1*	3
11	20	3	2*	3	3	3	2*	1*
Exactitude de classification (%)			54.54	54.54	45.45	81.81	36.36	36.36

Tableau II.2: Classification SVM multi-classes en utilisant la stratégie « un contre tous ».

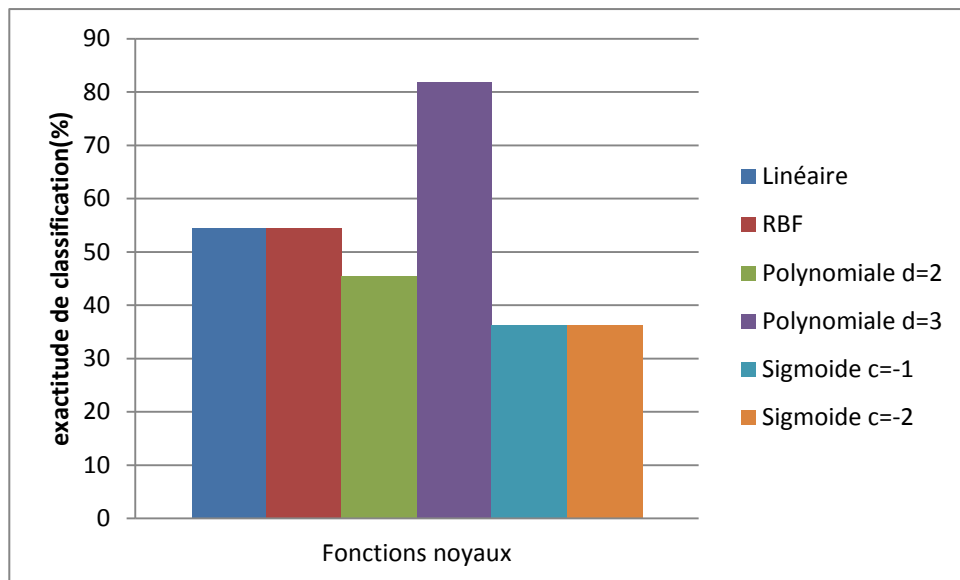


Figure II.12: représentation de l'exactitude de classification des différentes fonctions noyaux avec la stratégie *un contre tous* pour le cas de *pollution uniforme*.

Echantillons (n°)	Conductivité (mS/cm)	Classe assignée	Résultats obtenus avec les différentes fonctions noyaux					
			Linéaire	RBF	Polynomial		Sigmoïde	
					d=2	d=3	c=-1	c=-2
1	1.2	1	1	3*	2*	1	1	3*
2	1.2	1	1	1	1	1	1	1
3	3.5	1	1	3*	2*	2*	1	1
4	3.5	1	3*	3*	1	1	3*	1
5	6	2	2	2	2	2	1*	3*
6	6	2	3*	3*	2	3*	2	2
7	6	2	3*	3*	2	3*	2	2
8	10	3	2*	3	3	3	1*	3
9	10	3	2*	3	3	3	2*	2*
10	20	3	2*	3	2*	3	2*	3
11	20	3	2*	3	3	3	3	2*
Exactitude de classification (%)			36.36	54.54	72.72	72.72	54.54	63.63

Tableau II.3: Classification SVM multi-classes en utilisant la stratégie « un contre un ».

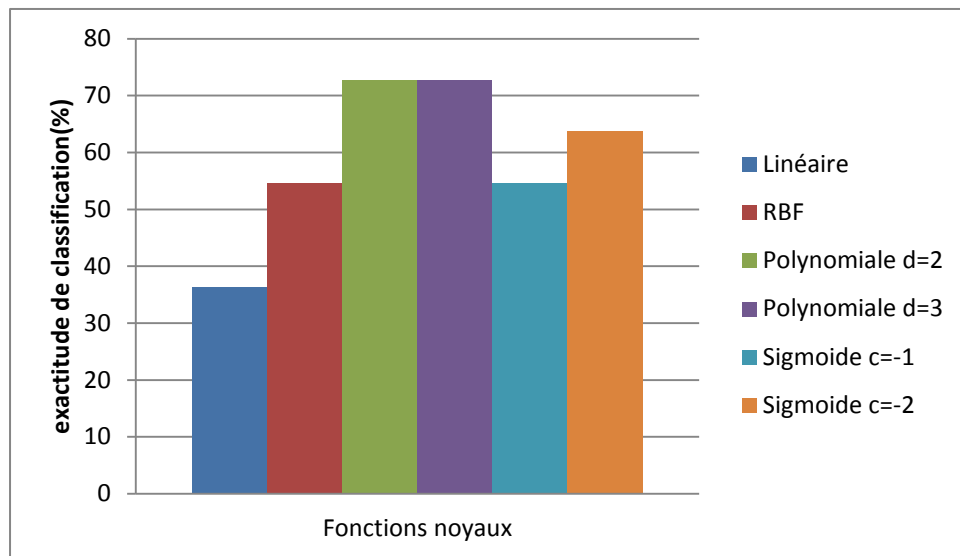


Figure II.13 : représentation de l'exactitude de classification des différentes fonctions noyaux avec la stratégie *un contre un* pour le cas de *pollution uniforme*.

II.9.2. Cas de pollution non uniforme :

Echantillons (n°)	Conductivité (mS/cm)	Classe assignée	Résultats obtenus avec les différentes fonctions noyaux					
			Linéaire	RBF	Polynomial		Sigmoïde	
					d=2	d=3	c=-1	c=-2
1	1.2	1	1	1	1	1	1	1
2	3.5	1	1	1	1	1	1	1
3	3.5	1	1	1	1	1	1	1
4	1.2	1	1	1	1	1	1	1
5	6	2	1*	1*	2	2	1*	1*
6	6	2	1*	3*	3*	3*	1*	1*
7	10	3	3	3	2*	1*	3	3
8	20	3	2*	3	3	3	3	3
9	20	3	2*	3	3	3	2*	3
Exactitude de classification (%)			55.55	77.78	77.78	77.78	66.66	77.78

Tableau II.4: Classification SVM multi-classes en utilisant la stratégie « un contre tous ».

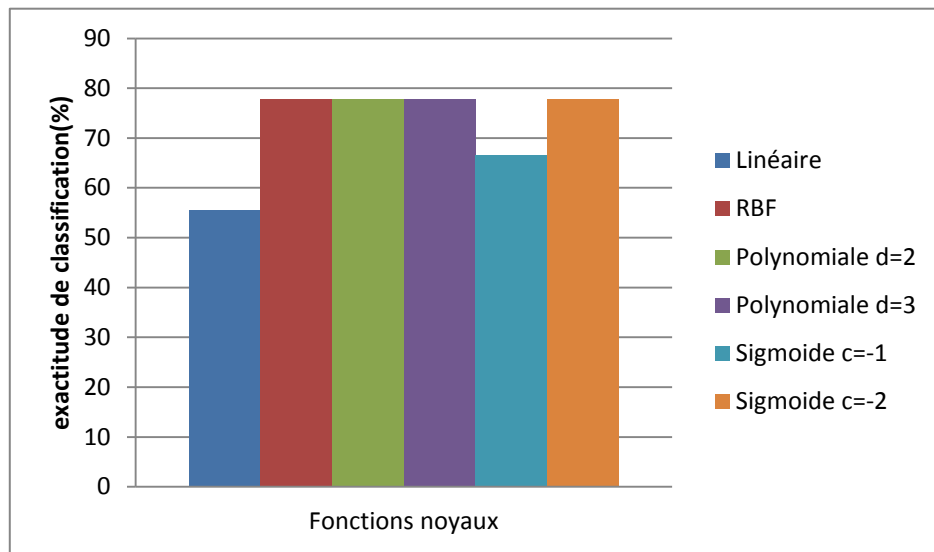


Figure II.14 : représentation de l'exactitude de classification des différentes fonctions noyaux avec la stratégie **un contre tous** pour le cas de **pollution non uniforme**.

Echantillon (n°)	Conductivité (mS/cm)	Classe assignée	Résultats obtenues avec les différentes fonctions noyaux					
			Linéaire	RBF	Polynomial		Sigmoïde	
					d=2	d=3	c=-1	c=-2
1	1.2	1	1	1	1	1	1	1
2	3.5	1	1	1	1	1	1	1
3	3.5	1	1	1	1	1	1	1
4	1.2	1	1	1	1	1	1	1
5	6	2	2	2	2	2	1*	1*
6	6	2	2	3*	3*	3*	2	2
7	10	3	3	3	2*	1*	2*	2*
8	20	3	2*	3	3	3	2*	2*
9	20	3	3	3	3	1*	2*	2*
Exactitude de classification (%)			88.89	88.89	77.78	66.66	55.55	55.55

Tableau II.5: Classification SVM multi classe en utilisant la stratégie « un contre un ».

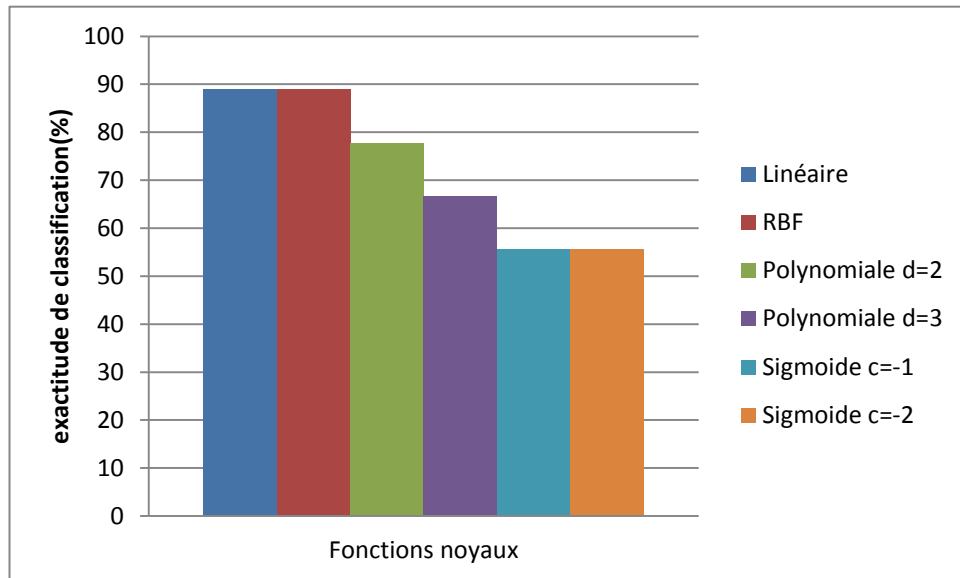


Figure II.15 : représentation de l'exactitude de classification des différentes fonctions noyaux avec la stratégie **un contre un** pour le cas de **pollution non uniforme**.

II.10. Interprétation des résultats :

II.10.1. Pollution uniforme :

Stratégie un contre tous :

Dans le cas de la stratégie « un contre tous », l'utilisation du noyau polynomial d'ordre $d=3$ a donné le meilleur résultat avec neuf bonnes classifications sur onze donc un pourcentage de 81.82%. Les deux noyaux RBF et linéaire donnent des résultats similaires avec six bonnes classifications sur onze ce qui correspond à un pourcentage de 54.54%. Tandis que, le noyau sigmoïde ne donne pas de bons résultats pour cette stratégie « un contre tous ». Nous notons un pourcentage d'exactitude de 36.36% pour les deux cas $c=-1$ et $c=-2$.

Stratégie un contre un :

Dans le cas de la stratégie « un contre un », l'utilisation du noyau polynomial d'ordre $d=2$ et $d=3$ donne les meilleurs résultats avec huit bonnes prédictions sur onze donc un pourcentage de 72.72% pour les deux cas. Le noyau sigmoïde pour le paramètre $c = -2$ donne aussi un bon résultat avec un pourcentage de 63.63%. Le noyau sigmoïde avec $c = -1$ et le noyau RBF donnent dans ce cas les mêmes résultats avec six classifications adéquates sur onze donc une exactitude de classification de 54.54%. Cependant le noyau linéaire reste le noyau le moins précis pour cette stratégie un contre un avec seulement quatre

classifications justes sur onze ce qui donne un pourcentage de 36.36%.

II.10.2. Pollution non uniforme :

Stratégie un contre tous :

Dans le cas de la stratégie « un contre tous », les quatre noyaux RBF, polynomial d'ordre $d=2$, polynomial d'ordre $d=3$ et sigmoïde avec le paramètre $c = -2$ donnent des résultats performants avec sept classifications justes sur neuf donc un pourcentage de 77.77%. Le noyau sigmoïde de paramètre $c = -1$ donne des résultats moins performants par rapport aux autres noyaux avec six bonnes classifications sur neuf ce qui veut dire 66.66%. Enfin pour le noyau linéaire, les résultats de classifications sont moyens avec 55.55% donc cinq classifications justes sur neuf.

Stratégie un contre un :

Dans le cas de la stratégie « un contre un », les 2 noyaux RBF et linéaire donnent de bons résultats avec huit classifications justes sur neuf donc un pourcentage de 88.88%. Le noyau polynomial donne des résultats moins performants que les deux premiers noyaux avec sept bonnes classifications sur neuf pour $d=2$ et six sur neuf pour $d=3$. Pour le noyau sigmoïde avec $c = -1$ et $c = -2$, les résultats de classifications sont intermédiaires avec cinq classifications justes sur neuf ce qui donne un pourcentage de 55.55%.

Conclusion :

Dans la première partie de ce chapitre, nous avons présenté les concepts de la méthode SVM, dont nous avons expliqué la démarche de construction de l'hyperplan optimal dans les cas des données linéairement séparables et non linéairement séparables; en se basant sur les fondements mathématiques pour le cas binaire (à deux classes). Aussi nous avons défini quelques fonctions noyaux utilisés par les SVM non linéaires pour la transformation et présentation des vecteurs d'entrée dans l'espace des caractéristiques.

Nous avons ensuite décrit la démarche suivie pour généraliser cette technique pour le cas de multi-classes à partir du cas binaire en utilisant les trois approches : la discrimination binaire, Un-contre-Un ou Un-contre-Tous.

Dans la seconde partie, nous avons présenté la procédure que nous avons utilisée pour la construction des modèles SVM ainsi que les résultats de classification pour les deux

stratégies avec les différentes fonctions noyaux.

Nous concluons pour le cas de notre base de données que la stratégie « un contre tous » avec le noyau polynomial d'ordre $d=3$ pour le cas de pollution uniforme est le meilleur modèle à utiliser. Pour le cas de pollution non uniforme, le meilleur modèle a utilisé est celui de la stratégie « un contre un » avec la fonction noyau RBF.

On peut expliquer ça par le fait que les fonctions RBF et polynomiale projettent les données dans un espace de dimension plus grand qu'une fonction linéaire. De ce fait, les données seront plus facilement séparables.

En dernier lieu, il faut mentionner que la méthode SVM est une méthode d'apprentissage supervisé performante et fiable pour la classification.

Conclusion générale

Ce travail nous a permis d'appliquer de nouvelles méthodes de traitement de signal pour la classification de la sévérité de pollution des isolateurs HT, en vue de diagnostiquer l'état de surface de l'isolateur tout en se référant à la norme CEI 71 pour classer le degré de sévérité de pollution a trois classes différentes : faible pollution, moyenne pollution et forte pollution.

En premier lieu, nous avons décomposé les signaux du courant de fuite par la DWT pour avoir les caractéristiques des dix premiers détails. Ensuite, l'étude de la variation de l'écart type, de maximum, de minimum et de la moyenne en fonction du niveau de décomposition et en fonction de la conductivité de l'agent polluant (1.2, 3.5, 6, 10 et 20 mS/cm) nous renseigne sur la sévérité de pollution. Ces dernières caractéristiques sont utilisées comme entrées pour le modèle SVM que nous avons construit.

En deuxième lieu, nous avons approfondi nos connaissances sur les types d'apprentissage, la théorie des SVM comme une des méthodes d'apprentissage supervisé et son utilisation pour la classification dans le cas binaire et le cas multi-classes afin de les utiliser pour la classification de sévérité de pollution des isolateurs de haute tension par la suite.

Les résultats de cette classification montrent que la meilleure stratégie dans le cas de pollution uniforme est la stratégie « un contre tous » avec la fonction noyau polynomiale d'ordre $d=3$ avec un pourcentage de 81.82 %. Pour le cas de pollution non uniforme la meilleur stratégie été « un contre un » avec la fonction noyau RBF et un pourcentage de 88.89%.

L'intérêt principal de la méthode SVM est sa facilité d'emploi, la mise en œuvre des algorithmes des SVM est peu coûteuse, mais le point noir est la durée des phases de tests qui sont assez longues lorsqu'il s'agit de régler les paramètres de la fonction noyau. Parmi d'autres avantages des SVM on cite : qu'ils possèdent des fondements théoriques solides, une optimisation très simple, et un temps de calcul court.

Les SVM peuvent également s'utiliser pour des tâches de prédiction des variables continues en fonction d'autres variables (température, évolution des marchés, etc.). Le champ d'application des SVM est donc large.

Références bibliographiques

- [1] M. Tegar, « Modélisations d'isolateurs pollués soumis à divers paramètres électrogéométriques », Thèse de Doctorat d'Etat, Département d'électrotechnique, Ecole Nationale Polytechnique, Alger, 2003.
- [2] M. A. B. Elkosaïr, E. E. I. Sharaine, M. M Awad, H. E. M Zoroura, M. M. Khalifa et A.Nossier, « Performance de chaînes d'isolateurs haute tension soumises à la pollution du désertique », rapport 33-05, France, 1970.
- [3] W.Heise, G.F.luxa, G.Revrery, M. P. Verma, « Estimation de la méthode d'essais sous pollution artificielle par couche solide », CIGRE, rapport 33-09, 1982.
- [4] A. Cimador et S. Vilet, « La Pollution des Isolateurs », EDF-Epure, Rapport N° 27, France, Juillet 1990.
- [5] T. Suda, « Frequency Characteristics of Leakage Current Waveforms of Artificially Polluted Suspension insulators », IEEE Trans. Dielectr. Electr. Insul, Vol. 8, pp.705 -709, 2005.
- [6] X.Yan Shi, C.Sun and Z. Zhang, « Evaluating the Safety Condition of Porcelain Insulators by the Time and Frequency Characteristics of LC Based on Artificial Pollution Tests », IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 17, N° 2; April 2010, pp1-9.
- [7] O. Bennouna, O. Robin, H. Chafouk et J.P. Roux, « Diagnostic et détection de défauts des systèmes embarqués dans l'automobile », Rapport scientifique, Saint Etienne du Rouvray, France.
- [8] P. Savicky et J. Fürnkranz, « Combining pairwise classifiers with stacking », International symposium on intelligent data analysis, Berlin, Germany, pp. 219-229, 2003.
- [9] Harzelli Imadeddine et Maanani Yacine, « Mesure du courant de fuite et de la charge électrique d'un isolateur pollué de haute tension. Caractérisation de son état de surface en

utilisant la transformée en ondelette », Projet de fin d'études, Ecole nationale polytechnique, Alger, Juin 2011.

[10] Landrin Patrick Flandrin, « Temps-fréquence », Edition Hes, Collection traitement du signal, 1998.

[11] Fabrice Chane-Ming, Franck Molinaro et Jean Leveau, « Applications de l'analyse en ondelettes sur un signal géophysique », Laboratoire de Physique de l'Atmosphère Université de la Réunion, France, Septembre 1997.

[12] Amirouche Annouche et Mohamed Said Berkani, « Traitement et analyse du signal du courant de fuite pour la caractérisation des surfaces isolantes polluées par apprentissage supervisé », PFE, Ecole nationale polytechnique, Alger, Juin 2011.

[13] H.V. Luong, « Étude de la méthode de la transformation en ondelette et l'application à la compression des images », Rapport de stage, 15 juillet 2005.

[14] C.Chang and C.J. Lin, « LIBSVM: A library for support vector machines », <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.

[15] B. Moula, A. Mekhaldi et M. Tegar, « Comparaison des Performances d'un Isolateur Haute Tension Soumis aux Conditions de Pollution Artificielle », CNHT'2011–8ème Conférence Nationale sur la Haute Tension, Tiaret, 09-11 Mai 2011, pp 33-37.

[16] A.Cavallini et G.C. Montanari, « Investigations on Leakage Current and Phase Angle Characteristics of Porcelain and Polymeric Insulator under contaminated Conditions », IEEE Transactions on Dielectrics and Electrical Insulation Vol. 16, No. 2; April 2009, pp1 -10.

[17] J. Kharroubi, « Étude de Techniques de Classement, Machines à Vecteurs Supports pour la Vérification Automatique du Locuteur », Thèse de Doctorat, Ecole nationale supérieure des télécommunications, France, 2002.

[18] Ludovic Mercier, « Les machines à vecteurs support pour la classification en imagerie hyperspectrale : implémentation et mise en œuvre », centre d'enseignement de Grenoble, France, février 2010.

- [19] Pascal Vincent, « Modèles à noyaux à structure locale », Thèse PhD, Université de Montréal, Canada, Octobre 2003.
- [20] V. Vapnik, « The nature of statistical learning theory », Springer, New York, 1995.
- [21] Emmanuel Jakobowicz, « La discrimination à plus de deux classes Comparaison de plusieurs approches issues des Support Vector Machines (SVM) », Ecole Polytechnique Fédérale de Lausanne, Suisse.
- [22] Mohamadally Hasan et Fomani Boris, « SVM : Machines à Vecteurs de Support ou Séparateurs à Vastes Marges », Thèse de Doctorat, Versailles St Quentin, France, Janvier 2006.
- [23] C.M. Bishop, « Pattern Recognition and Machine Learning », General library system, Information Science and Statistics, Springer, New York, USA, 2006.
- [24] A. Aizerman, E.M. Braverman and L.I. Rozoner, « Theoretical foundations of the potential function method in pattern recognition learning », Automation and Remote Control, 28(1964), pp. 821-837.
- [25] F. Amidi et k. Ouerdane, « Les effets de la pollution sur les ouvrage électrique à 30kV en région littorales », Projet de fin d'études, Université des sciences et technologies Houari Boumediene, Alger, Juin 1998.
- [26] S. Knerr, L. Personnaz, and G. Dreyfus, « Single-layer learning revisited: a stepwise procedure for building and training a neural network », Ecole superieur de physique et de chimie industrielles de la vie de Paris, Vol.F68, France, 1990.
- [27] U. H. G. Krebel, « Pairwise classification and support vector machines », The MIT Press, Cambridge, 1999, pp. 255-268.
- [28] Gilles Lebrun, « Sélection de modèles pour la classification supervisée avec des SVM (Séparateurs à Vaste Marge).Application en traitement et analyse d'images », Université de Caen/Basse-Normandie, France, Novembre 2006.
- [29] « Artificial pollution tests on high-voltage insulators to be used on AC systems », IEC 85-1, 1991.