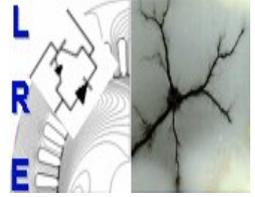




Ecole Nationale Polytechnique
Département d'Electrotechnique
Laboratoire de Recherche en Electrotechnique



Mémoire de Master en Electrotechnique

Présenté par :

ZEBIRI Karim

Intitulé

Classification de la sévérité de pollution des isolateurs par la méthode de l'Analyse Linéaire Discriminante

Directeur(s) du mémoire :

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

Jury:

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

ENP 2013

ملخص :

في العمل الحالي، يتم إجراء تصنيف شدة تلوث عوازل التوتر العالي بناء 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 أداء تصنيف لشدة التلوث باستعمال طريقة LDA. ثانياً، سنقوم بتحليل إشارات التوتر المطبق باستعمال CWT لتقدير حالة سطح النموذج.

الكلمات الرئيسية:

شدة تلوث، تلوث منتظم، تلوث غير منتظم، التيار المتسرب، التصنيف، تقنية الموجات المستمرة، تقنية الموجات المتقطعة، تحليل التمايز الخطي.

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'effectuer la classification de la sévérité de pollution par la méthode LDA. En deuxième lieu, nous analysons les signaux de la tension appliquée par la CWT pour estimer l'état de surface du modèle.

Mots clés :

Sévérité de pollution, pollution uniforme, pollution non uniforme, courant de fuite, classification, transformée en ondelettes discrète, transformée en ondelettes continue, analyse discriminante linéaire.

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 classify the severity of pollution using LDA method. After that, we analyze the signals of the voltage applied using the CWT to estimate the surface quality of the model.

Key words:

Pollution severity, uniform pollution, not uniform pollution, leakage current, supervised training, classification, Discrete Wavelet Transform, Continuous Wavelet Transform, Discriminant Linear Analysis.

Tables des matières

Introduction générale:	1
------------------------------	---

Chapitre I : Théorie de l'Analyse Linéaire Discriminante

Introduction :	3
I.1. Analyse discriminante :	3
I.2. La règle de décision de Bayes :	4
I.3. Les principales méthodes d'Analyse Discriminante :	5
I.3.1. Analyse Discriminante Quadratique (QDA) :	5
I.3.2. Analyse Discriminante Linéaire (LDA) :	5
I.3.3. Analyse Factorielle Discriminante (FDA) :	6
I.4. La classification avec l'analyse discriminante linéaire :	8
Conclusion :	10

Chapitre II : Application de la LDA pour la classification de la sévérité de pollution

Introduction :	11
II.1. Rappel sur la transformée de Fourier :	11
II.2. La transformée en ondelettes :	12
II.2.1. L'ondelette d'analyse:	13
II.2.3. L'analyse multi-résolution (AMR):	14
II.2.4. La transformée en ondelettes discrètes (DWT) :	14
II.2.4.1. Niveau maximal de décomposition :	15
II.2.4.2. Choix de l'ondelette analysante :	15
II.3. Décomposition du signal du courant de fuite par DWT :	16
II.4. Calcul des caractéristiques des coefficients de détails d'ondelette :	17
II.4.1. Etude de la standard déviation des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :	18
II.4.2. Etude des maximums des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :	20

II.4.3. Etude des minimums des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :	23
II.4.4. Etude des moyennes des détails du signal du courant de fuite pour une pollution avec 100 g/l de Kaolin :	25
II.5. Les Classes de pollution :	27
II.6. Application de l'analyse linéaire discriminante (LDA) pour la classification de la sévérité de pollution des isolateurs :	28
II.6.1. Composition du vecteur d'état :	28
II.6.2. Base de données pour le cas de la pollution uniforme :	29
II.6.3. Base de données pour le cas de la pollution non uniforme :	29
II.7. Validation et résultat de classification :	29
Conclusion :	31

Chapitre III : Caractérisation de l'état de surface des isolateurs par la technique des ondelettes

Introduction	32
III.1. La transformée en ondelettes continues (CWT) :	32
III.1.1. Scalogramme :	32
III.1.2. Calcul des coefficients de la CWT :	33
III.2. Pollution uniforme :	35
III.2.1 Influence de la conductivité volumique:	35
III.2.2. Influence de la concentration en kaolin :	39
Conclusion :	43
Conclusion générale:	45

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 contournement des isolateurs sous haute tension et en présence d'impuretés conductrices est lié à l'apparition des décharges électriques qui progressent à leurs surfaces conductrices jusqu'au court-circuit. Il peut être caractérisé par deux phases successives [3]:

- Initiation d'une décharge électrique instable à un certain seuil de tension à cause d'un assèchement local de la surface conductrice.
- Progression de la décharge électrique avec l'augmentation du courant de fuite sous des conditions de pollution favorables, pour former un arc de contournement stable qui modifie partiellement la conductivité du milieu.

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 [4].

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 [5].

L'apprentissage supervisé a émergé ces dernières années comme un outil efficace pour la classification dans de nombreux domaines [6]. L'Analyse Linéaire Discriminante (LDA) est une de ces techniques de classification.

Le présent travail consiste à utiliser le signal du courant de fuite pour extraire des critères de jugement sur l'état de surface des isolateurs pollués et servir pour la classification de la sévérité de pollution des isolateurs pollués donné.

Dans le premier chapitre, nous présentons la théorie de la classification par l'Analyse Discriminante.

Dans le deuxième chapitre, nous donnons des rappels sur la transformée en ondelettes, plus particulièrement la décomposition du signal du courant de fuite en utilisant la transformée en ondelette discrète (DWT). Nous présentons par la suite les différentes classes associées à chaque niveau de pollution selon la CEI 61. A la fin de ce chapitre nous utilisons la technique de la LDA pour faire la classification à partir des caractéristiques des signaux du courant de fuite pour les deux cas de pollution uniforme et non uniforme.

Dans le troisième chapitre, nous mettons en application l'analyse par la technique des ondelettes. Pour la pollution uniforme, la transformée en ondelette continue (CWT) est appliquée sur les signaux du courant de fuite pour un éventuel diagnostic de l'état de surface.

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.

Chapitre I:

Théorie de

l'Analyse

Linéaire

Discriminante

Introduction :

L'objectif de la classification supervisée est d'affecter des individus à des groupes définis a priori à partir des mesures effectuées sur des variables. Dans ce contexte, les analyses discriminantes linéaire et quadratique sont parmi les méthodes les plus populaires. Elles sont fondées sur des hypothèses de multinormalité [7].

L'analyse discriminante classique suppose que l'échantillon d'apprentissage et l'échantillon test qui contient les individus à classer, sont issus d'une même population. Depuis les travaux de Fisher (1936) qui a introduit une règle de discrimination linéaire entre deux groupes, de nombreuses évolutions ont été proposées [8].

Dans le présent chapitre, nous allons présenter la théorie concernant l'analyse discriminante ainsi que les principales méthodes génératives dont la plus connue est l'Analyse Discriminante Linéaire (LDA).

I.1. Analyse discriminante :

L'analyse discriminante est le nom donné à la classification dans le cadre supervisé. Elle se distingue de la classification non supervisée par le fait que des observations dont on connaît l'appartenance aux classes sont disponibles pour apprendre la règle de décision (on parlera aussi parfois de classifieur). Ces observations, dites d'apprentissage, supervisent la construction du classifieur [9].

Le problème de l'analyse discriminante est de prédire l'appartenance d'une observation x décrit par p variables explicatives, à une classe parmi k classes $C_1 \dots C_k$ définies a priori. Afin de prédire l'appartenance de l'observation x à une des k classes, nous disposons d'un ensemble d'apprentissage :

$$A = \{(x_1, C_1), \dots, (x_n, C_n), x_i \in \mathbb{R}^p, C_i \in \{1, \dots, k\}\} \quad (\text{I.1})$$

Où le vecteur x_i est la $i^{\text{ème}}$ observation et C_i indique le numéro de la classe à laquelle x_i appartient. Nous allons donc utiliser l'échantillon A pour construire une règle de décision δ qui associe tout vecteur x de \mathbb{R}^p à la classe correspondante.

$$\delta: \mathbb{R}^p \rightarrow \{1, \dots, k\} \quad (\text{I.2})$$

$$x \rightarrow C$$

On distingue classiquement deux objectifs principaux en Analyse Discriminante:

-Descriptif : l'aspect descriptif vise à trouver une représentation qui permet l'interprétation des groupes grâce aux variables explicatives.

-Décisionnel : dans ce cas, on cherche à définir la bonne affectation d'un nouveau individu dont on ne connaît que les valeurs des variables explicatives. Cet aspect est particulièrement apprécié dans des domaines où l'aspect diagnostic est essentiel [10].

Dans notre travail, nous nous intéresserons plus particulièrement à l'aspect décisionnel qui est le plus important et souvent le plus délicat. Nous allons tout d'abord rappeler le modèle probabiliste de la discrimination avant de décrire les principales méthodes paramétriques d'Analyse Discriminante.

I.2. La règle de décision de Bayes :

Les méthodes paramétriques d'Analyse Discriminante font une hypothèse de normalité des classes [10]. Les densités de probabilité des variables explicatives conditionnellement aux classes $p(x|C_i), \forall i = 1, \dots, k$ sont supposées celles de lois normales $N(\mu_i; S_i)$ de moyennes μ_i et de matrice de variance S_i :

$$p(x|C_i) = \frac{1}{(2\pi)^{\frac{p}{2}} (\det S_i)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} (x - \mu_i)^t S_i^{-1} (x - \mu_i)\right) \quad (I.3)$$

La règle de décision optimale δ^* , dite règle de Bayes, affecte le point x à la classe qui a la probabilité a posteriori maximum. La règle de décision δ^* prend la forme suivante [69]:

$$x \in C_{i^*}, \text{ si } i^* = \operatorname{argmax}_{i=1, \dots, k} \{p(C_i|x)\} \quad (I.4)$$

La formule de Bayes permet d'obtenir la probabilité que l'individu x provienne de la classe C_i :

$$p(C_i|x) = \frac{p(C_i|x)\pi_i}{p(x)} \quad (I.5)$$

$$\text{Où } \pi_i \text{ est la probabilité a priori de la classe } C_i \text{ et } p(x) = \sum_{i=1}^k p(x|C_i) \quad (I.6)$$

Par conséquent et comme les dénominateurs de l'équation (I.5) sont communs pour chacune des k classes, la règle de Bayes consiste donc à affecter x à la classe C_{i^*} si:

$$i^* = \operatorname{argmax}_{i=1, \dots, k} \{\pi_i p(x|C_i)\} \quad (I.7)$$

Afin de faciliter l'écriture des règles de décision par la suite, nous allons définir la fonction de coût K_i conditionnellement à la classe $C_i, \forall i = 1, \dots, k$ de la façon suivante:

$$K_i : \mathbb{R}^p \rightarrow \mathbb{R}$$

$$x \rightarrow -2 \log(\pi_i p(x|C_i)) \quad (\text{I.8})$$

Avec cette notation, la règle de Bayes consiste donc à affecter x à la classe C_{i^*} si [10]:

$$i^* = \operatorname{argmin}_{i=1,\dots,k} \{K_i(x)\} \quad (\text{I.9})$$

I.3. Les principales méthodes d'Analyse Discriminante :

Dans cette section, nous allons décrire brièvement les principales méthodes paramétriques de l'Analyse Discriminante.

I.3.1. Analyse Discriminante Quadratique (QDA) :

Si l'on suppose que les densités conditionnelles des classes $p(x|C_i)$ sont gaussiennes, on obtient alors l'Analyse Discriminante Quadratique (QDA) qui doit son nom au fait qu'elle réalise des séparations quadratiques entre les classes [9]. La fonction de coût K_i prend dans ce cas la forme suivante :

$$K_i(x) = (x - \mu_i)^t S_i^{-1} (x - \mu_i) + \log(\det S_i) - 2\log(\pi_i) + C^{te} \quad (\text{I.10})$$

Où la constante C^{te} représente une quantité commune à toutes les classes et n'intervient donc pas dans la règle de décision.

La figure I.1 (gauche) montre la frontière de décision de QDA pour un jeu de données artificielles de dimension 2.

I.3.2. Analyse Discriminante Linéaire (LDA) :

L'Analyse Discriminante Linéaire (LDA) est basée sur le critère de la maximisation de l'écart entre les classes et la minimisation de la dispersion autour de la moyenne de chaque classe [9]. Cette analyse combine linéairement les données de manière à maximiser la discrimination entre les classes. Le nombre d'axes générés par la LDA est inférieur ou égal au nombre d'étiquettes différentes.

L'analyse discriminante maximise la répartition des N observations dans leurs classes respectives, contrairement à l'analyse factorielle qu'on va voir par la suite qui maximise la variance des N observations en fonction des variables [9].

La LDA est généralement considérée comme une méthode d'analyse discriminante à part entière plutôt que comme une méthode de régularisation de QDA, car elle fait l'hypothèse supplémentaire d'égalité des matrices de variances, i.e. $\forall i = 1, \dots, k, S_i = S$.

L'analyse discriminante linéaire doit également son nom au fait qu'elle réalise des séparations linéaires entre les classes [9].

Les termes $\log(\det S)$ et $x^t S^{-1} x$ ne dépendant plus des classes comme dans le cas de QDA, mais ils peuvent être utilisés pour le calcul de la règle de décision. La fonction de coût K_i prend la forme suivante :

$$K_i(x) = \mu_i^t S^{-1} x - 2\mu_i^t S^{-1} x - 2\log(\pi_i) + C^{te} \quad (\text{I.11})$$

Les estimateurs des proportions et des moyennes sont les mêmes que ceux de QDA et l'estimateur du maximum de vraisemblance de la matrice S est la matrice de covariance intra-classe empirique W :

$$\hat{S} = W = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^n s_{ij} (x_j - \hat{\mu}_i) (x_j - \hat{\mu}_i)^t \quad (\text{I.12})$$

La Figure I.1 (droite) montre la frontière de décision de LDA pour des données artificielles de dimension 2. En pratique, l'analyse discriminante linéaire est fréquemment utilisée car elle offre un bon compromis entre pertinence et complexité. D'autre part, elle fournit des résultats robustes aux fluctuations sur les hypothèses de normalité des classes et d'égalité des matrices de variance [9].

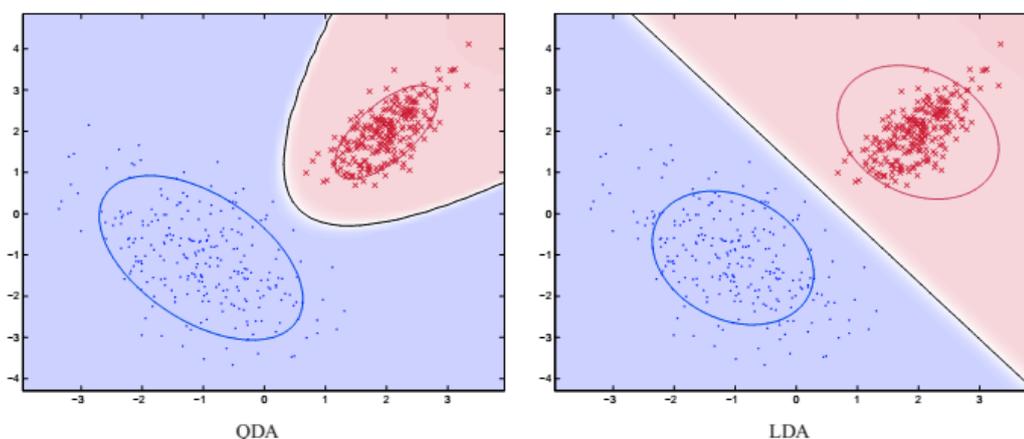


Figure I.1: Frontières de décision de l'Analyse Discriminante Quadratique (QDA) et de l'Analyse Discriminante Linéaire (LDA) sur un même jeu de données en dimension 2.

I.3.3. Analyse Factorielle Discriminante (FDA) :

L'Analyse Factorielle Discriminante combine une étape qui relève de la réduction de dimension et une étape de discrimination [10]. En effet, effectuer une Analyse Factorielle Discriminante consiste à projeter les données de \mathbb{R}^p sur les $d = (k - 1)$ axes discriminants qui maximisent le rapport de la variance interclasse et de la variance intra-classe, puis

d'apprendre la règle de décision δ^* sur les données projetées. Nous avons donc besoin de définir les matrices de variance inter et intra-classe. La matrice de variance interclasse est définie par:

$$B = \sum_{i=1}^k \pi_i (\mu_i - \mu) (\mu_i - \mu)^t \quad (\text{I.13})$$

$$\text{Où } \mu = \sum_{i=1}^k \pi_i \mu_i \quad (\text{I.14})$$

D'autre part, la matrice de variance intra-classe est définie par:

$$W = \sum_{i=1}^k \pi_i \Sigma_i \quad (\text{I.15})$$

Notons également que le théorème de Huyghens nous permet d'obtenir la relation suivante qui lie les matrices de variance inter et intra-classe à la matrice de variance totale:

$$S_{totale} = B + W \quad (\text{I.16})$$

Nous souhaitons trouver une représentation des données qui permet de discriminer les groupes le mieux possible. Pour ce faire, il faut que les projections des k centres de gravité soient le plus séparée possible, tandis que les données de chaque classe doivent se projeter de façon groupée autour du centre de gravité de leur classe. Nous recherchons donc une représentation des données qui maximise la variance interclasse et qui minimise la variance intra-classe [10]. Avec les notations et résultats précédents, les axes de la projection recherchés satisfont le problème d'optimisation suivant: $\max_u \frac{u' B u}{u' S_{totale} u}$ (I.17)

On sait que ce maximum est atteint pour u vecteur propre $S_{totale}^{-1} B$ associé à sa plus grande valeur propre. La Figure I.2 illustre le choix d'un axe de projection permettant de discriminer au mieux les classes [10].

Une fois la projection déterminée, on peut alors effectuer une Analyse Discriminante Linéaire (ou Quadratique). Cette stratégie qui combine réduction de dimension et discrimination est souvent profitable car les données de chaque classe n'occupent en général pas la totalité de l'espace et cela permet de réduire le nombre de paramètres à estimer. Cette méthode se révèle relativement efficace sur des données de grande dimension.

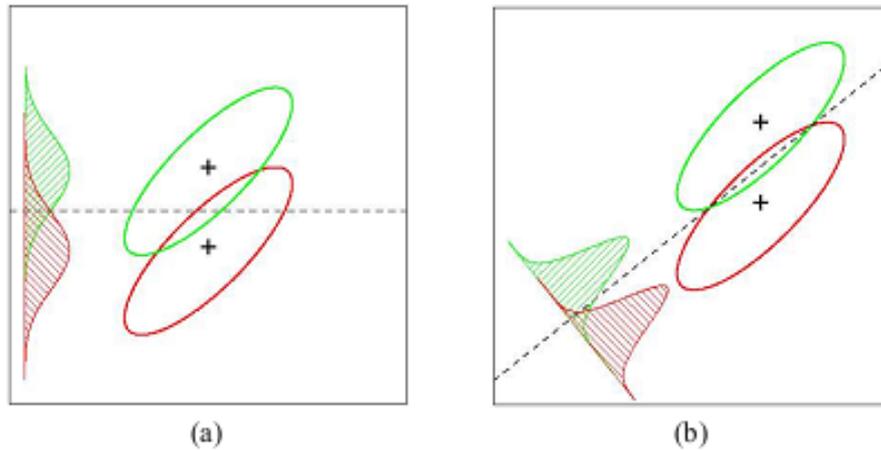


Figure I.2: l'axe principal de la figure(a) ne permet pas de discriminer efficacement les deux groupes alors que celui de la figure(b) possède un bon pouvoir discriminant.

I.4. La classification avec l'analyse discriminante linéaire :

En se basant sur la théorie de l'analyse discriminante linéaire, le principe de classification est le suivant :

- Les vecteurs à classifier x_i , $i = 1 \dots n$.
- Les classes correspondantes C_i , $i=1, \dots, m$. Tel que m est le nombre des classes.
- Le centre de gravité des vecteurs x_i : $\mu = \frac{1}{n} \sum_{k=1}^n x_k$ (I.18)
- le centre de gravité des classes c_i : $\mu_i = \frac{1}{n_i} \sum_{x_k \in C_k} x_k$ (I.19)

La LDA nous permet de calculer les mesures suivantes :

$$\text{La variance de la } i^{\text{ème}} \text{ classe : } S_i = \sum_{x_k \in C_k} (x_k - \mu_i)(x_k - \mu_i)^t \quad (\text{I.20})$$

$$\text{La matrice de variance intra-classes : } W = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_k - \mu_i)(x_k - \mu_i)^t \quad (\text{I.21})$$

$$\text{La matrice de variance interclasse : } B = \frac{1}{n} \sum_{i=1}^m (\mu_k - \mu)(\mu_k - \mu)^t \quad (\text{I.22})$$

$$\text{La matrice totale de variance : } S_{\text{totale}} = W + B \quad (\text{I.23})$$

Fisher définit la séparation entre un certain nombre de classes en maximisant le rapport , tel que : $S = \frac{B^2}{W^2}$ (I.24)

Le numérateur de l'expression IV.24 traduit la séparation des classes, alors que le dénominateur reflète la compacité des classes. Plus le rapport est grand, plus les classes sont séparées. Ainsi, d'une manière générale, pour m classes, et pour un paramètre α

donné, l'expression précédente devient :

$$S = \sum_{i=1}^m \sum_{j=1}^{m-1} \frac{\mu_i(\alpha) - \mu_j(\alpha)}{n_i \sigma_i^2(\alpha) - n_j \sigma_j^2(\alpha)} \quad (\text{I.25})$$

σ_i^2 est la $i^{\text{ème}}$ variance des vecteurs de la classe C_i , tel que :

$$\sigma_i^2(\alpha) = \frac{1}{n_i} \sum_{j=1}^{n_i} [x_{ij}(\alpha) - \mu_i(\alpha)]^2 \quad (\text{I.26})$$

De ce fait, le critère de Fisher a pour but de maximiser la variance interclasse, et de minimiser la variance intra-classe. Ainsi, dans le cas des deux classes et pour un paramètre, le critère de Fisher s'exprime de la manière suivante :

$$S(\alpha) = \frac{\mu_1(\alpha) - \mu_2(\alpha)}{n_1 \sigma_1^2(\alpha) - n_2 \sigma_2^2(\alpha)} \quad (\text{I.36})$$

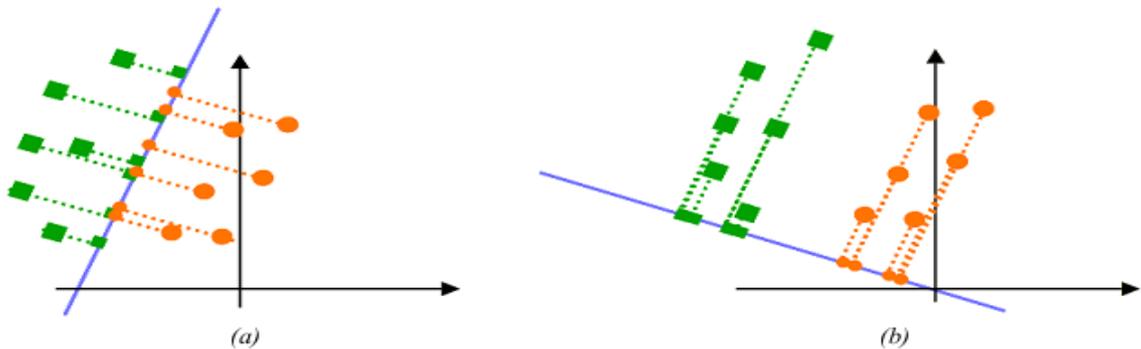


Figure I.3: Exemple d'une séparation de deux classes avec la droite discriminante de Fisher.
(a) mauvaise séparation. (b) bonne séparation.

Pour les deux cas de pollution, les résultats obtenus montrent clairement que la méthode SVM surpasse la méthode LDA. Ce résultat peut être expliqué par le fait que la méthode SVM cherche l'hyperplan qui sépare les exemples positifs des exemples négatifs, en garantissant que la marge entre le plus proche des positifs et des négatifs soit maximale, alors intuitivement, cela garantit un bon niveau de généralisation car de nouveaux exemples pourront ne pas être trop similaires à ceux utilisés pour trouver l'hyperplan mais être tout de même situés franchement d'un côté ou de l'autre de la frontière.

Conclusion :

Dans ce premier chapitre nous avons approfondie nos connaissances sur l'analyse discriminante et ces trois méthodes de classification QDA, LDA et l'analyse discriminante factorielle afin de l'utiliser pour la classification de la sévérité de pollution des isolateurs de haute tension.

Chapitre II:

*Application de la
LDA pour la
classification de la
sévérité de
pollution*

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 [15].

Dans la première partie du présent chapitre, nous abordons d'une manière explicite la transformée en ondelettes discrète. 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.

Dans la seconde partie, nous allons d'abord présenter les différentes classes associées à chaque concentration en kaolin. Ensuite, nous appliquerons la méthode de la LDA afin de classer le degré de sévérité de pollution des isolateurs pour les deux types de pollution en se basant sur les caractéristiques des signaux du courant de fuite.

II.1. Rappel sur la transformée de Fourier :

La transformée en ondelettes, comme la transformée de Fourier, est une transformation d'un espace d'entrée (espace-temps) vers un autre espace (espace fréquentiel) [11]. Dans le cas de la transformée de Fourier classique, la base de l'espace d'entrée est par définition $\{e(j\omega t), \omega \in \mathbb{R}\}$, comme nous le montre la formule exprimant la transformée inverse de Fourier:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega) \exp(j\omega t) d\omega \quad (\text{II.1})$$

Dans le cas de la transformée de Fourier numérisée ou discrète (Fast Fourier Transform), les variables t et $\omega = 2\pi f$ prennent un nombre fini de valeurs espacées de $T = \Delta t$:

$$f[n] = \frac{1}{N} \sum_{k=0}^N \hat{f}[k] \exp\left(\frac{j2\pi kn}{N}\right) \quad (\text{II.2})$$

La base devient discrète : $\{e_k[n] = \exp\left(\frac{j2\pi kn}{N}\right), 0 \leq k \leq N\}$ (II.3)

Dans les deux cas, la base de l'espace fréquentiel est constituée de distributions de Dirac $\{\delta(\omega)\}$, la variable ω prenant des valeurs discrètes (FFT) ou variant de manière continue dans \mathbb{R} .

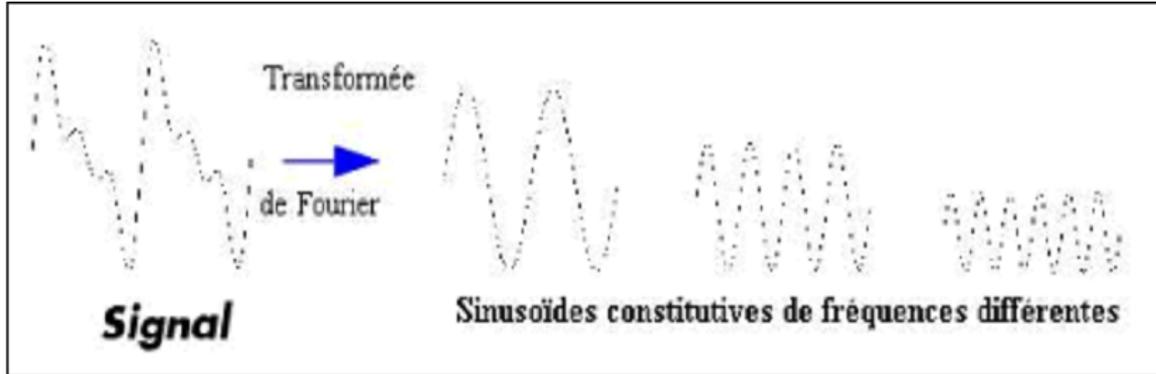


Figure II.1: Principe de la transformée de Fourier.

II.2. 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. Dans la transformée en ondelettes, comme dans l'analyse de Fourier, on cherche à transformer un signal quelconque en une série de nombres que l'on pourra ensuite utiliser pour reconstruire au mieux le signal d'origine [12]. Cependant dans la transformée en ondelettes, on utilise plusieurs niveaux de résolution pour examiner le signal et faire ressortir les différentes variations [16].

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 [18].

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 $\Psi(t)$:

$$\Psi_{(b,a)}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad , b \in \mathbb{R}, a > 0 \quad (\text{II.4})$$

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 [13].

II.2.1. L'ondelette d'analyse:

Un grand nombre d'ondelettes analysantes a été construit. Plusieurs types de ces ondelettes ont été rapportés dans la littérature dont certaines portent le nom de leurs créateurs. Les plus connues sont représentées dans la figure suivante : Daubechies, Gaussienne, Haar, Mexican Hat, Meyer et Morlet [14].

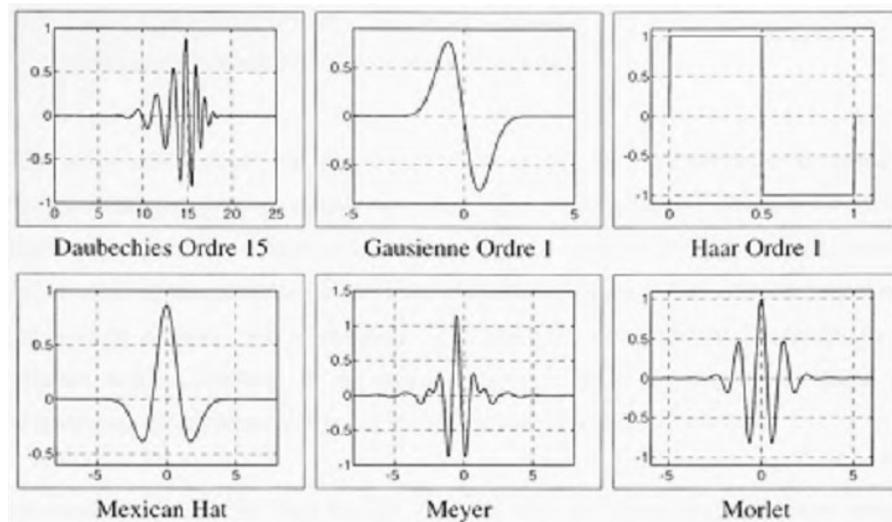


Figure II.2: Forme d'onde des ondelettes couramment utilisées.

Le choix de l'ondelette analysante n'est pas arbitraire. Pour être admissible, l'ondelette $\Psi(t)$ doit respecter les conditions suivantes :

- $\Psi(t)$ ainsi que son carré doivent être intégrables :

$$\int |\Psi(t)| dt < \infty \quad \text{et} \quad \int |\Psi(t)|^2 dt < \infty \quad (\text{II.5})$$

Cette condition implique que l'énergie de l'ondelette doit être finie.

- si $\widehat{\Psi}(t)$ représente la TF de $\Psi(t)$, alors : $\int \frac{|\widehat{\Psi}(\omega)|^2}{\omega} d\omega < \infty$ (II.6)

Cette condition implique que la moyenne de $\Psi(t)$ soit nulle ($\widehat{\Psi}(0) = 0$ où intégrale de $\Psi(t)$ dt=0).

II.2.3. 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 : [21]

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

II.2.4. 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 [17].

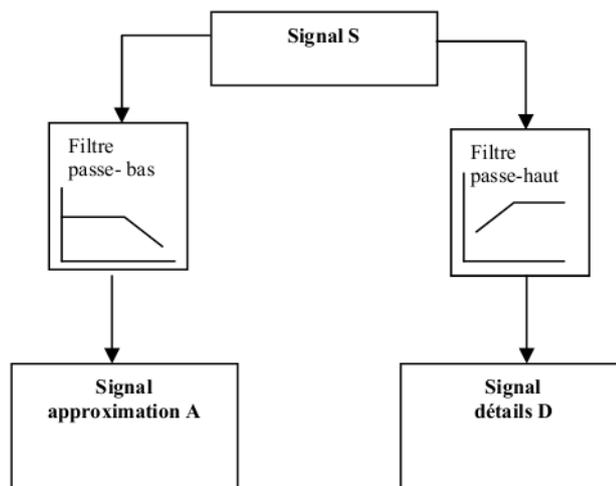


Figure II.3: décomposition du signal s en approximations et détails.

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 [17].

II.2.4.1. 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 [23]:

$$J_{ful} = \text{fix}(\log_2(N)) \quad (\text{II.7})$$

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 .

II.2.4.2. 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 [19,20].

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

$$\Psi(t) = e^{-\frac{t^2}{2}} * (e^{-j\omega t} - e^{-\frac{\omega^2}{2}}) \quad (\text{II.8})$$

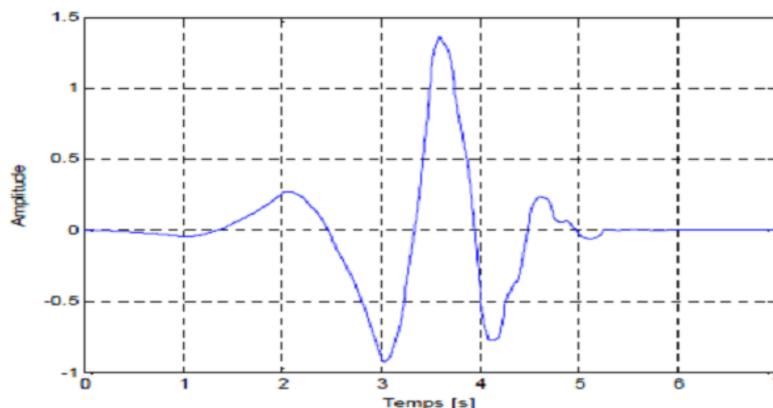


Figure II.4: Forme d'onde de l'ondelette de Daubechies 4.

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.

II.3. 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 (II.10). 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.

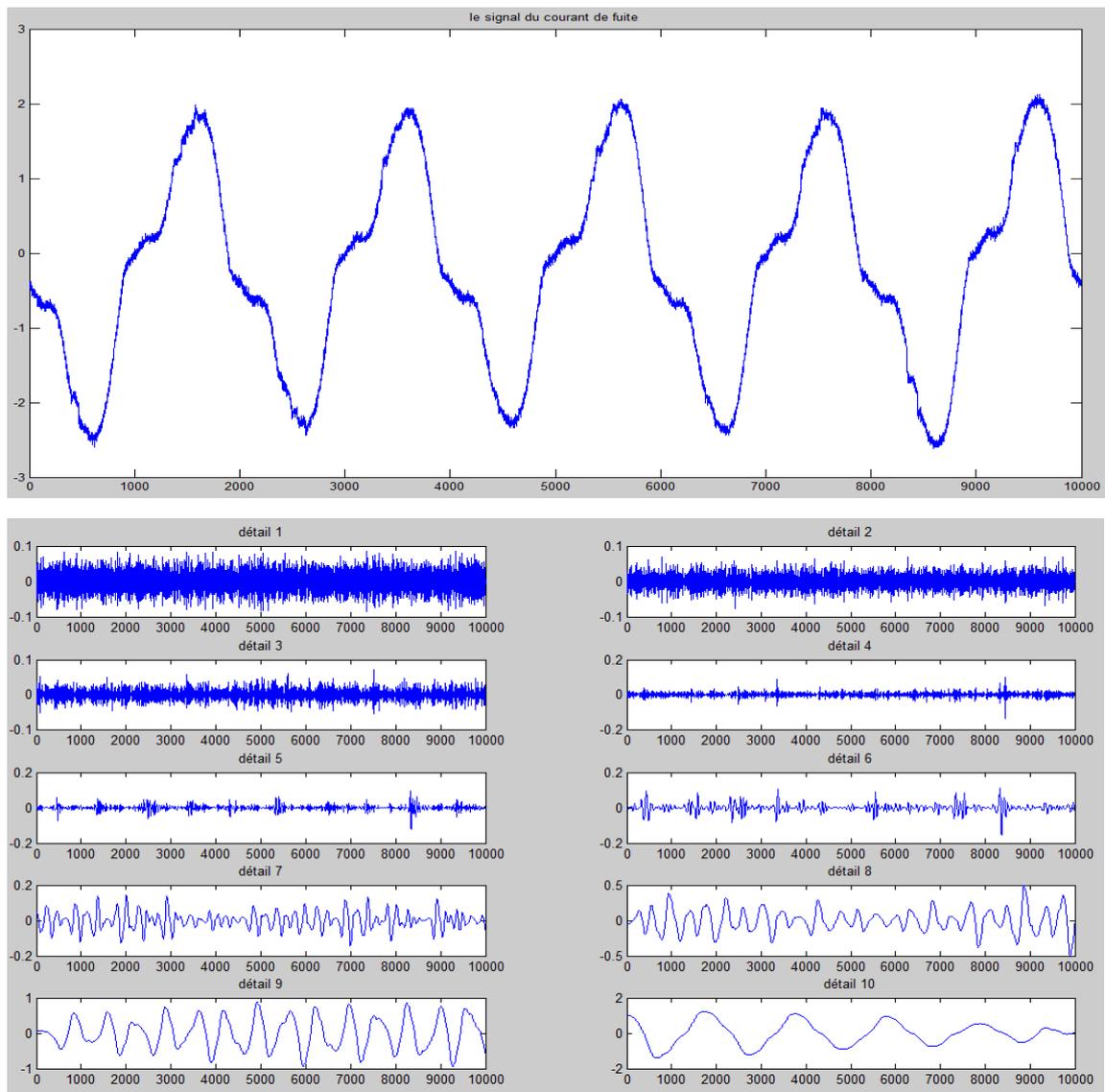


Figure II.5: Représentation de la forme d'onde du signal du courant de fuite et ses dix détails.

La base de données utilisée dans notre étude a été réalisée par Mr B.MOULA [22], 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.

II.4. 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 des maxima, minima, moyennes ainsi que la standard déviation (*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. La standard déviation (*écart type*) est calculée pour les détails afin d'identifier les transitions d'énergie présentes dans le signal du courant de fuite [24]. Elle est donnée par l'expression :

$$\delta = \sqrt{\frac{\sum_{i=1}^N (d_n - d_m)^2}{N-1}} \quad (\text{II.9})$$

δ : la standard déviation (*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.

II.4.1. Etude de la standard déviation 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 le standard déviation (*écart type*) pour les dix détails en utilisant la formule (II-7). Après avoir obtenu les valeurs de la standard déviation (*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.

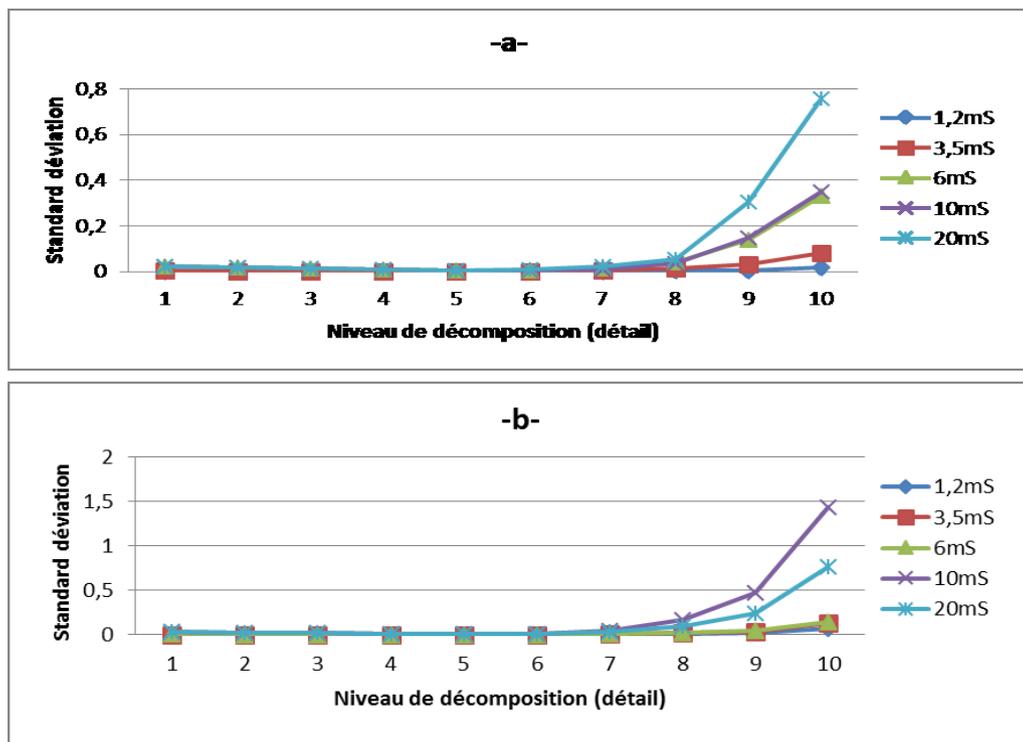


Figure II.6: Evolution de la standard déviation (écart type) des détails du signal du courant de fuite pour une tension appliquée de 5 kV_{eff}. a) pollution uniforme. b) pollution non

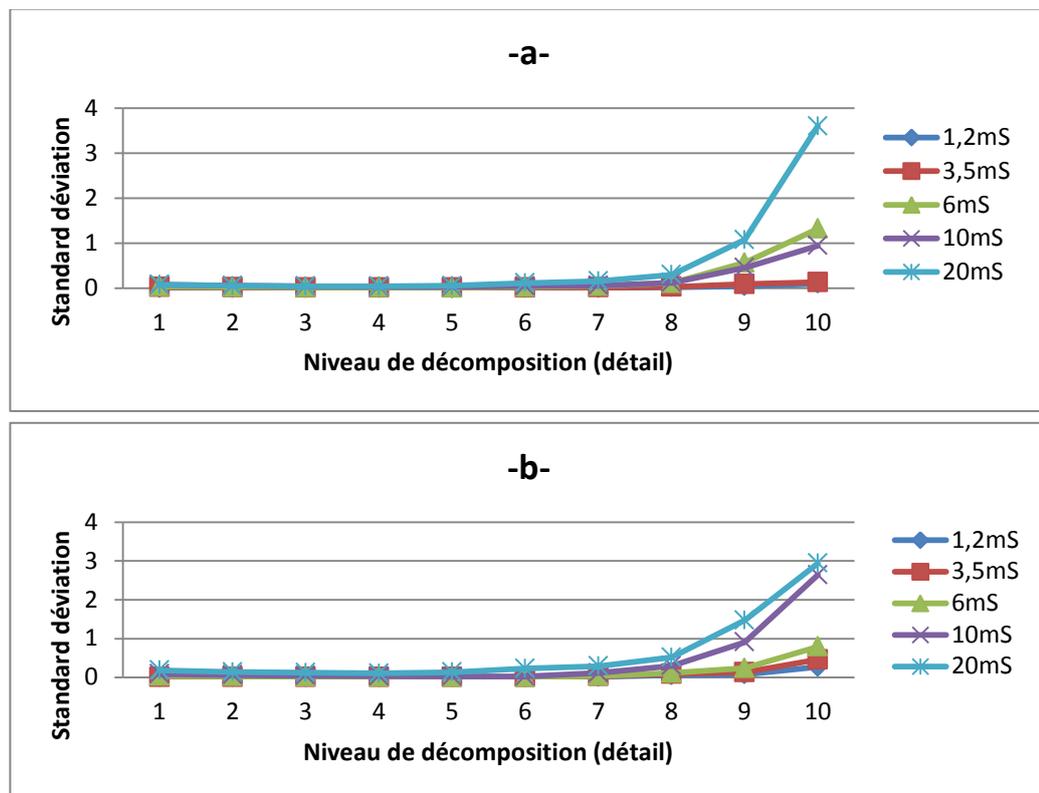


Figure II.7: Evolution de la standard déviation (écart type) des détails du signal du courant de fuite pour une tension appliquée de 15 kV_{eff}. a) pollution uniforme. b) pollution non

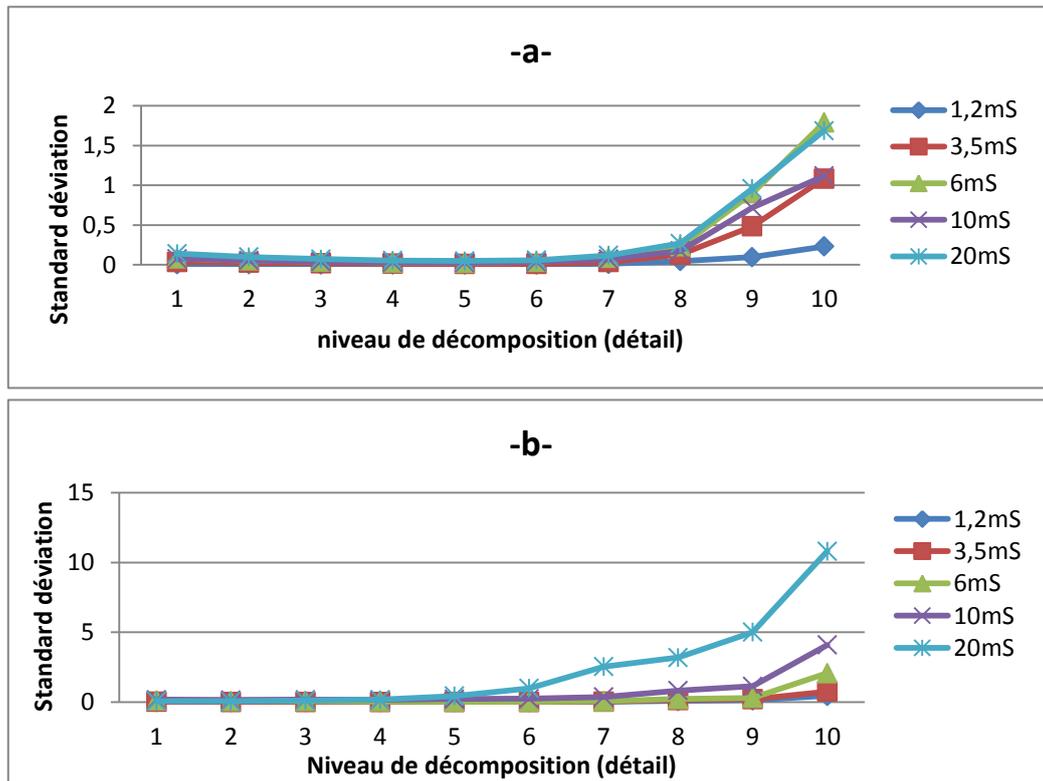


Figure II.8: Evolution de la standard déviation (écart type) des détails du signal du courant de fuite pour une tension appliquée de $25 \text{ kV}_{\text{eff}}$. a) pollution uniforme. b) pollution non uniforme.

Interprétation :

Nous constatons que les valeurs de la standard déviation augmentent avec le niveau de décomposition. Cette augmentation est vérifiée pour tous les niveaux de tension appliqués ainsi pour les deux types de pollution. La standard déviation est d'autant plus importante que la conductivité volumique de la solution polluante est élevée. Nous observons que l'augmentation de la standard déviation 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 la standard déviation peuvent être considérées comme information importante sur le degré de sévérité de pollution car leurs valeurs sont liées à la conductivité volumique de l'agent polluant.

II.4.2. Etude des maximums 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 maximums 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.

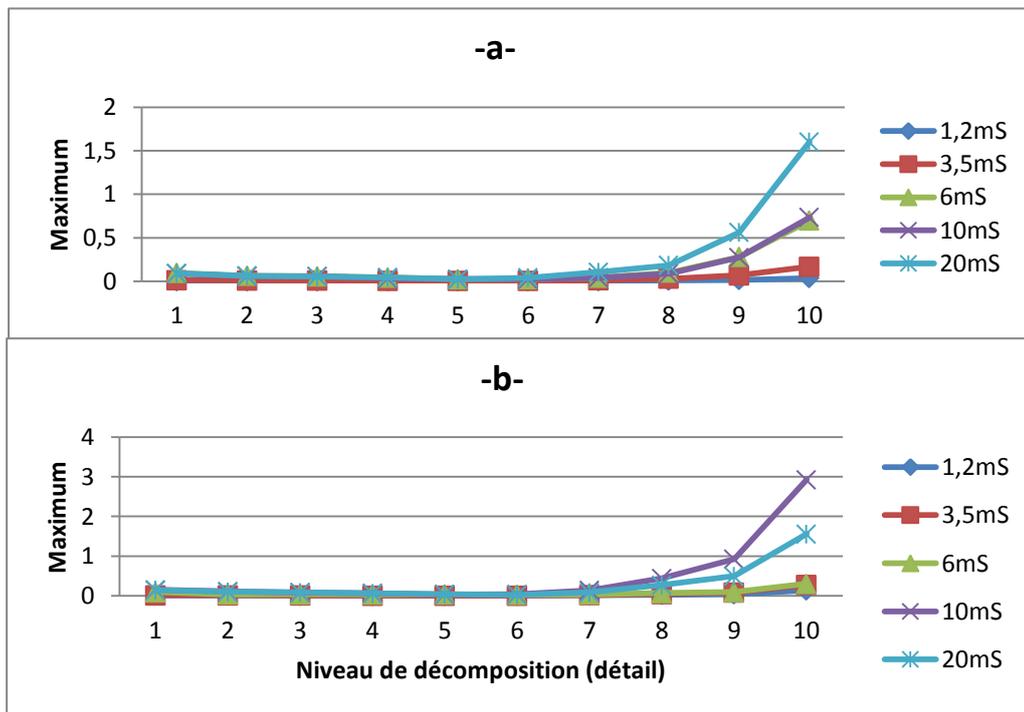


Figure II.9: Evolution des valeurs des maxima des détails du signal du courant de fuite pour une tension appliquée de 5 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

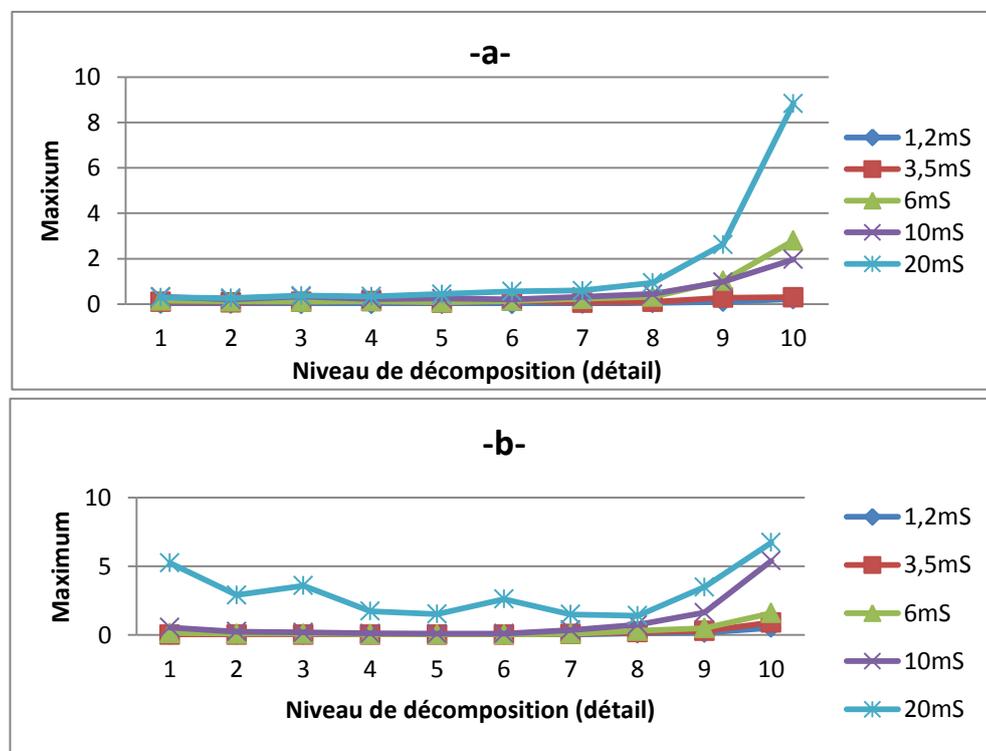


Figure II.10: Evolution des valeurs des maxima des détails du signal du courant de fuite pour une tension appliquée de 15 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

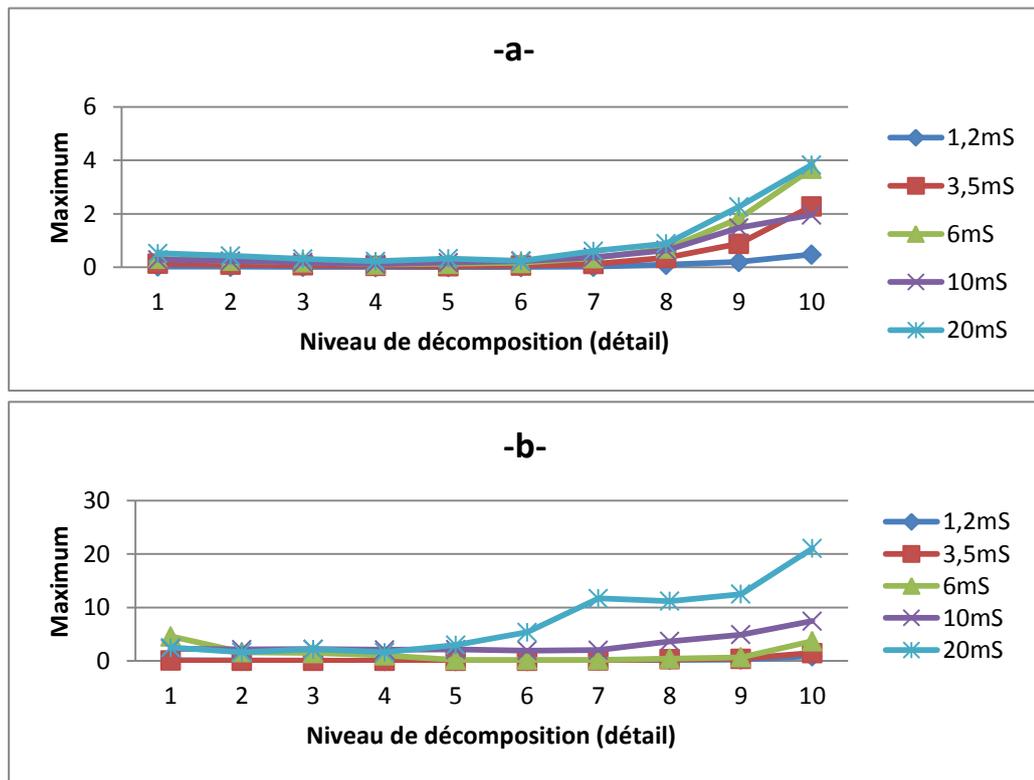


Figure II.11: Evolution des valeurs des maximums des détails du signal du courant de fuite pour une tension appliquée de 25 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

Interprétation :

Quelques soit les niveaux de tension appliquée, les valeurs des maximums en générale 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 maximums 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 maximums 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.

II.4.3. Etude des minimums 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 minimums 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.

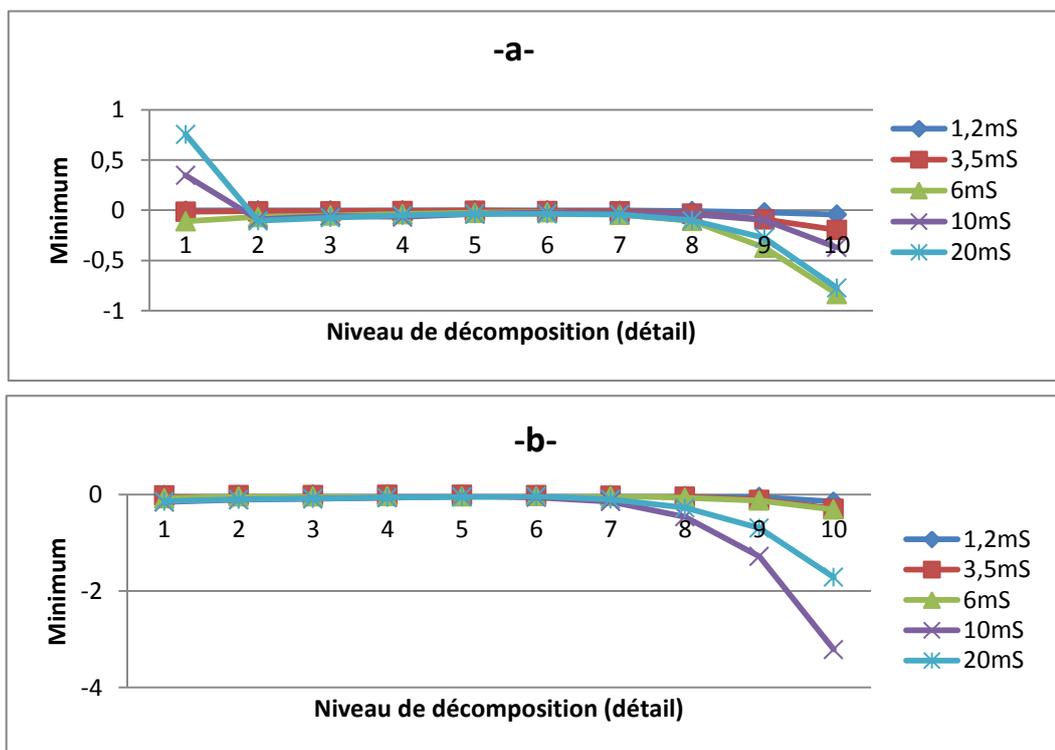


Figure II.12: Evolution des valeurs des minimums des détails du signal du courant de fuite pour une tension appliquée de 5 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

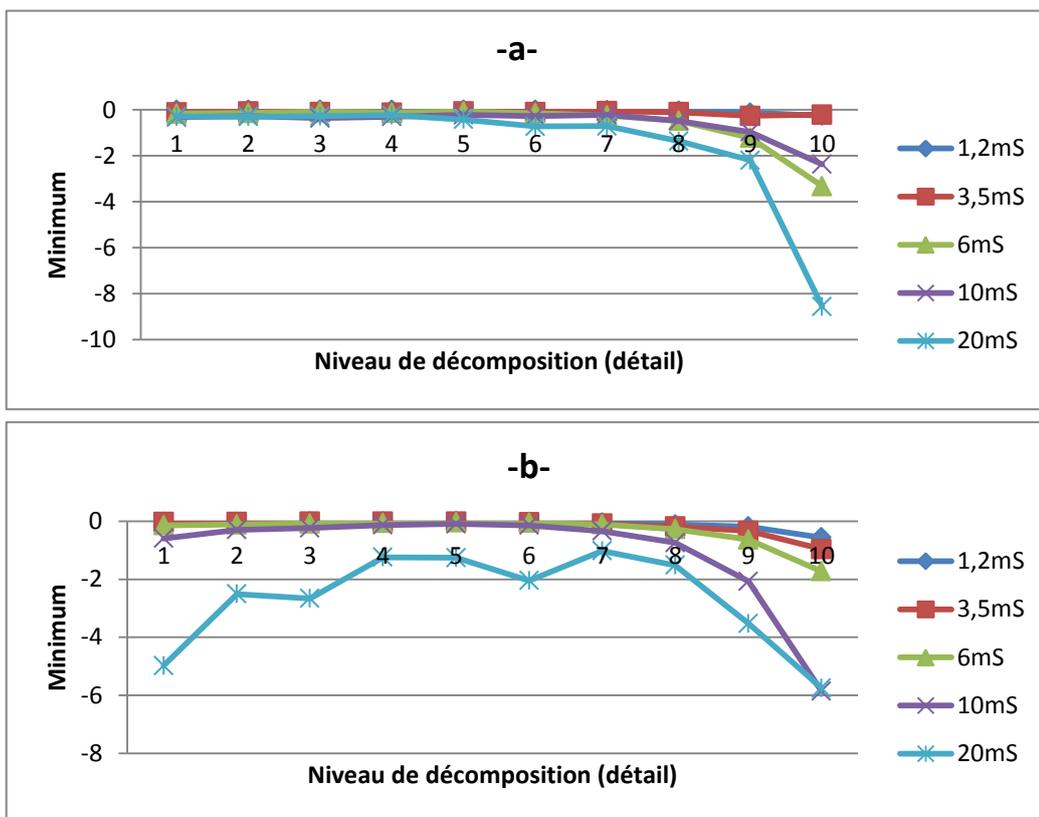


Figure II.13: Evolution des valeurs des minimums des détails du signal du courant de fuite pour une tension appliquée de 15 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

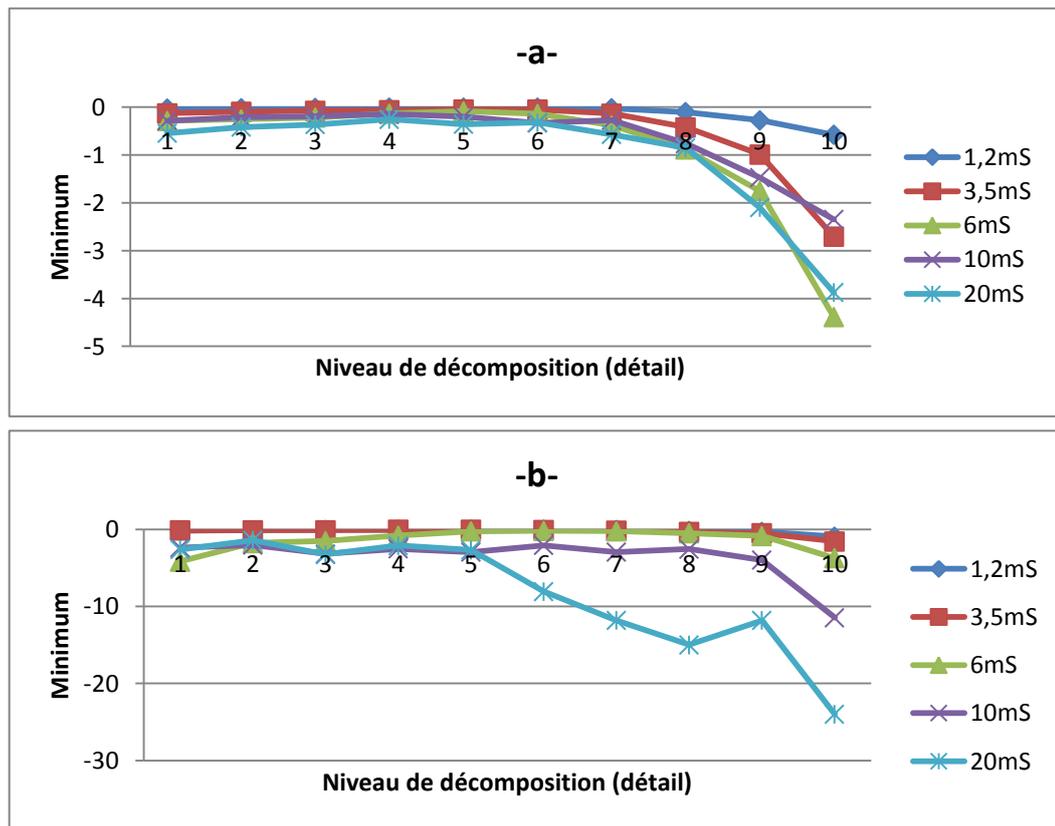


Figure II.14: Evolution des valeurs des minimums des détails du signal du courant de fuite pour une tension appliquée de 25 kV_{eff}. a) pollution uniforme. b) pollution non uniforme.

Interprétation :

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

II.4.4. 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.

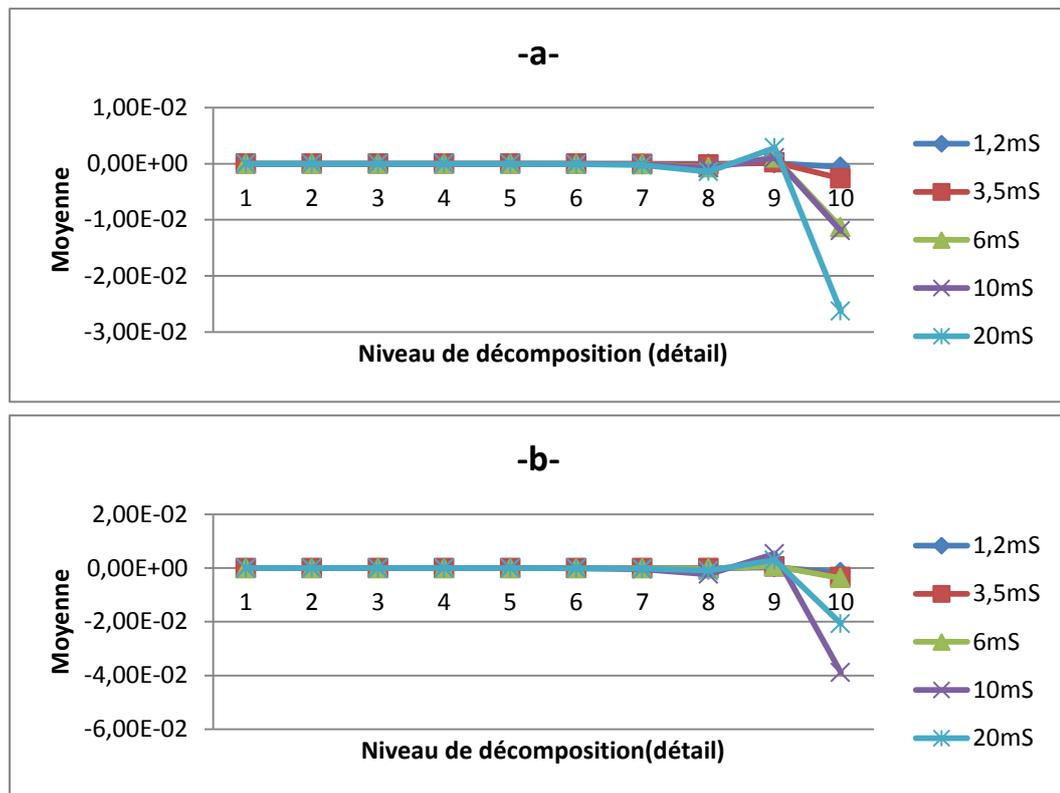


Figure II.15: Evolution des valeurs des moyennes des détails du signal du courant de fuite pour une tension appliquée de 5 kVeff. a) pollution uniforme. b) pollution non uniforme.

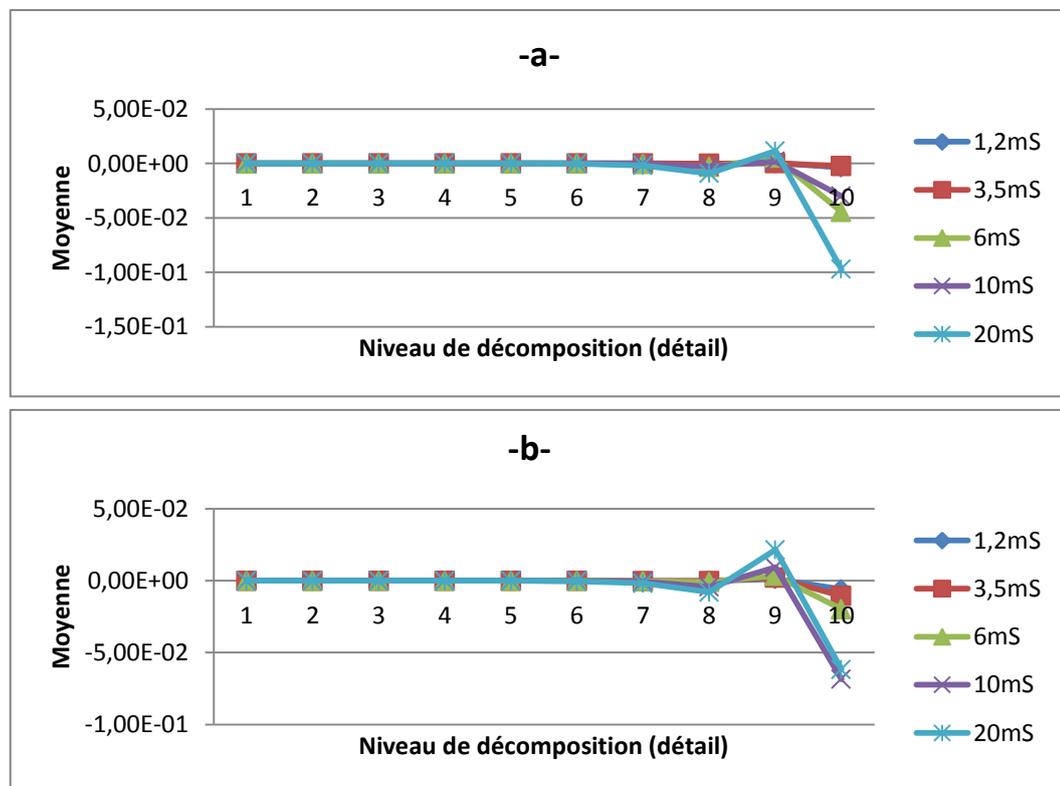


Figure II.16: Evolution des valeurs des moyennes des détails du signal du courant de fuite pour une tension appliquée de 15 kVeff. a) pollution uniforme. b) pollution non uniforme.

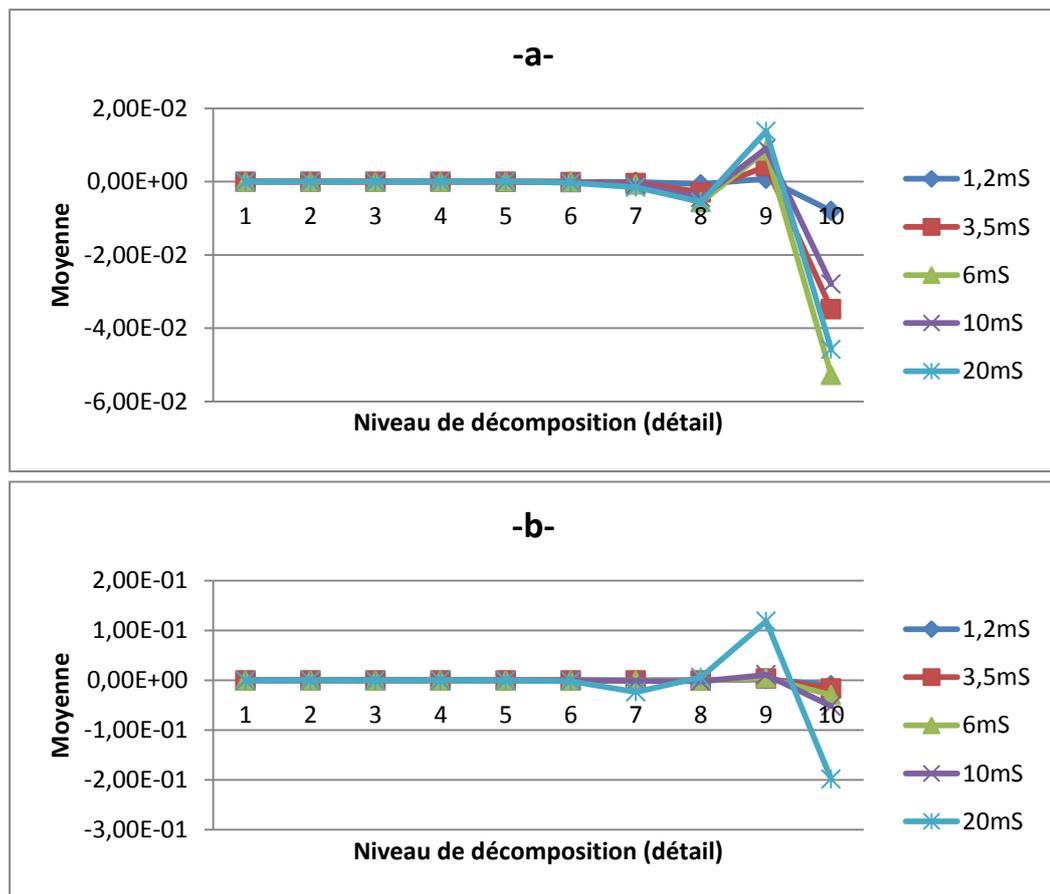


Figure II.17: Evolution des valeurs des moyennes des détails du signal du courant de fuite pour une tension appliquée de 25 kVeff. a) pollution uniforme. b) pollution non uniforme

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'utiliser comme une quatrième information pour la prédiction de l'état de surface des isolateurs.

II.5. 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 [59], 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 [59].

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 [59].

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 [59].

En tenant compte de la base de données de B. Moula [22] et des normes définies ci-dessus, nous proposons un tableau (Tableau IV-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.6. Application de l'analyse linéaire discriminante (LDA) pour la classification de la sévérité de pollution des isolateurs :

II.6.1. 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 (1×40), 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 standards déviations, les minimums, les maximums 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
---	---	---	---

II.6.2. 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 [22] 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.6.3. 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 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 LDA.

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

II.7. Validation et résultat de classification :

Les résultats obtenus par la classification en utilisant l'Analyse Discriminante Linéaire sont données dans les deux tableaux suivants:

Echantillon n°	Conductivité (mS/cm)	Classe assignée	Résultat obtenu
1	1.2	1	1
2	1.2	1	1
3	3.5	1	1
4	3.5	1	3*
5	6	2	1*
6	6	2	3*
7	6	2	2
8	10	3	2*
9	10	3	3
10	20	3	3
11	20	3	2*
Exactitude de classification (%)			54.54

Tableau II.2: Résultats de la classification en utilisant la LDA dans le cas de pollution uniforme.

Echantillon n°	Conductivité (mS/cm)	Classe assignée	Résultat obtenu
1	1.2	1	1
2	3.5	1	1
3	3.5	1	2*
4	1.2	1	1
5	6	2	2
6	6	2	1*
7	10	3	2*
8	20	3	1*
9	20	3	2*
Exactitude de classification (%)			44.44

Tableau II.3: Résultats de la classification en utilisant la LDA dans le cas de pollution

Interprétation des résultats :

La classification avec l'analyse discriminante linéaire dans le cas de la pollution uniforme a donné six bonnes classifications sur un ensemble de onze échantillons ce qui correspond à 54.54%.

En ce qui concerne la pollution non uniforme, la classification a donné quatre bonnes classifications sur un ensemble de neuf échantillons donc un pourcentage d'exactitude de classification de 45.44%.

Conclusion :

Dans ce chapitre nous avons fait une 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.

Ensuite, nous avons utilisé la LDA pour faire la classification du degré de sévérité de pollution des isolateurs. Les résultats de classification étaient moyennes avec des pourcentages de 54.54% et 45.44% pour la pollution uniforme et non uniforme respectivement.

Finalement, on peut dire que l'analyse discriminante linéaire (LDA) est une méthode d'apprentissage supervisé qui peut être utilisée pour la classification sauf qu'elle donne des résultats de classification moyennes.

Chapitre III:

*Caractérisation de
l'état de surface des
isolateurs par la
technique des
ondelettes*

Introduction

Dans ce chapitre, nous rappelons la théorie des ondelettes continue (CWT) afin de l'appliquer pour la caractérisation de l'état de surface des isolateurs pollués. En premier lieu, nous exploitons les signaux de la tension appliquée pour tracer les scalogrammes et les courbes des coefficients d'ondelettes pour différentes conductivités (3.5, 10 et 20 mS/cm).

En deuxième lieu, nous traçons les scalogrammes et les courbes des coefficients d'ondelettes pour différentes concentrations en kaolin (40, 100 et 200 g/l) en gardant la conductivité volumique et la tension appliquée constantes.

III.1. La transformée en ondelettes continues (CWT) :

La transformée en ondelette continue (CWT) utilise des ondelettes dont l'enveloppe rétrécit avec l'échelle en gardant le même nombre d'oscillations. Fondamentalement, la transformée en ondelette se calcule de la même manière que la transformée de Fourier fenêtrée, autrement dit on regarde la corrélation au sens du produit scalaire entre le signal à analyser et la famille de fonction $\Psi_{a,b}(t)$ [25, 26, 27].

Une ondelette est une onde qui oscille à fréquence constante et qui s'annule au bout d'un instant donné. La fonction analysante peut être choisie en toute liberté contrairement à la transformée de Fourier qui se restreint uniquement aux fonctions sinusoïdales.

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{b}} \Psi\left(\frac{t-a}{b}\right) \quad (u \in R_+, S) \quad (\text{III.1})$$

Ψ : l'ondelette mère

a : l'échelle de l'ondelette

b : la translation de l'ondelette

III.1.1. Scalogramme :

Le scalogramme d'un signal est défini par : $P_w = (b, \xi) = |W_f(a, b)|^2 = |W_f(a, \frac{n}{\xi})|^2$ (III.2)

Le scalogramme normalisé vaut : $\frac{n}{\xi} \cdot P_w f(b, \xi)$ (III.3)

La normalisation revient à multiplier l'expression par un facteur égal à $\frac{\xi}{n}$. n est le

centre de fréquence de l'ondelette élémentaire qui a une énergie concentrée dans un intervalle positif centré en n/s dont la taille est donnée par $1/s$. Dans le domaine temps-fréquence de l'analyse par ondelettes, un atome d'ondelette (Figure II.12) est représenté par le rectangle de Heisenberg de largeur temporelle $s * \sigma_t$ et de largeur fréquentielle $\frac{1}{s * \sigma_w}$; σ_t et σ_w correspondent aux dimensions du rectangle de Heisenberg pour l'ondelette mère. La largeur temporelle et fréquentielle restent constantes pour toute la transformée car :

$$S_{\text{rect}} = \sigma_w * \sigma_t.$$

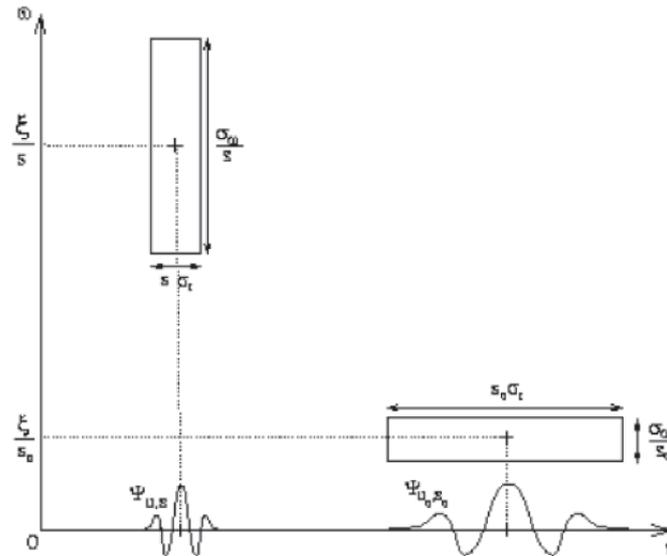


Figure III.1: Atome de la transformée en ondelettes.

III.1.2. Calcul des coefficients de la CWT :

La procédure de calcul des coefficients s'effectue en multipliant le signal par l'ondelette analysante puis en calculant l'intégrale du produit. Ce processus est accompli en quatre étapes [18]:

- Dès le départ, on compare l'ondelette à une portion du signal.
- On calcule le coefficient C qui étudie le degré de similitude entre la portion du signal et l'ondelette analysante. Ces derniers sont similaires lorsque la valeur numérique du coefficient C est élevée. Plus précisément, lorsque l'énergie du signal et celle de l'ondelette sont relativement proches, le coefficient C peut être interprété comme un facteur de corrélation.

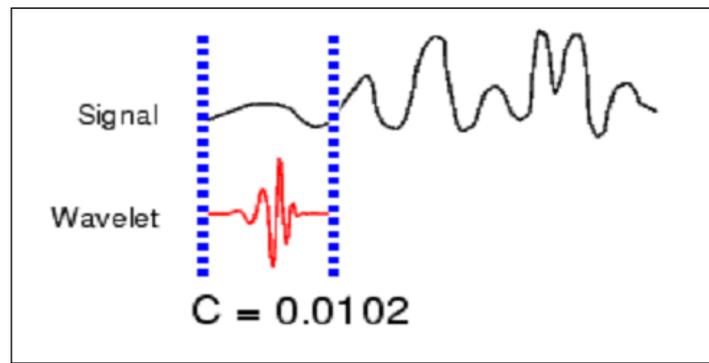


Figure III.2: Comparaison de l'ondelette à une portion du signal

- On translate l'ondelette vers la droite jusqu'au balayage total du signal (Figure III.3).

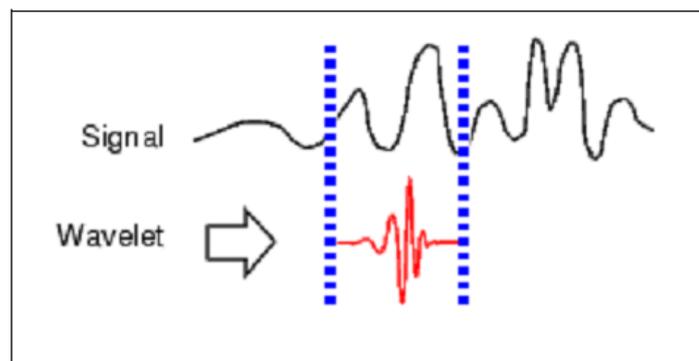


Figure III.3: Balayage du signal par l'ondelette analysante.

- On passe ensuite à une échelle plus élevée qui induit une dilatation temporelle de l'ondelette analysante. Cela a pour effet d'augmenter la portion de comparaison (Figure III.4).

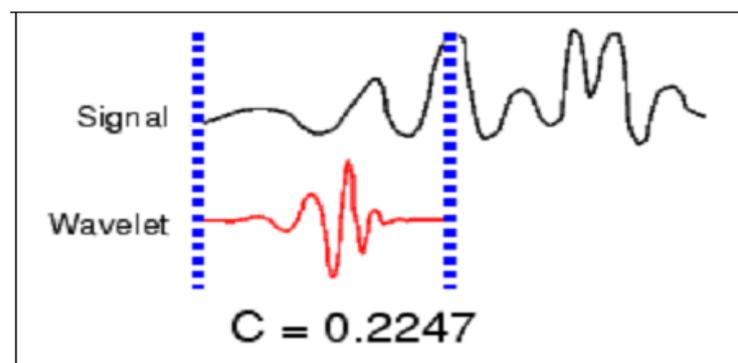


Figure III.4: Dilatation de l'ondelette et balayage du signal.

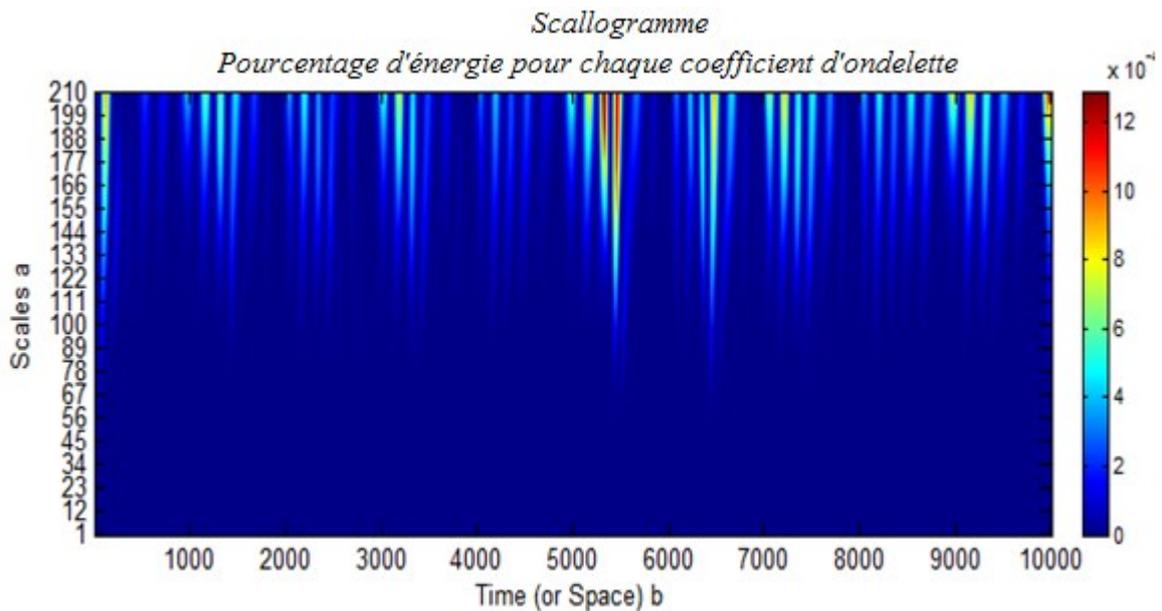
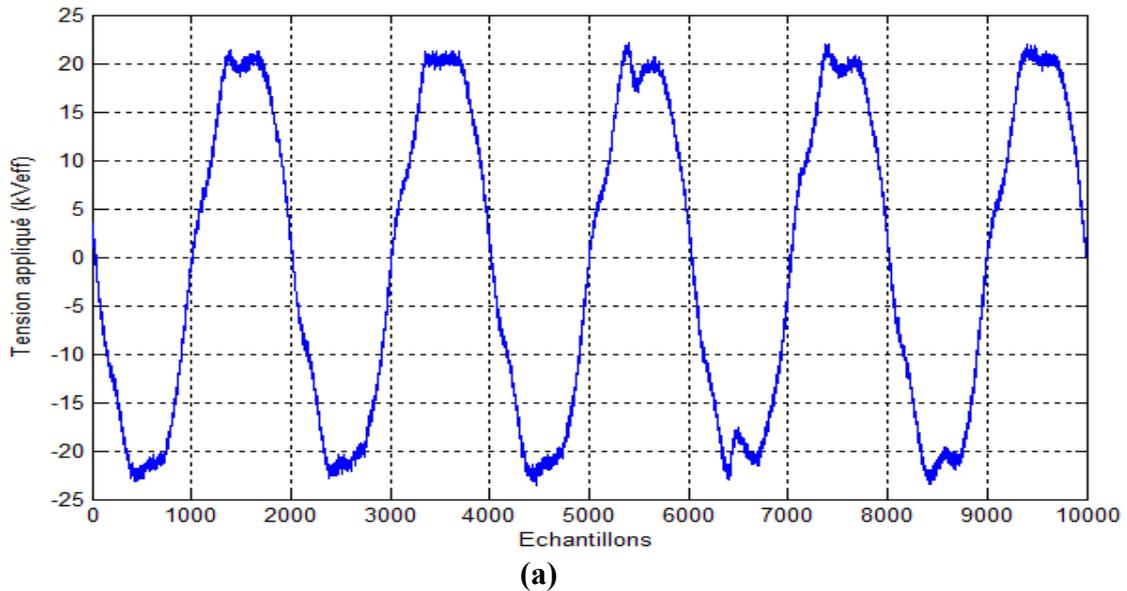
Cette opération est répétée pour chaque échelle jusqu'au calcul des coefficients pour toutes les échelles et pour différentes portions du signal. Les coefficients constituent le résultat d'une régression effectuée sur le signal original.

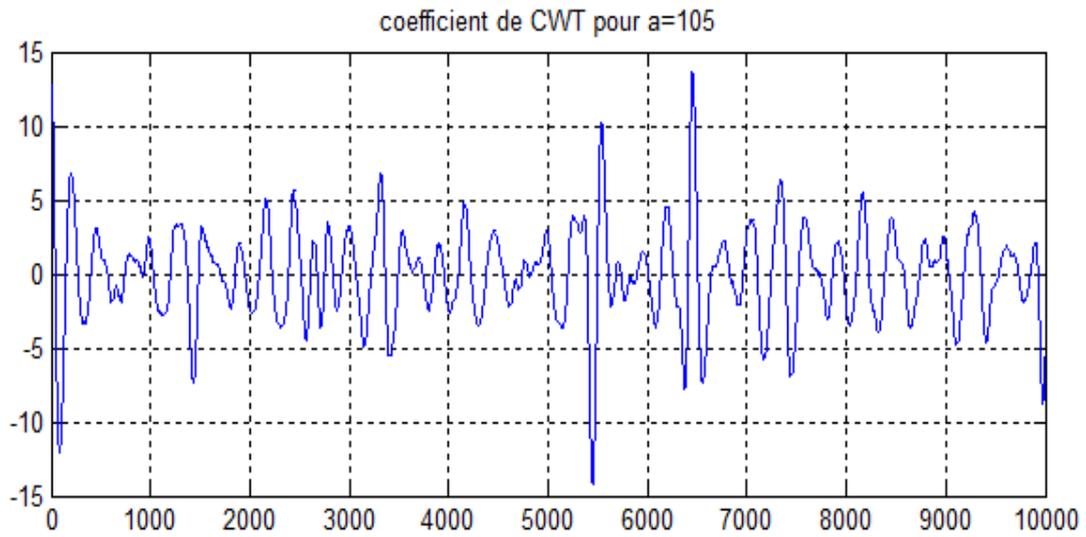
III.2. Pollution uniforme :

Le traitement des signaux de la tension appliquée est effectué en utilisant la technique de la CWT.

III.2.1 Influence de la conductivité volumique:

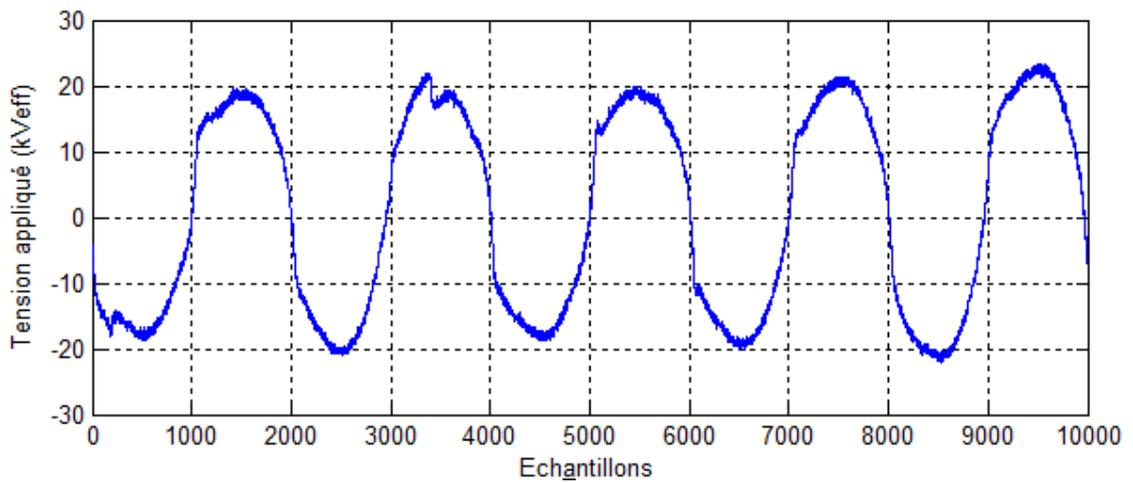
Pour étudier l'influence de la conductivité volumique on a tracé les scalogrammes des signaux du courant de fuite pour trois conductivités différentes 3.5, 10 et 20 mS, en fixant la concentration du kaolin (100 g/l) et la tension appliquée (15 kV_{eff}).



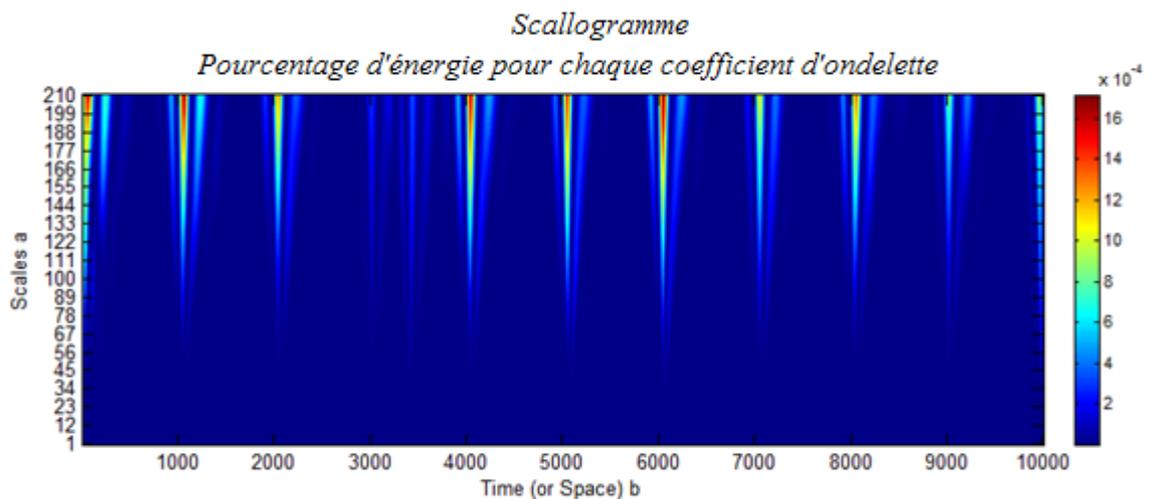


(c)

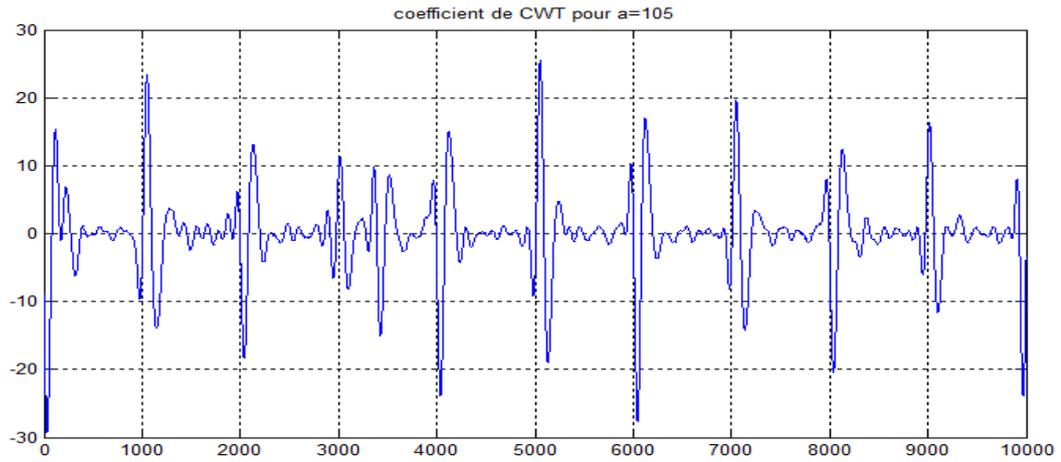
Figure III.5 : (a) Forme d'onde de la tension appliquée (b) Scalogramme correspondant (c) Coefficients de la CWT pour $a = 105$ Conductivité 3.5 mS/cm.



(a)

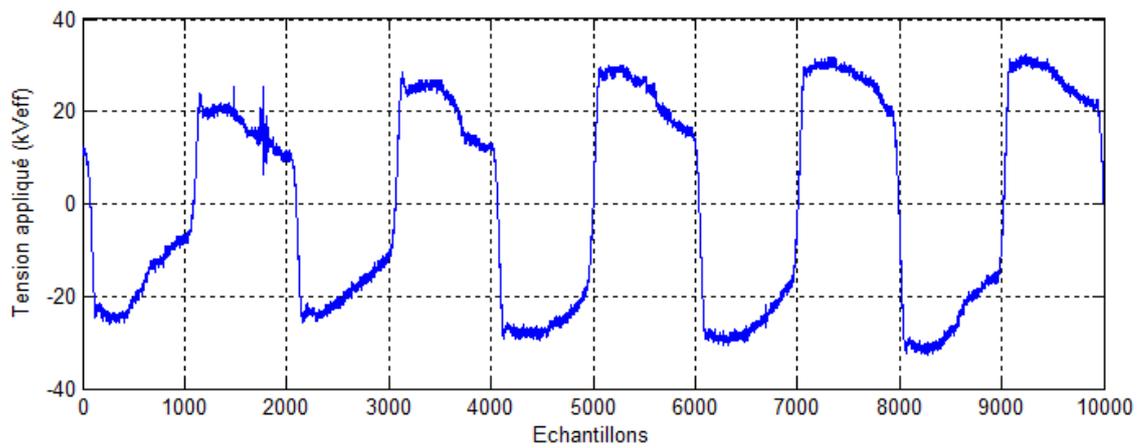


(b)

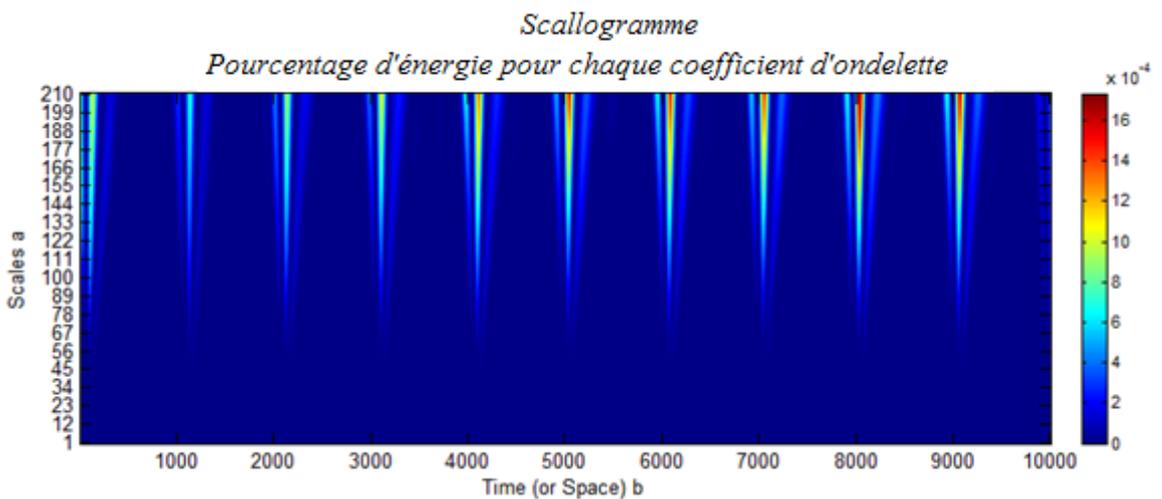


(c)

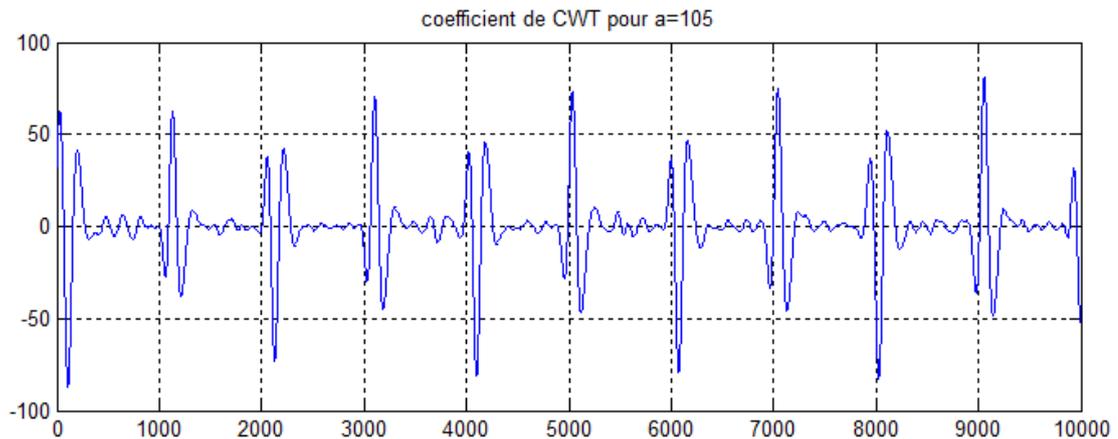
Figure III.6 : (a) Forme d'onde de la tension appliquée (b) Scalogramme correspondant (c) Coefficients de la CWT pour $a = 105$. Conductivité 10 mS/cm



(a)



(b)



(c)

Figure III.7: (a) Forme d'onde de la tension appliquée (b) Scalogramme correspondant (c) Exemple des Coefficients de la CWT pour $a = 105$. Conductivité 20 mS/cm

Discussion des résultats :

Conductivité volumique 3.5mS/cm :

Pour la conductivité de 3.5 mS/cm, nous remarquons que le signal de la tension appliquée affiche des distorsions au niveau de la crête et au voisinage de l'axe des temps. Son scalogramme montre que les coefficients ont des valeurs maximales (couleur rouge intense). Pour des échelles moyennes et grandes (de 133 à 210) c-à-d des perturbations de basses et moyennes fréquences apparaissent. Cela signifie que ces perturbations oscillent à des fréquences faibles et moyennes indiquant que l'activité des décharges parallèles à la surface est moindre.

Les coefficients de la CWT pour l'échelle $a = 105$ montrent qu'une petite corrélation entre l'oscillation de l'ondelette de Daubechies 4 et celle de la transformée en ondelette de la tension. Cette corrélation est plus en plus évidente dans les deux points qui apparaissent au milieu du signal de la CWT et aux extrémités.

Conductivité volumique 10 mS/cm :

Pour la conductivité de 10 mS/cm, nous remarquons des distorsions dans tout le signal de la tension appliquée plus précisément dans les crêtes. Son scalogramme montre que les coefficients ont une valeur maximale dans plusieurs régions du signal pour des échelles plus faibles (de 122 à 210). Cela est expliqué par l'apparition des fréquences un peu plus élevées indiquant l'augmentation des décharges partielles au niveau de la surface du modèle. L'apparition de ces décharges partielles, engendre l'assèchement du modèle.

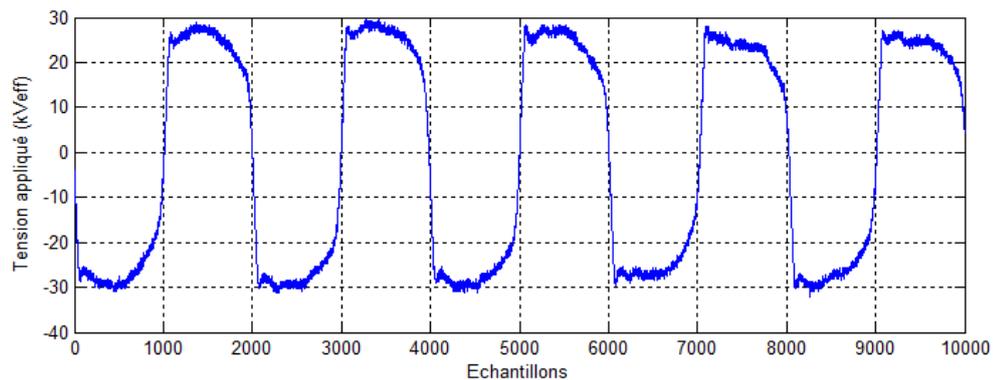
Les coefficients de la CWT pour une échelle de $a=105$ montre une similitude entre l'ondelette analysante Daubechies 4 et la transformée en ondelette continue du signal étudié.

Conductivité volumique 20 mS/cm :

Pour cette conductivité, nous constatons que les perturbations sur l'onde de la tension appliquée s'intensifient avec des chutes de tension plus remarquables au niveau de la crête. Le scalogramme montre que ces chutes de tension sont causées à des instants espacés de la même manière ou on peut remarquer des coefficients très élevés de la CWT pour des grandes échelles allons de 144 à 210. Cela est expliqué par l'apparition des basses fréquences et donc des perturbations et des décharges moindre à la surface.

III.2.2. Influence de la concentration en kaolin :

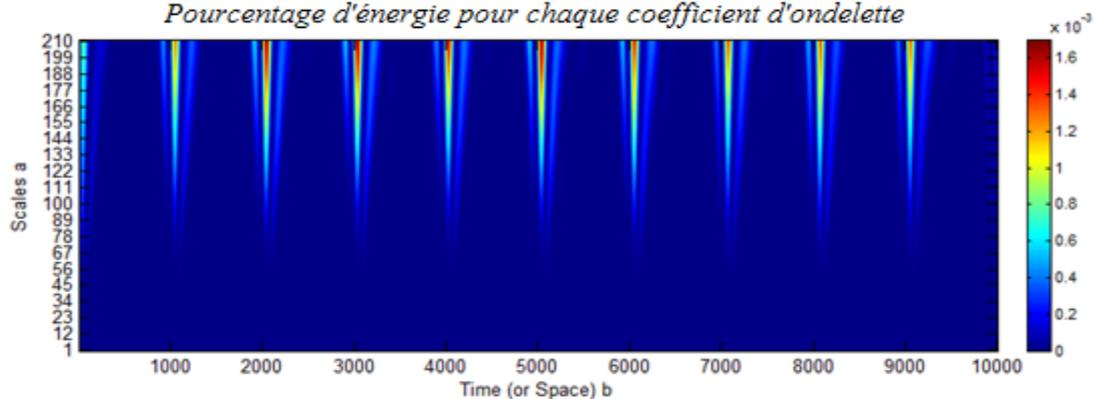
Pour étudier l'influence de la concentration en kaolin nous, avons tracé les scalogrammes des signaux de la tension appliquée pour trois concentrations différentes 40, 100 et 200 g/l, en fixant la conductivité volumique à 20 mS/cm.



(a)

Scalogramme

Pourcentage d'énergie pour chaque coefficient d'ondelette



(b)

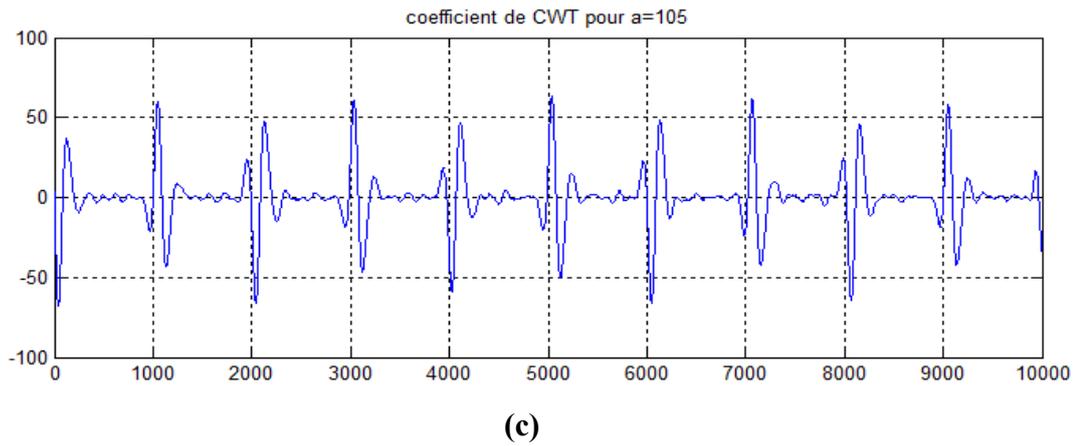
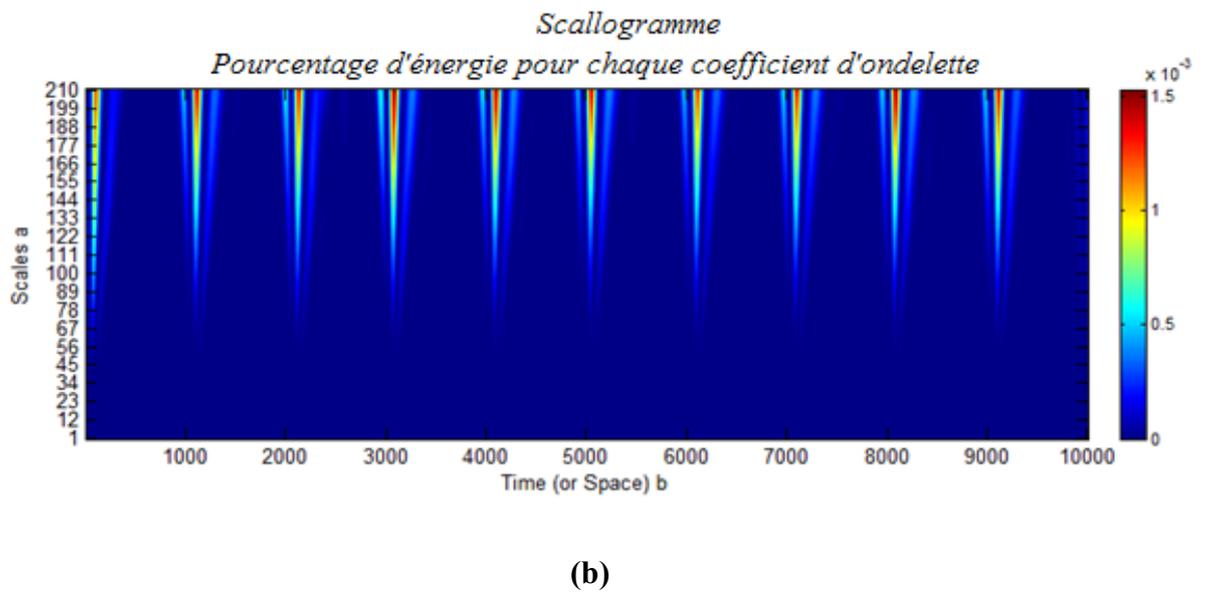
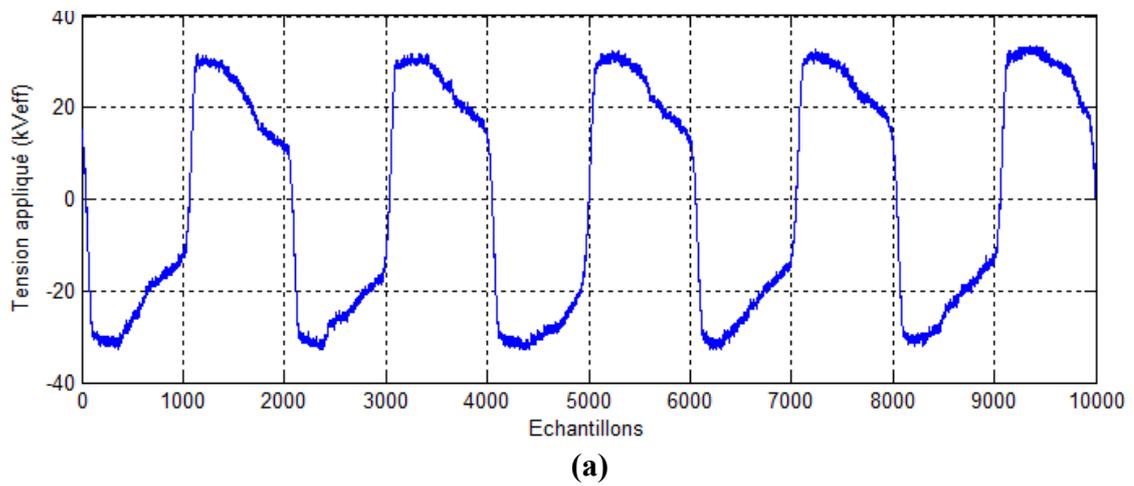
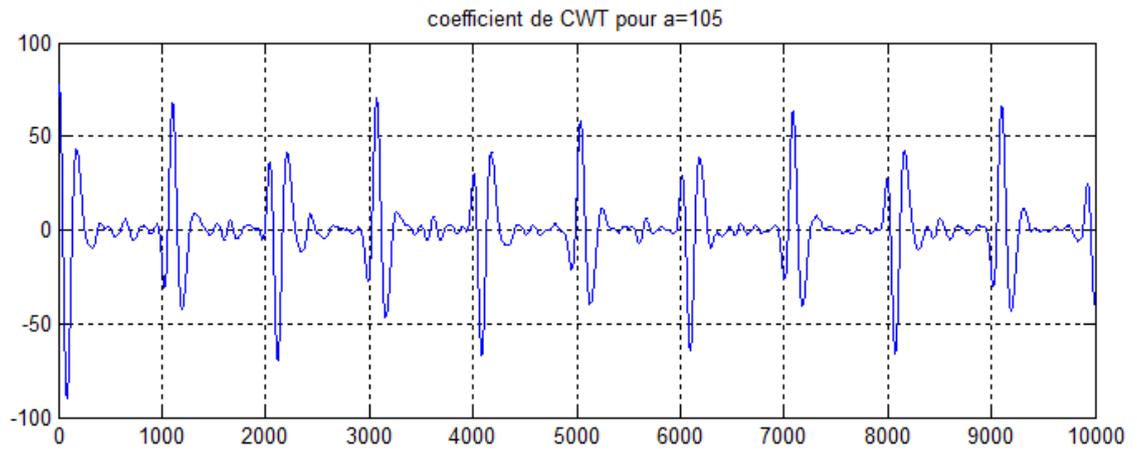


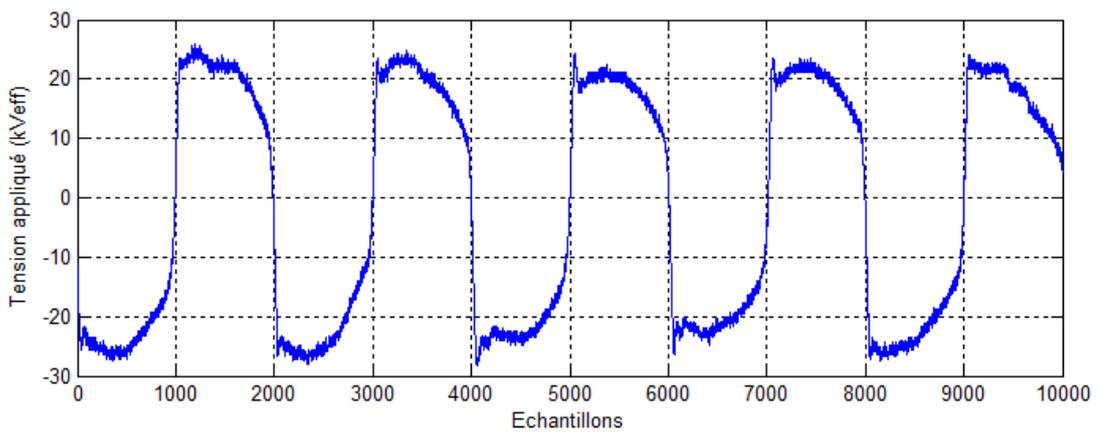
Figure III.8 : Concentration en kaolin de 40g/l : (a) Forme d'onde de la tension appliquée 25 kV_{eff} (b) Scalogramme correspondant (c) Exemple des Coefficients de la CWT pour a = 105



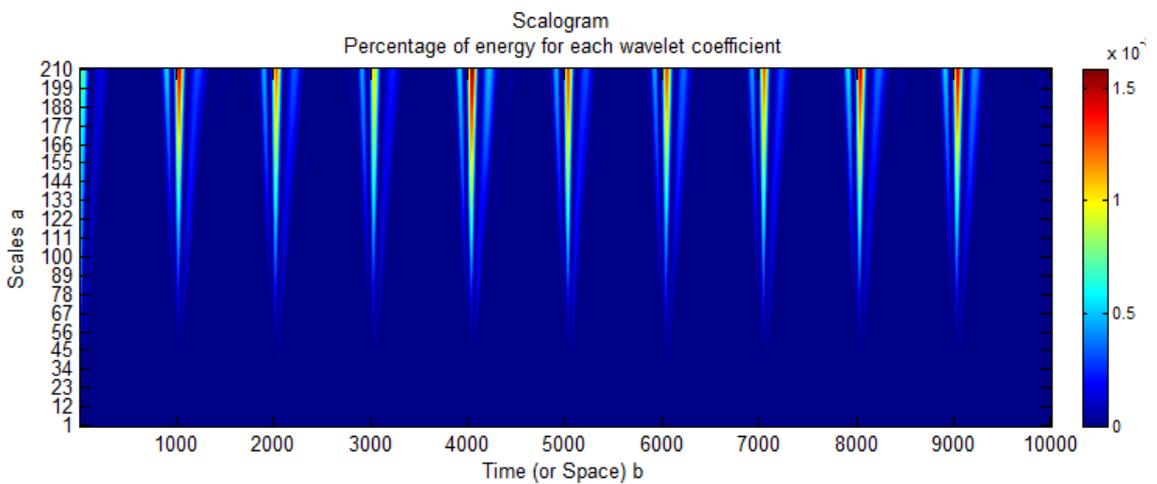


(c)

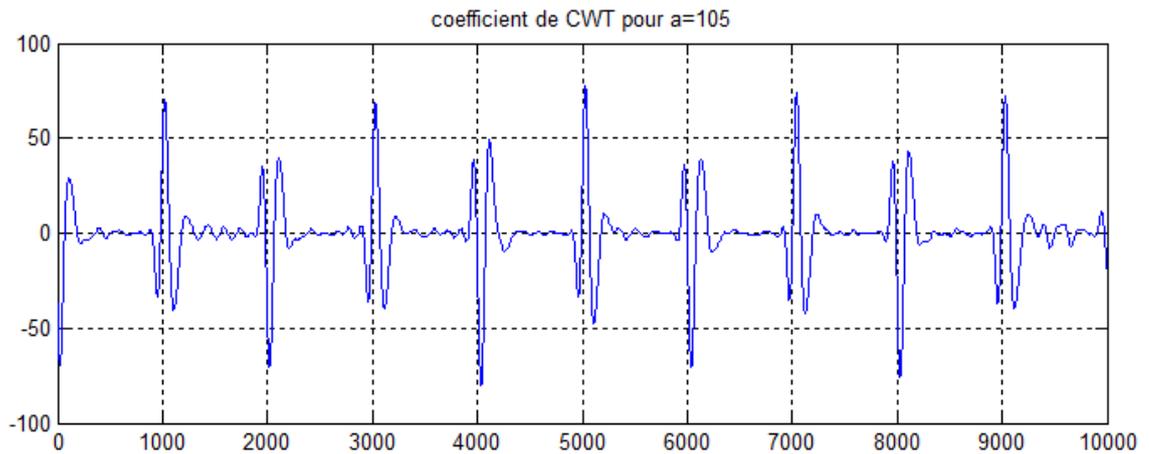
Figure III.9 : Concentration en kaolin de 100g/l : (a) Forme d'onde de la tension appliquée 20 kV_{eff} (b) Scalogramme correspondant (c) Coefficients de la CWT pour $a = 105$



(a)



(b)



(c)

Figure III.10 : Concentration en kaolin de 200 g/l : (a) Forme d'onde de la tension appliquée $25 kV_{eff}$ (b) Scalogramme correspondant (c) Coefficients de la CWT pour $a = 105$

Discussion des résultats :

Concentration en kaolin de 40g/l :

Pour une conductivité de 20mS/cm et une concentration en kaolin de 40g/l, les distorsions occupent pratiquement toutes les régions du signal de la tension appliquée. Le scalogramme indique que les coefficients de la CWT sont très intenses pour des échelles plus élevés à partir de 166 à 210, ce qui justifie l'apparition des décharges partielles de basses fréquences.

Les coefficients de la CWT pour une échelle de $a=105$ montrent une similitude entre l'ondelette analysante Daubechies 4 et la transformée continue du signal appliqué au niveau des crêtes du signal de la tension appliquée ou les distorsions sont plus visibles.

Concentration de kaolin 100g/l :

Pour cette concentration en kaolin de 100g/l et avec une conductivité élevée, les perturbations sont remarquables dans toutes les régions du signal de la tension appliquée. Le scalogramme indique que les coefficients de la CWT sont très intenses pour des échelles un peu élevés entre 155 et 210, ce qui est expliqué par l'apparition des décharges partielles de basses fréquences.

Les coefficients de la CWT pour une échelle de $a=105$ montrent une corrélation entre l'ondelette analysante Daubechies 4 et la CWT de la tension appliquée.

Concentration de kaolin 200g/l :

Pour une concentration en kaolin de 200g/l, les distorsions occupent pratiquement toutes les régions du signal de la tension appliquée. Le scalogramme indique que les coefficients de la CWT sont très intenses pour des échelles plus réduites à partir de 144. Ce qui peut être traduit par l'apparition des décharges partielles de basses et moyennes fréquences.

Les coefficients de la CWT pour une échelle $a=105$ montre une bonne corrélation entre l'ondelette analysante Daubechies 4 et la transformée continue du signal au niveau des crêtes positives et négatives du signal de la tension appliquée.

L'apparition des fréquences un peu plus élevées indiquent l'augmentation des décharges partielles au niveau de la surface de l'isolateur pollué.

Conclusion :

Le scalogramme représente une projection des coefficients calculés pour des échelles données sur un plan orthogonal. D'après l'étude faite dans ce chapitre, l'application de la CWT sur les signaux de la tension appliquée et le scalogramme montre leur efficacité pour le diagnostic de la sévérité de pollution des isolateurs.

Concernant l'analyse par le scalogramme, nous concluons que l'augmentation de la conductivité peut être observée par la couleur des coefficients qui spécifient leurs valeurs selon qu'elles soient faibles ou élevées. Un coefficient élevé pour une grande échelle indique une oscillation du signal à basse fréquence et inversement pour une échelle de plus petite valeur. Dans notre cas, la fréquence des perturbations occasionnées sur le signal de la tension augmente avec la sévérité de pollution.

Le diagnostic de l'état de surface des isolateurs peut être fait à travers une échelle bien déterminée (105 dans notre cas) qui nous renseigne sur la pollution du modèle grâce aux coefficients de cette même échelle. Lorsque la corrélation entre la CWT du signal et l'ondelette analysante est quasi inexistante avec une faible amplitude, la conductivité

de la surface est faible. Cette corrélation est autant plus apparente avec l'augmentation de la conductivité volumique qui fait augmenter l'amplitude des coefficients de la CWT ainsi que le nombre d'oscillations qui ressemblent beaucoup à celles de l'ondelette choisie.

La concentration du kaolin n'as pas une grande influence sur l'amplitude des coefficients de la CWT. Pour les trois cas de concentration de kaolin étudiés, nous concluons aussi qu'il y a une corrélation entre l'ondelette analysante (Daubechies 4) et les coefficients de la CWT.

*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.

Nous avons approfondi nos connaissances sur la théorie de la LDA comme une des méthodes d'apprentissage supervisé et son utilisation pour la classification afin de l'utiliser pour classer la sévérité de pollution des isolateurs de haute tension.

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 la standard déviation, 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 pour l'entraînement du modèle LDA que nous avons construit. Les résultats de cette classification montrent que la LDA est une méthode de classification qui donne des résultats moyens pour les deux types de pollution.

En deuxième lieu, Le diagnostic de l'état de surface des isolateurs par la technique des ondelettes nous a également permis de caractériser la sévérité de la pollution grâce au signal de la tension appliquée (pollution uniforme) via la CWT par l'intermédiaire du scalogramme qui est une projection des coefficients calculés pour des échelles données sur un plan orthogonal et pendant l'activité des décharges parallèles. En effet, plus la conductivité volumique de la solution polluante et la concentration en kaolin augmentent, plus les distorsions de la tension tendent vers les hautes fréquences. Mais l'influence de la conductivité volumique est plus remarquable que celle de la concentration de kaolin.

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] M.T.Gencoglu, M.Cebeci, « The pollution flashover on high voltage insulators », Electric Power Systems Research 78, 2008, pp. 1914–1921.
- [4] 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.
- [5] A. Cimador et S. Vilet, « La Pollution des Isolateurs », EDF-Epure, Rapport N° 27, France, Juillet 1990.
- [6] T. Suda, « Frequency Characteristics of Leakage Current Waveforms of Artificially Polluted Suspension insulators », IEEE Trans. Dielectr. Electr. Insul, Vol. 8, pp.705 -709, 2005.
- [7] Stéphane Verdun, Véronique Cariou et El Mostafa Qannari, « Discrimination et classification supervisée en référence à des prototypes », Rapport de recherche, Ecole Nationale Vétérinaire Agroalimentaire et de l'Alimentation Nantes Atlantique –ONIRIS, 2010.
- [8] Julien Jacques et Christophe Biernacki, « Analyse discriminante sur données binaires lorsque les populations d'apprentissage et de test sont différentes », Laboratoire de Statistiques et Analyse des Données, Université Pierre Mendès, France.
- [9] Charles Bouveyron, « Modélisation et classification des données de grande dimension application à l'analyse d'image », Thèse de Doctorat, Université JOSEPH FOURIER – Grenoble1, Septembre 2006.

- [10] Charles Bouveyron, Stéphane Girard et Cordelia Schmid, « Analyse Discriminante de Haute Dimension », Rapport scientifique, Institut nationale de recherche en informatique et en automatique, France, Janvier 2005.
- [11] Michel Vacher et Dan Istrate, « Notes de lecture sur la Transformée en Ondelettes : Communication Langagière et Interaction Personne-Système », Grenoble, France, Septembre 2004.
- [12] S.M. Debbal et F. Bereksi-Reguig, « Analyse spectro-temporelle des bruits cardiaques par les transformées discrete et continue d'ondelettes », Sciences & Technologie B – N°23, Juin 2005, pp. 5-15.
- [13] Landrin Patrick Flandrin, « Temps-fréquence », Edition Hes, Collection traitement du signal, 1998.
- [14] Lee-Pierre Belley, « Séparation de sources aveugle utilisant la transformée en ondelettes continues », Thèse de Doctorat, Montréal, Janvier 2008.
- [15] 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.
- [16] P. Savicky et J. Fürnkranz, « Combining pairwise classifiers with stacking », International symposium on intelligent data analysis, Berlin, Germany, pp. 219-229, 2003.
- [17] 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 », Projet de fin d'études, Ecole nationale polytechnique, Alger, Juin 2011.
- [18] 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.

- [19] C.Chang and C.J. Lin, « LIBSVM: A library for support vector machines », <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.
- [20] M. Moreira, E. Mayoraz, « Improved pairwise coupling classification with correcting classifiers », European conference on machine learning, Germany, pp. 160-171, 21-24, April 1998.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] M.A.Douar et M.C.Bouzidi, « Caractérisation de l'état de surface d'un isolateur pollué à l'aide des signaux du courant de fuite et de la tension appliquée », Projet de fin d'étude, Département de Génie Electrique, Ecole Nationale Polytechnique d'Alger, 2009.
- [26] M.G.E. Schneiders, « Wavelets in Control Engineering », Master's thesis, Eindhoven University of Technology, August 2001.
- [27] A.Bouzida, « Diagnostic de défauts de la machine asynchrone à cage par la technique d'ondelette », Mémoire de Magister, Ecole nationale polytechnique d'Alger 2008.