



Ecole Nationale Polytechnique
Département de Génie Electrique
Laboratoire de Recherche en Electrotechnique



Mémoire de Magister en Electrotechnique

Option : Réseaux Electriques et Haute Tension

Présenté par :

KADRI Farid

Ingénieur d'état en Electrotechnique de l'ENP

Intitulé

PRÉDICTION PAR RÉSEAUX DE NEURONES ARTIFICIELS DE LA VARIATION NON LINÉAIRE DES PROPRIÉTÉS DU PRC VIEILLI THERMIQUEMENT

Soutenu publiquement le **28 /06/2012** devant le jury composé de :

Président :	Boubakeur Ahmed	Professeur	ENP
Rapporteur :	Oudjaout Mohamed	Maitre Assistant	ENP
Examineurs :	Tahmi Redouane	Maitre de Conférences	ENP
	Nezli Lazhari	Maitre de Conférences	ENP

ENP 2012

Remerciements

Les travaux de recherche développés dans ce mémoire ont été réalisés au sein du Laboratoire de Recherche en Electrotechnique (LRE) –Haute Tension- de l'Ecole Nationale Polytechnique d'Alger (ENP).

Je remercie très sincèrement mon encadreur Monsieur OUADJAOUT Mohamed, Enseignant-chercheur en Mathématiques à l'Ecole Nationale Polytechnique, qui m'a offert la possibilité de réaliser ce travail sous sa direction. Je le remercie également pour ses discussions profitables, ses conseils judicieux et ses suggestions. Je lui suis reconnaissant de la confiance qu'il m'a témoignée tout au long de ce travail.

Je tiens à exprimer également mes vifs remerciements à Monsieur BOUBAKEUR Ahmed, Professeur à l'Ecole Nationale Polytechnique, pour son suivi et sa disponibilité. Je le remercie aussi pour ses discussions profitables, ses conseils judicieux et ses suggestions.

Merci aussi à tous mes collègues et amis de l'Ecole Nationale Polytechnique.

Dédicace

À la mémoire de ma mère

SOMMAIRE

INTRODUCTION GENERALE	1
-----------------------------	---

CHAPITRE I : INTRODUCTION AUX RESEAUX DE NEURONES ARTIFICIELS

I.1. OBJECTIFS.....	3
I.2. HISTORIQUE.....	3
I.3. APPLICATIONS.....	5
I.4. LES RESEAUX DE NEURONES ARTIFICIELS.....	7
I.4.1. LE NEURONE BIOLOGIQUE.....	7
I.4.2. STRUCTURE.....	7
I.4.3. LE NEURONE FORMEL (MODELES MATHEMATIQUES).....	9
I.4.4. STRUCTURE.....	9
I.4.5. COMPORTEMENT	9
I.4.6. CARACTERISTIQUES DU NEURONE FORMEL	10
I.4.6.1. FACTEURS DE PONDERATION	10
I.4.6.2. BIAIS.....	11
I.4.6.3. FONCTION D'ACTIVATION	11
I.4.7. CARACTERISTIQUES DES RNA	12
I.4.7.1. COUCHE D'ENTREE.....	12
I.4.7.2. COUCHE CACHEE	12
I.4.7.3. COUCHE DE SORTIE.....	12
I.4.8. CLASSIFICATION DES RESEAUX DE NEURONES.....	13
I.4.9. SELON L'ARCHITECTURE :.....	13
I.4.10. SELON L'APPRENTISSAGE	14
I.5. GRANDS TYPES DE RESEAUX.....	15
I.5.1. PERCEPTRONS MULTI-COUCHES.....	15
I.5.2. RESEAUX DE COHONEN	16
I.5.2.1. VECTOR QUANTIZATION.....	16
I.5.2.2. SOM (SELF ORGANIZING MAP).....	16
I.5.2.3. LVQ (LEARNING VECTOR QUANTIZATION).....	16
I.5.3. RESEAUX DE HOPFIELD.....	17
I.5.4. RESEAUX A RESONANCE ADAPTATIVE (ART).....	17

I.5.5.	RESEAU A FONCTION RADIALE (RBF).....	17
I.6.	L'ALGORITHME DE RETRO- PROPAGATION.....	17
I.6.1.	INITIALISATION DES POIDS.....	19
I.6.2.	ETAPE D'APPRENTISSAGE DE L'ALGORITHME DE RETRO PROPAGATION.....	19
I.6.3.	NORMALISATION DES DONNEES.....	20
I.6.4.	FONCTION DE PERFORMANCE.....	20
I.6.5.	OVERFITTING.....	21
I.7.	CONCLUSION.....	21

CHAPITRE II : INTRODUCTION AUX ALGORITHMES GENETIQUES

II.1.	INTRODUCTION.....	22
II.2.	HISTORIQUE.....	22
II.3.	LES CINQ ELEMENTS PRINCIPAUX DES ALGORITHMES GENETIQUES.....	25
II.4.	FONCTIONNEMENT DES ALGORITHMES GENETIQUES.....	26
II.4.1.	LE CODAGE.....	28
II.4.2.	L'OPERATEUR DE SELECTION.....	30
II.4.3.	L'OPERATEUR DE CROISEMENT OU CROSSOVER.....	33
II.4.4.	L'OPERATEUR DE MUTATION.....	35
II.4.5.	L'OPERATEUR DE REMPLACEMENT.....	37
II.5.	CONCLUSION.....	39

CHAPITRE III : VIEILLISSEMENT THERMIQUE DU PRC

III.1.	INTRODUCTION.....	40
III.2.	PRESENTATION DU MATERIAU.....	40
III.3.	TRAVAUX EFFECTUÉS.....	43
III.4.	POLARISATION DANS LES POLYMERES.....	43
III.5.	LES PRINCIPALES CARACTERISTIQUES ETUDIEES.....	46
III.5.1.	VARIATION DE L'INDICE DE PERTES DIELECTRIQUES.....	46
III.5.2.	VARIATION DE LA CONSTANTE DIELECTRIQUE.....	48
III.5.3.	INFLUENCE DU VIEILLISSEMENT THERMIQUE SUR LA TENSION SEUIL DE D.P.....	50
III.5.4.	VARIATION DE LA PERTE DE MASSE EN FONCTION DU TEMPS DE VIEILLISSEMENT.....	51

CHAPITRE IV : APPROCHE SUIVIE, RESULTATS ET DISCUSSIONS

IV.1.	INTRODUCTION.....	53
IV.2.	LA STRATEGIE SUIVIE DANS CE TRAVAIL.....	54
IV.2.1.	LE NOMBRE DE COUCHES CACHEES :.....	55
IV.2.2.	LE NOMBRE DE NEURONES DES COUCHES CACHEES :.....	56

<i>IV.3. L'UTILISATION DES ALGORITHMES GENETIQUES DANS CE TRAVAIL</i>	<i>57</i>
<i>IV.4. L'ORGANIGRAMME DE L'APPROCHE SUIVIE.....</i>	<i>58</i>
<i>IV.5. RESULTATS ET DISCUSSION</i>	<i>59</i>
<i>IV.5.1. LE DOMAINE DE CONFIANCE DE LA PREDICTION.....</i>	<i>59</i>
<i>IV.5.2. LA TAILLE DE LA BASE DE DONNEES D'APPRENTISSAGE</i>	<i>59</i>
<i>IV.5.3. PREDICTION DE LA VARIATION DE L'INDICE DE PERTES DIELECTRIQUES</i>	<i>59</i>
<i>IV.5.4. PREDICTION DE LA VARIATION DE LA CONSTANTE DIELECTRIQUE</i>	<i>65</i>
<i>IV.5.5. PREDICTION DE LA VARIATION DE LA TENSION SEUIL DE D.P.....</i>	<i>67</i>
<i>IV.5.6. PREDICTION DE LA VARIATION DE LA PERTE DE MASSE.....</i>	<i>69</i>
<i>CONCLUSION GENERALE</i>	<i>72</i>

Introduction Générale

Au cours de la dernière décennie, les algorithmes d'apprentissage statistique ont suscité beaucoup d'intérêt, que se soit sur un plan académique, ou bien dans la gestion d'entreprises. Ils ont été implémentés avec succès pour l'accomplissement de tâches prédictives reliées à des processus statistiques observés, pour lesquels on peut identifier plusieurs variables explicatives [1].

Ce travail se concentre sur une classe particulière de ces algorithmes :

Les réseaux de neurones artificiels et les algorithmes génétiques.

Les réseaux de neurones tirent leur puissance de modélisation dans leur capacité à déceler des dépendances fortes qui implique plusieurs variables liées.

Cependant, et comme pour toutes les techniques qui font partie de l'ensemble des techniques de l'intelligence artificielle, les réseaux de neurones artificiels présentent des limites et des insuffisances. Ceci nous amène à combiner plusieurs techniques intelligentes afin de surmonter ces limites.

Dans le présent travail, nous étudions et traiterons le problème du choix des paramètres structurels relatifs aux réseaux de neurones artificiels. Il s'agit du nombre de couches cachées, et du nombre de neurones dans chaque couche cachée.

Pour cela, nous avons fait appel à une autre technique intelligente connue dans le domaine de la recherche par sa capacité d'optimisation : il s'agit des algorithmes génétiques, que nous avons appliqués à la prédiction des variations non linéaires des caractéristiques du PRC (Polyéthylène Réticulé Chimiquement) vieilli thermiquement. C'est l'objet de ce travail.

Ce mémoire est composé de quatre chapitres :

Le premier chapitre est une introduction générale aux réseaux de neurones artificiels, en décrivant le principe de fonctionnement ainsi que les types les plus étudiés dans la recherche et les plus utilisés en pratique. L'algorithme d'apprentissage « Back Propagation » est également décrit.

Dans le deuxième chapitre, nous exposons les grandes lignes des algorithmes génétiques, les règles de fonctionnement, et cela pour faciliter la compréhension de la technique hybride utilisée dans ce travail.

Nous tenons à noter que les informations présentées dans le premier et le deuxième chapitres ne sont qu'une synthèse bibliographique à partir d'un certains nombres de livres et de thèses qui ont traités ces techniques [1-18].

Le troisième chapitre est consacré à la description de quelques travaux qui ont étudié le vieillissement thermiques du PRC utilisé dans les câbles à moyenne tension [19] [20]. Il s'agit des travaux effectués par M. Mecheri (thèse de doctorat ENP -2008-), et M. Boukezzi (thèse de doctorat ENP -2007-). Les résultats expérimentaux présentés dans ce chapitre vont être comparés à nos résultats obtenus par l'application de la technique hybride RNA-AG.)

Dans le quatrième et dernier chapitre, nous présentons les résultats obtenus par l'application de la technique hybride RNA-AG à la prédiction des variations non linéaires des caractéristiques du PRC (Polyéthylène Réticulé Chimiquement) vieilli thermiquement, ces résultats ont été comparés aux résultats pratiques obtenus par Mecheri et Boukezzi, On termine ce chapitre par quelques perspectives envisagées pour améliorer et compléter le travail effectué dans le cadre de ce mémoire de magister.

Chapitre I.

Introduction aux Réseaux de Neurones Artificiels

I.1. OBJECTIFS

Le cerveau humain contient environ 100 milliards de neurones. Ces neurones nous permettent, entre autre, de lire ce texte tout en maintenant une respiration régulière permettant d'oxygéner le sang, en actionnant notre cœur qui assure une circulation efficace de ce sang pour nourrir les cellules, etc. [1].

Chacun de ces neurones est par ailleurs fort complexe. Essentiellement, il s'agit de tissu vivant et de chimie. Les spécialistes des neurones biologiques (ceux qui œuvrent en neurophysiologie) commencent à peine à comprendre quelques uns de leurs mécanismes internes. On croit en général que leurs différentes fonctions neuronales, y compris celle de la mémoire, sont stockées au niveau des connexions (synapses) entre les neurones. C'est ce type de théorie qui a inspiré la plupart des architectures de réseaux de neurones artificiels que nous aborderons dans ce chapitre. L'apprentissage consiste alors soit à établir de nouvelles connexions, soit à en modifier celles existantes. Ceci nous amène à poser une question fondamentale : en ce basant sur nos connaissances actuelles, peut-on construire des modèles approximatifs de neurones et les entrainer pour, éventuellement, réaliser des tâches utiles (comme la prédiction par exemple)? Généralement, on peut répondre par «oui», même si les réseaux que nous allons développer ne possèdent qu'une infime fraction de la puissance du cerveau humain, et c'est l'objectif de ce travail pour montrer comment on peut y arriver. [1]

I.2. HISTORIQUE

Deux ingrédients sont à la base de tout avancement des connaissances. Premièrement, il importe de posséder un nouveau concept, ou un nouveau point de vue à propos d'un sujet, qui vient jeter la lumière là où il n'y avait qu'obscurité. Par exemple, considérons le cœur humain. A différentes époques, on le considérait comme le centre de l'âme ou encore comme une source de chaleur. Au 17e siècle, les médecins ont commencé à le considérer comme une pompe et ont donc conçu des expériences pour tenter de comprendre son fonctionnement, ce qui a éventuellement permis une compréhension

du système sanguin. Sans le concept de pompe, une compréhension du cœur et du système sanguin en général était simplement hors d'atteinte [2].

Deuxièmement, il importe aussi de posséder des outils technologiques permettant de construire des systèmes concrets. Par exemple, on connaissait les théories physiques permettant d'envisager la conception d'une bombe atomique bien avant d'être capable de réaliser une telle bombe. On savait aussi mathématiquement reconstruire des images de radiographie en coupe (tomographie) bien avant de posséder les ordinateurs et les algorithmes capables d'effectuer efficacement les calculs requis dans un temps raisonnable [2].

Brièvement, les premières recherches remontent à la fin du 19^e et au début du 20^e siècle. Elles ont consisté en de travaux multidisciplinaires en physique, en psychologie et en neurophysiologie par des scientifiques tels Hermann Von Helmholtz, Ernst Mach et Ivan Pavlov [4] [7]. À cette époque, il s'agissait de théories plutôt générales sans modèles mathématiques précis d'un neurone. On s'entend pour dire que la naissance du domaine des réseaux de neurones artificiels remonte aux années 1940 avec les travaux de Warren McCulloch et Walter Pitts [7] qui ont montré qu'avec de tels réseaux, on pouvait, en principe, calculer n'importe quelle fonction arithmétique ou logique. Vers la fin des années 1940, Donald Hebb a ensuite proposé une théorie fondamentale pour l'apprentissage [7].

La première application concrète des réseaux de neurones artificiels est survenue vers la fin des années 1950 avec l'invention du réseau dit «perceptron» par Frank Rosenblatt [8].

Rosenblatt et ses collègues ont construit un réseau et démontré ses habilités à reconnaître des formes. Malheureusement, il a été démontré par la suite que ce perceptron simple ne pouvait résoudre qu'une classe limitée de problèmes.

Dans la même période, Bernard Widrow et Ted Hoff ont proposé un nouvel algorithme d'apprentissage pour entraîner un réseau adaptatif de neurones linéaires, dont la structure et les capacités sont similaires au perceptron [8].

Vers la fin des années 1960, un livre publié par Marvin Minsky et Seymour Papert [7] est venu jeter beaucoup d'ombre sur le domaine des réseaux de neurones. Entre autres choses, ces deux auteurs ont démontré les limitations des réseaux développés par Rosenblatt et Widrow-Hoff. Beaucoup de personnes ont été influencés par cette démonstration qu'ils ont généralement mal interprétée. Ils ont conclu à tort que le

domaine des réseaux de neurones était un cul de sac et qu'il fallait cesser de s'y intéresser (et de financer la recherche dans ce domaine), d'autant plus qu'on ne disposait pas à l'époque d'ordinateurs suffisamment puissants pour effectuer des calculs complexes [8].

Heureusement, certains chercheurs ont persévéré, en développant de nouvelles architectures et de nouveaux algorithmes plus puissants. En 1972, Teuvo Kohonen et James Anderson [7] [8] ont développé indépendamment et simultanément de nouveaux réseaux pouvant servir de mémoires associatives. Egalement, Stephen Grossberg a investigué ce qu'on appelle les réseaux auto-organisés [8].

Dans les années 1980, une pierre d'achoppement a été levée par l'invention de l'algorithme de rétro-propagation des erreurs. Cet algorithme est la réponse aux critiques de Minsky et Papert formulées à la fin des années 1960 [7]. C'est ce nouveau développement, généralement attribué à David Rumelhart et James McClelland, mais aussi découvert plus ou moins en même temps par Paul Werbos et par Yann LeCun [1], qui a littéralement ressuscité le domaine des réseaux de neurones. Depuis, c'est un domaine où bouillonnent constamment de nouvelles théories, de nouvelles structures et de nouveaux algorithmes. Nous tenterons dans ce chapitre d'en survoler les principaux [1].

1.3. APPLICATIONS

Les réseaux de neurones servent aujourd'hui à toutes sortes d'applications dans divers domaines.

Par exemple, on a développé un autopilote pour avion, ou encore un système de guidage pour automobile. On a conçu des systèmes de lecture automatique de chèques bancaires et d'adresses postales. On produit des systèmes de traitement du signal pour différentes applications militaires, un système pour la synthèse de la parole.

Des réseaux sont aussi utilisés pour bâtir des systèmes de vision par ordinateur, pour effectuer des prévisions sur les marchés monétaires, pour évaluer le risque financier ou en assurance, pour différents processus manufacturiers, pour le diagnostic médical, pour l'exploration pétrolière ou gazière, en robotique, en télécommunication, et on peut citer beaucoup d'autres applications. Bref, les réseaux de neurones ont aujourd'hui un impact considérable et, il y a fort à parier, que leur importance ira grandissant dans le futur [1].

Industry	Applications
Aerospace	High-performance aircraft autopilot, flight path simulation, aircraft control systems, autopilot enhancements, aircraft component simulation, and aircraft component fault detection
Automotive	Automobile automatic guidance system, and warranty activity analysis
Banking	Check and other document reading and credit application evaluation
Defense	Weapon steering, target tracking, object discrimination, facial recognition, new kinds of sensors, sonar, radar and image signal processing including data compression, feature extraction and noise suppression, and signal/image identification
Electronics	Code sequence prediction, integrated circuit chip layout, process control, chip failure analysis, machine vision, voice synthesis, and nonlinear modeling
Entertainment	Animation, special effects, and market forecasting
Financial	Real estate appraisal, loan advising, mortgage screening, corporate bond rating, credit-line use analysis, credit card activity tracking, portfolio trading program, corporate financial analysis, and currency price prediction
Industrial	Prediction of industrial processes, such as the output gases of furnaces, replacing complex and costly equipment used for this purpose in the past
Insurance	Policy application evaluation and product optimization
Manufacturing	Manufacturing process control, product design and analysis, process and machine diagnosis, real-time particle identification, visual quality inspection systems, beer testing, welding quality analysis, paper quality prediction, computer-chip quality analysis, analysis of grinding operations, chemical product design analysis, machine maintenance analysis, project bidding, planning and management, and dynamic modeling of chemical process system
Medical	Breast cancer cell analysis, EEG and ECG analysis, prosthesis design, optimization of transplant times, hospital expense reduction, hospital quality improvement, and emergency-room test advisement
Oil and gas	Exploration
Robotics	Trajectory control, forklift robot, manipulator controllers, and vision systems
Speech	Speech recognition, speech compression, vowel classification, and text-to-speech synthesis

Securities	Market analysis, automatic bond rating, and stock trading advisory systems
Telecommunications	Image and data compression, automated information services, real-time translation of spoken language, and customer payment processing systems
Transportation	Truck brake diagnosis systems, vehicle scheduling, and routing systems

Tableau I-1 : Les domaines d'application des RNA [6]

I.4. LES RESEAUX DE NEURONES ARTIFICIELS

I.4.1. LE NEURONE BIOLOGIQUE

Les cellules nerveuses, appelées neurones, sont les éléments de base du système nerveux central. Celui-ci en possède environ cent milliards. Les neurones possèdent de nombreux points communs dans leur organisation générale et leur système biologique, avec les autres cellules. Ils présentent cependant des caractéristiques qui leur sont propres et se trouvent dans cinq fonctions spécialisées qu'ils assurent [3]:

- ✓ Recevoir des signaux en provenance de neurones voisins,
- ✓ Intégrer ces signaux,
- ✓ Engendrer un flux nerveux,
- ✓ Conduire le flux nerveux et le transmettre à un autre neurone capable de le recevoir.

I.4.2. STRUCTURE

Un neurone est constitué de trois parties :

- Le corps cellulaire,
- Les dendrites,
- L'axone.

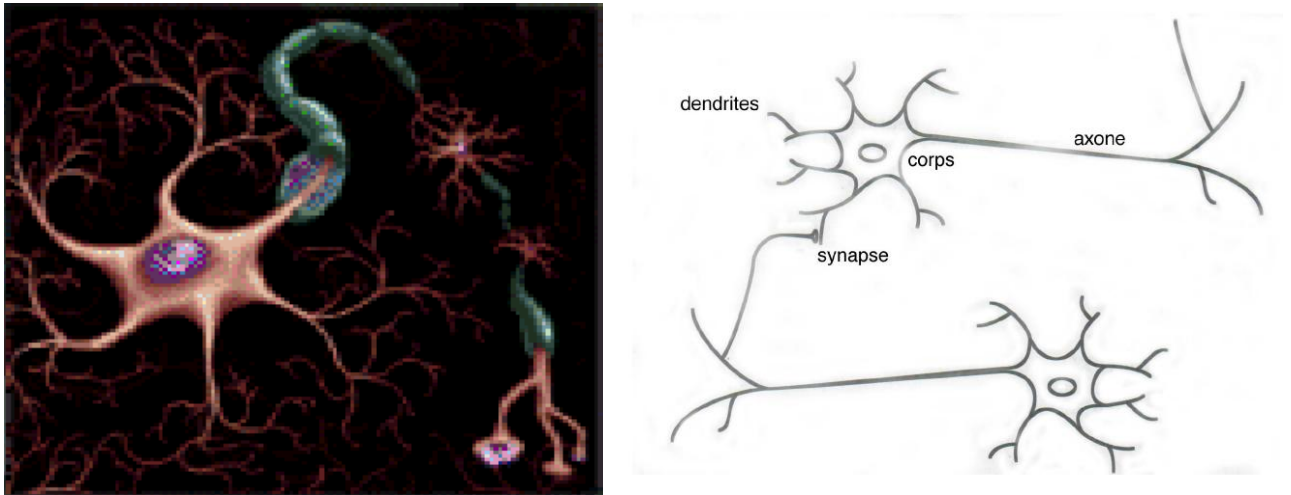


Figure I-1 : Le neurone biologique [1] [3]

A. LE CORPS CELLULAIRE

Il contient le noyau du neurone et effectue les transformations biochimiques nécessaires à la synthèse des enzymes et des autres molécules qui assurent la vie du neurone. Sa forme est pyramidale ou sphérique dans la plupart des cas. Elle dépend souvent de sa position dans le cerveau.

B. LES DENDRITES

Chaque neurone possède une « chevelure » de dendrites. Celles-ci sont de fines extensions tubulaires. Elles se ramifient, ce qui les amène à former une espèce d'arborescence autour du corps cellulaire. Elles sont les récepteurs principaux du neurone pour capter les signaux qui lui parviennent.

C. L'AXONE

L'axone, qui est la fibre nerveuse, sert de moyen de transport pour les signaux émis par le neurone. Il se distingue des dendrites par sa forme et par les propriétés de sa membrane externe. En effet, il est généralement plus long que les dendrites, et se ramifie à l'extrémité, là où il communique avec d'autres neurones, alors que les ramifications des dendrites se produisent plutôt près du corps cellulaire. Pour former le système nerveux, les neurones sont connectés les uns aux autres suivant des répartitions spatiales complexes. Les connexions entre deux neurones se font en des endroits appelés synapses où ils sont séparés par des espaces synaptiques. [3]

1.4.3. LE NEURONE FORMEL (MODELES MATHEMATIQUES)

Les réseaux de neurones biologiques réalisent facilement un certain nombre d'applications telles que la reconnaissance de formes, le traitement du signal, l'apprentissage par l'exemple, la mémorisation, la généralisation. Ces applications sont pourtant, malgré tous les efforts déployés en algorithmique et en intelligence artificielle, à la limite des possibilités actuelles. C'est à partir de l'hypothèse que le comportement intelligent émerge de la structure et du comportement des éléments de base du cerveau que les réseaux de neurones artificiels se sont développés. Les réseaux de neurones artificiels sont des modèles. À ce titre, ils peuvent être décrits par leurs composants, leurs variables descriptives et les interactions des composants. [2]

1.4.4. STRUCTURE

La figure I-2 montre la structure d'un neurone artificiel. Chaque neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones amont. A chacune de ces entrées est associé un poids w , abréviation de weight (poids), représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones aval. A chaque connexion est associé un poids.

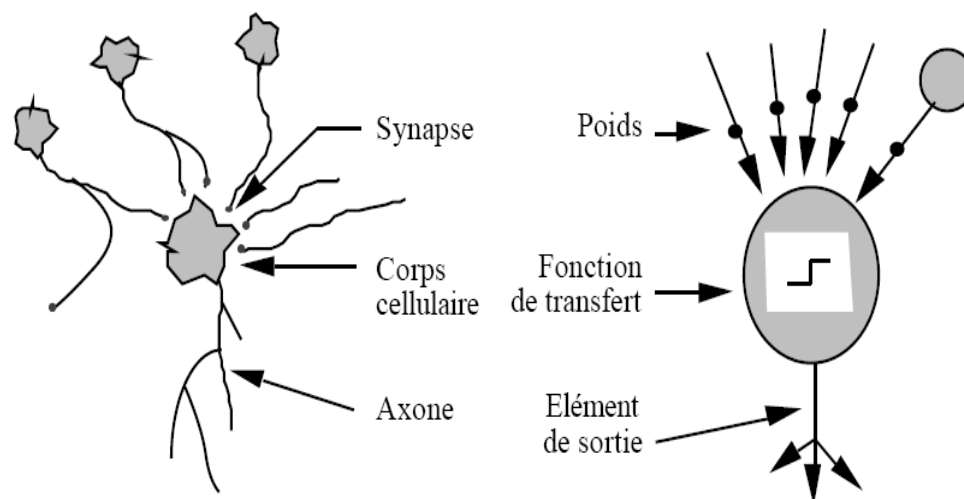


Figure I-2 : Mise en correspondance neurone biologique / neurone artificiel [2]

1.4.5. COMPORTEMENT

Un neurone est essentiellement constitué d'un intégrateur qui effectue la somme pondérée de ses entrées. Le résultat n de cette somme est ensuite transformé par une

fonction de transfert f qui produit la sortie \mathbf{a} du neurone. les R entrées du neurone correspondent au vecteur $\mathbf{p} = [p_1 \ p_2 \ \dots \ p_R]^T$, alors que $\mathbf{w} = [w_{1,1} \ w_{1,2} \ \dots \ w_{1,R}]^T$ représente le vecteur des poids du neurone. La sortie n de l'intégrateur est donnée par l'équation suivante :

$$\begin{aligned} n &= \sum_{j=1}^R w_{1,j} p_j - b_j \\ &= w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R - b \end{aligned}$$

Que l'on peut aussi écrire sous forme matricielle :

$$\mathbf{n} = \mathbf{w}^T \mathbf{p} - \mathbf{b}$$

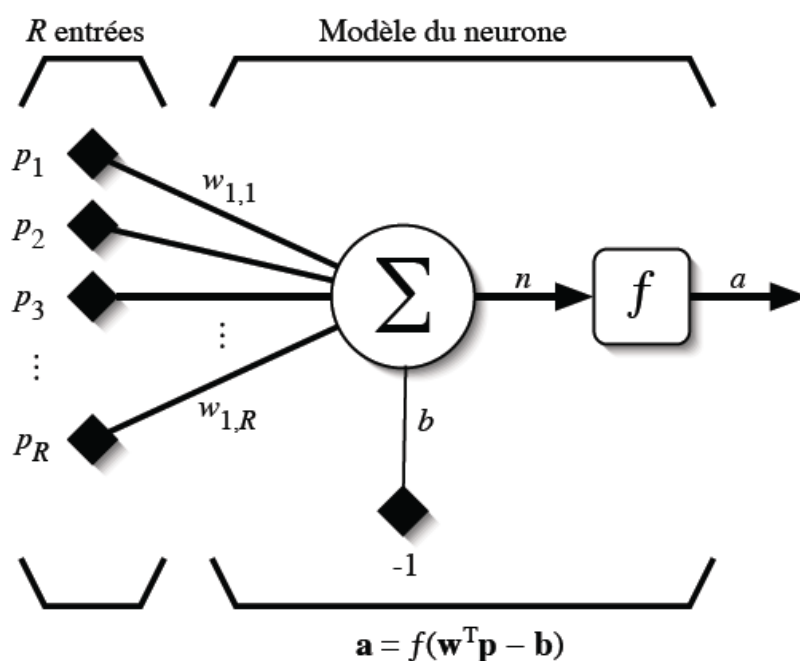


Figure I-3 : Modèle d'un neurone artificiel [5] [7]

1.4.6. CARACTERISTIQUES DU NEURONE FORMEL

1.4.6.1. FACTEURS DE PONDERATION

Les valeurs $w_{1,1}, w_{1,2}, w_{1,3} \dots w_{1,R}$ sont des facteurs de pondération liés à chaque nœud pour déterminer la force du vecteur de ligne d'entrée $R = [p_1 \ p_2 \ p_3 \ \dots, \ p_R]^T$.

Chaque entrée est multipliée par le poids associé de la connexion de neurone : X^tW .

1.4.6.2. BIAIS

Le seuil interne b du nœud est la grandeur compensée qui affecte le lancement de la sortie n du nœud comme suit, [4] :

$$n = \sum_{j=1}^R w_{1,j} p_j - b_j$$

1.4.6.3. FONCTION D'ACTIVATION

Dans cette sous-section, cinq fonctions d'activation les plus communes sont présentées. Une fonction d'activation exécute un fonctionnement mathématique sur la sortie du signal. Toutes les fonctions d'activation décrites ici sont également supportées par MATLAB Tableau 3. [4]










Nom de la fonction	Relation d'entrée/sortie	Icône	Nom Matlab
seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlim
seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlims
linéaire	$a = n$		purelin
linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$		satlin
linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$		satlins
linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$		poslin
sigmoïde	$a = \frac{1}{1+\exp^{-n}}$		logsig
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
compétitive	$a = 1$ si n maximum $a = 0$ autrement		compet

Tableau I-2 : Fonctions d'activation des RNA [4]

1.4.7. CARACTERISTIQUES DES RNA

1.4.7.1. COUCHE D'ENTREE

Le nombre de neurones dans cette couche correspond au nombre d'entrées au réseau de neurone. Cette couche se compose de nœuds passifs, c.-à-d, qui ne participent pas à la modification réelle de signal, mais transmet seulement le signal à la couche suivante.

1.4.7.2. COUCHE CACHEE

Cette couche a un nombre arbitraire des couches avec un nombre arbitraire des neurones. Les nœuds dans cette couche participent à la modification du signal, par conséquent, ils sont en activité.

1.4.7.3. COUCHE DE SORTIE

Le nombre de neurones dans la couche de sortie correspond au nombre de sorties du réseau de neurone. Les nœuds dans cette couche sont actifs. [1]

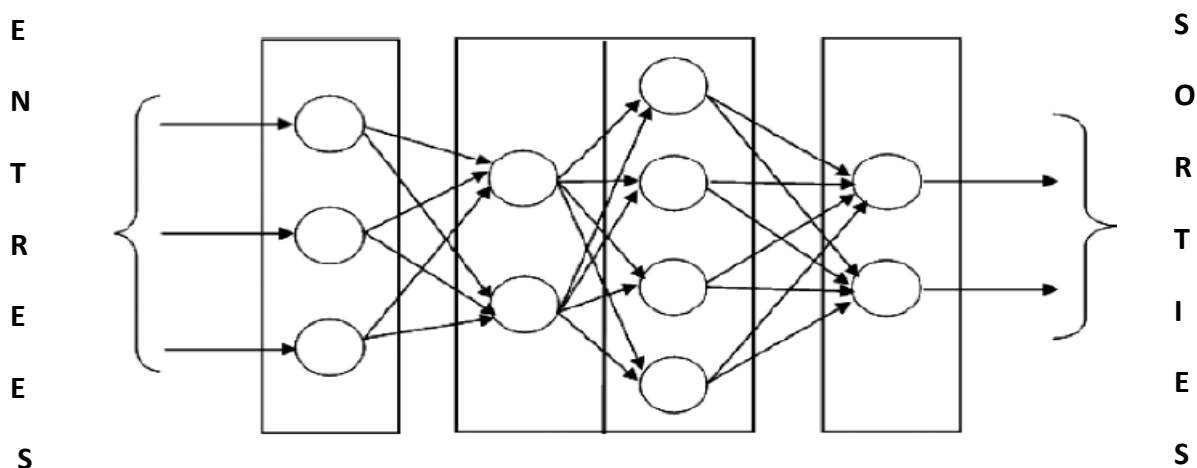


Figure I-4 : Architecture générale d'un réseau de neurone

1.4.8. CLASSIFICATION DES RESEAUX DE NEURONES

Toutes les caractéristiques des réseaux de neurones sont ensuite exploitées à travers la propriété principale des réseaux de neurones qu'est l'apprentissage.

Cet apprentissage permet de configurer les poids synaptiques ainsi que les fonctions d'activation afin d'adopter un comportement désiré.

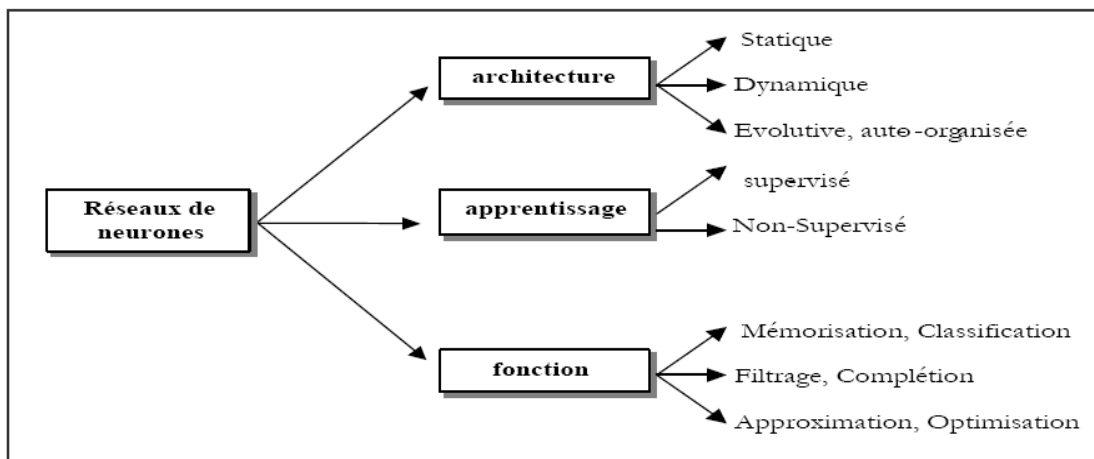


Figure I-5 : Les différentes possibilités de classification des réseaux des neurones [3]

1.4.9. SELON L'ARCHITECTURE :

1.4.9.1. RESEAUX STATIQUES

Ces réseaux sont constitués de plusieurs couches. Chaque neurone d'une couche reçoit ses entrées à partir des neurones de la couche précédente. Dans de tels réseaux, il n'existe pas de retour d'informations. Ils peuvent être utilisés dans les problèmes de classification et d'approximation de fonctions [6].

1.4.9.2. RESEAUX DYNAMIQUES

Contrairement aux précédents, ces réseaux contiennent des retours d'informations. Cet aspect leur confère la possibilité de commander et d'identifier les systèmes dynamiques [6].

1.4.9.3. RESEAUX AUTO-ORGANISE

Les réseaux de neurones auto-organisés sont des réseaux qui changent leurs structures internes pendant l'apprentissage. Ainsi, les neurones se regroupent topologiquement

suivant la représentation des exemples. Ces réseaux sont des dérivés des modèles de Kohonen [3].

1.4.10. SELON L'APPRENTISSAGE

Parmi les propriétés désirables pour un réseau de neurones, la plus fondamentale est sûrement la capacité d'apprendre de son environnement, d'améliorer sa performance à travers un processus d'apprentissage. Mais qu'est-ce donc que l'apprentissage ? Malheureusement, il n'existe pas de définition générale, universellement acceptée, car ce concept touche à trop de notions distinctes qui dépendent du point de vue que l'on adopte [1].

Dans le contexte des réseaux de neurones artificiels, nous adopterons un point de vue pragmatique en proposant la définition suivante :

« L'apprentissage est un processus dynamique et itératif permettant de modifier les paramètres d'un réseau en réaction avec les stimuli qu'il reçoit de son environnement. Le type d'apprentissage est déterminé par la manière dont les changements de paramètres surviennent. » [1]

Cette définition implique qu'un réseau se doit d'être stimulé par un environnement, qu'il subisse des changements en réaction avec cette stimulation, et que ceux-ci provoquent dans le futur une réponse nouvelle vis-à-vis de l'environnement. Ainsi, le réseau peut s'améliorer avec le temps.

Nous allons passer en revue les deux types de principes – les plus connus – pouvant guider l'apprentissage d'un réseau de neurone.

1.4.10.1. APPRENTISSAGE SUPERVISE

L'apprentissage supervisé permet de déterminer les poids synaptiques à partir d'exemples étiquetés auxquels un expert a associé des réponses du réseau. Les paramètres du réseau sont donc modifiés de manière à minimiser l'erreur entre la

sortie cible (fournie par l'expert) et la sortie réelle du réseau [5] [3] [8]. Il y'a plusieurs exemples d'algorithmes d'apprentissage dirigés :

- L'algorithme least mean square (LMS) et sa généralisation, connus sous le nom de « retro propagation ».
- la loi Delta.
- Gradient descent.

1.4.10.2. APPRENTISSAGE NON-SUPERVISE

Les données fournies en entrée ne contiennent pas d'information sur la sortie désirée. L'apprentissage est réalisé à l'aide des règles qui modifient les paramètres du réseau en fonction des exemples fournis en entrée. On peut citer dans cette catégorie la loi de Hebb [5] [8].

1.5. GRANDS TYPES DE RESEAUX

Il existe plusieurs types de réseaux de neurones pour plusieurs applications.

On peut citer :

1.5.1. PERCEPTRONS MULTI-COUCHES

Le premier réseau de neurones que nous allons étudier s'appelle le «perceptron multicouche» (PMC). Ce type de réseau est dans la famille générale des réseaux à «propagation vers l'avant», c'est -à-dire qu'en mode normal d'utilisation, l'information se propage dans un sens unique, des entrées vers les sorties sans aucune rétroaction. Son apprentissage est de type supervisé, par correction des erreurs. Dans ce cas uniquement, le signal d'erreur est «rétropropagé» vers les entrées pour mettre à jour les poids des neurones [5] [8].

Un réseau multicouche n'est rien d'autre qu'un assemblage de couches concaténées les unes aux autres, de la gauche vers la droite, en prenant les sorties d'une couche et en les injectant comme les entrées de la couche suivante.

Le perceptron multicouche est un des réseaux de neurones les plus utilisés pour des problèmes d'approximation, de classification et de prédiction. Il est habituellement constitué de deux ou trois couches de neurones totalement connectés.

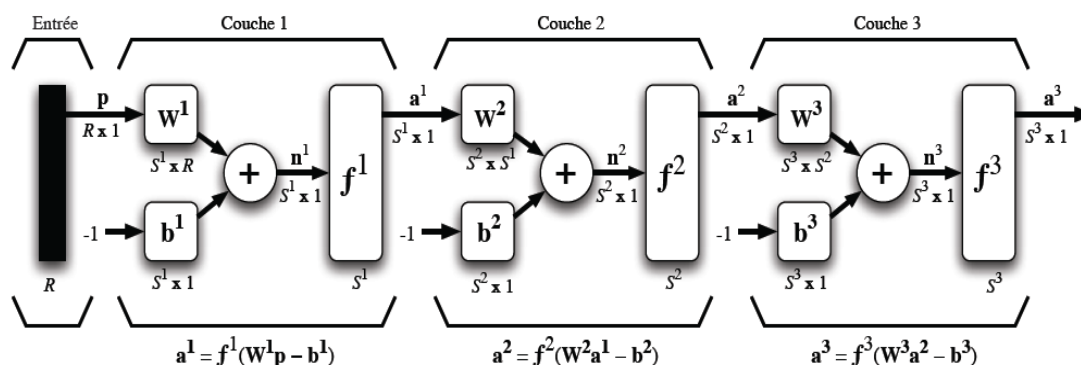


Figure I-6 : Représentation matricielle d'un réseau de trois couches [1]

1.5.2. RESEAUX DE KOHONEN

Les réseaux de Kohonen décrivent en fait trois familles de réseaux de neurones:

1.5.2.1. VQ (VECTOR QUANTIZATION)

Introduite par Grossberg (1976), la quantification vectorielle est une méthode généralement qualifiée d'estimateur non supervisé de densité. Elle permet de retrouver des groupes sur un ensemble de données [3] [5] [8].

1.5.2.2. SOM (SELF ORGANIZING MAP)

Les SOM sont issus des travaux de Fausett (1994) et Kohonen (1995). Ces réseaux sont très utilisés pour l'analyse de données. Ils permettent de cartographier en deux dimensions et de distinguer des groupes dans des ensembles de données [3] [5] [8].

1.5.2.3. LVQ (LEARNING VECTOR QUANTIZATION)

Les réseaux utilisant la méthode LVQ ont été proposés par Kohonen (1988). Des trois types de réseaux présentés ici, la LVQ est la seule méthode qui soit réellement adaptée à la classification de données par "recherche du plus proche voisin" [3] [5] [8].

1.5.3. RESEAUX DE HOPFIELD

Ces réseaux sont des réseaux récurrents, un peu plus complexes que les perceptrons multicouches. Chaque cellule est connectée à toutes les autres et les changements de valeurs de cellules s'enchaînent en cascade jusqu'à un état stable. Ces réseaux sont bien adaptés à la reconnaissance de formes [3].

1.5.4. RESEAUX A RESONANCE ADAPTATIVE (ART)

Ces modèles ont été développés par Carpenter et Grossberg (ART-1, ART-2, ARTMAP). Ils possèdent un réservoir de cellules de sortie qui ne sont utilisées que si nécessaire. Une cellule sera dite *recrutée* (committed) ou *libre* (uncommitted) si elle est utilisée ou non utilisée. Ces réseaux sont bien adaptés à la Classification et la Catégorisation au sein d'une classe [3].

1.5.5. RESEAU A FONCTION RADIALE (RBF)

Le réseau RBF est un réseau de neurones utilise l'apprentissage supervisé. Il s'agit d'une '*spécialisation*' d'un PMC. Le réseau possède trois couches comme le PMC, où la couche cachée contient des neurones RBF qui contiennent chacune une fonction d'activation gaussienne [1] [7] [9].

1.6. L'ALGORITHME DE RETRO-PROPAGATION

- Les réseaux de neurones les plus utilisés sont comme suit [4] :
 - Perceptron multi couche 'Multilayer Perceptrons' (MLP).
 - Réseau à fonction radiale (RBF network).
 - réseau de Hopfield.
 - Adaline.
 - réseau de Kohonen

Le réseau à rétro-propagation concerne les réseaux de neurones multicouches qui utilisent l'algorithme de rétro-propagation pour l'apprentissage. Le plus souvent, des réseaux de neurones Feed-forward multicouche sont utilisant pour l'algorithme de rétro-propagation. Dans ce cas, beaucoup de versions de l'algorithme de rétro-

propagation existent. Tous ces algorithmes contribuent principalement à une formation plus rapide. Nous citons comme exemple :

- Variable Learning Rate
- Resilient Backpropagation
- Conjugate Gradient Algorithm
- Quasi-Newton Algorithm
- Levenberg-Marquardt Algorithm
- Reduced Memory Levenberg-Marquardt Algorithm [4]

L'algorithme de rétro-propagation est l'un des algorithmes les plus connus pour prévoir les poids du réseau de neurones. Il a été développé par « David Rumelhart », Geoff Hinton et R.J. Williams en 1986 [4]. C'est l'un des algorithmes d'apprentissage les plus populaires dû à sa réussite du point de vue simplicité et applicabilité. L'algorithme se compose de deux phases : phase d'apprentissage et phase de rappel (validation & test). Dans la phase d'apprentissage, les poids sont initialisés au hasard puis les sorties sont calculées et comparées aux sorties désirées. Par la suite, l'erreur du réseau est calculée et utilisée pour régler les poids de la couche de sortie. D'une manière similaire, l'erreur du réseau est propagée vers l'arrière et employée pour mettre à jour les poids des couches précédentes. La phase d'apprentissage sera terminée quand la valeur d'erreur est inférieure à la valeur réglée par le superviseur. Un des désavantages de l'algorithme rétro-propagation est que la phase d'apprentissage est très longue, pendant la phase de rappel (validation & test), le réseau avec les poids finaux résultant du procédé d'apprentissage sont utilisés. L'algorithme de rétro propagation est l'extension de la « règle delta » utilisée dans l'apprentissage des réseaux de neurones à une couche (perceptron). [5] [6]

Les valeurs S_j de sortie calculées d'une configuration indiquée P_l d'entrée ne correspondent pas toujours à leurs valeurs prédéterminées (Désirées) D_j . L'erreur E_j est donnée par la différence de S_j et de D_j et doit être réduite au minimum par les modifications de poids. Afin de réaliser ceci, l'erreur de sortie va se propager vers l'arrière aux couches cachées. Soit l'erreur totale en moyenne quadratique E_j définie par :

$$E = \sum E_j = \frac{1}{n} \sum (D_j - S_j)^2$$

Ainsi, nous devons réduire au minimum l'erreur E, qui suit le changement des poids. Nous suivons la règle de delta pour régler le facteur α d'apprentissage. L'algorithme de descente de gradient est employé pour choisir la méthode de modification du poids. Nous utilisons ces deux techniques pour définir la modification du poids, [4] :

$$\Delta W_{kj} = -\alpha \frac{\partial E}{\partial W_{kj}}, 0 < \alpha \leq 1$$

$$W_{kj} = W_{kj} + \Delta W_{kj}$$

1.6.1. INITIALISATION DES POIDS

Si les poids initiaux du réseau de neurones sont proches du zéro, la sortie produira une valeur plutôt basse. C'est pour cela qu'on n'aura jamais la possibilité de représenter des systèmes non linéaire, car le réseau va tendre vers un système linéaire.

Les poids sont pratiquement initialisés suivant une loi aléatoire telle que la fonction de « Nguyen-Widrow » [4].

1.6.2. ETAPE D'APPRENTISSAGE DE L'ALGORITHME DE RETRO PROPAGATION

1. Initialisation aléatoire des poids.
2. Choisir la structure du réseau.
3. Choisir les fonctions d'activations des neurones.
4. Sélectionner le sous ensemble d'apprentissage et l'appliquer au réseau.
5. Calculer les sorties du réseau en se basant sur les poids initiaux et les entrées.
6. Calculer l'erreur entre les sorties calculées et les sorties désirées.
7. Propager l'erreur vers l'arrière et ajuster les poids de tel sorte à minimiser l'erreur de sortie, et cela en allant de la couche de sortie vers la couche d'entrée.
8. Répéter les étapes 4-7 jusqu'à ce que l'erreur soit inférieure au minimum prédéfini.

Après plusieurs de répétitions de ces opérations, l'erreur entre les sorties réelles et les sorties de cible devraient être réduites à une valeur acceptable, et le réseau est dit appris. A ce moment, le réseau peut être utilisé dans les phases de rappel (teste) ou de « généralisation » où les poids ne sont pas changés. [5]

1.6.3. NORMALISATION DES DONNEES

Les réseaux neuronaux exigent que leurs données d'entrées et de sortie soient normalisées pour avoir le même ordre de grandeur. La normalisation est très critique pour certaines applications. [5]. Elle a l'influence sur la qualité de la sortie produite. C'est une opération importante parce que le réseau apprend mieux avec un intervalle régularisé et bondé d'entrée plutôt que des valeurs arbitraires. C'est également important pour améliorer la capacité de généralisation du réseau.

En outre, si des fonctions d'activation sigmoïdes ou tan-sigmoïdes sont utilisées, elles sont particulièrement utilisées avec une plage de données se trouvant dans l'intervalle [0, 1] et [-1, 1] respectivement. [4]

1.6.4. FONCTION DE PERFORMANCE

Ces fonctions sont utilisées pour calculer l'erreur des réseaux de neurones Les fonctions suivantes sont supportées par l'environnement MATLAB

Fonction	formule
Mse	$\frac{1}{n} \sum_1^n (t_i - a_i)^2$
Msereg	$msereg = \gamma mse + (1 - \gamma) msw$
Sse	$\sum_1^n (t_i - a_i)^2$
Mae	$\frac{1}{n} \sum_1^n t_i - a_i $

Tableau I-2 : Fonctions de performance

1.6.5. OVERFITTING

D'un point de vue statistique, l'overfitting signifie équiper un modèle d'un surplus de paramètres. Un réseau de neurones qui apprend a généralement une bonne capacité de généralisation, si l'on gère mal l'apprentissage du réseau, par exemple si l'on arrive à trop minimiser l'erreur, le réseau aura « appris par cœur » les exemples, et si l'entrée n'est pas exactement celle qu'il aura apprise, il la classera mal. On appelle ce type de sur-apprentissage de l'overfitting. Le réseau se borne à ressortir exactement ce qu'on lui a appris sans généraliser [4]. Ceci est expliqué dans la figure I-7.

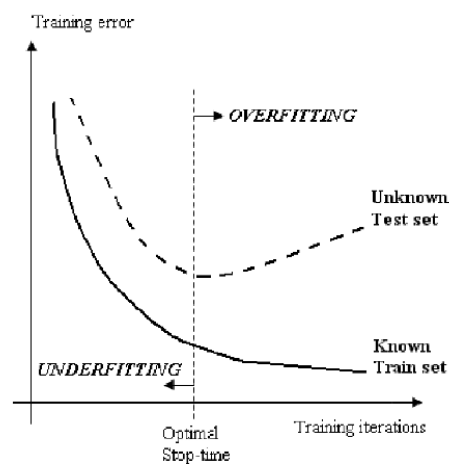


Figure I-7 : L'overfitting dans les réseaux de neurone

1.7. CONCLUSION

Nous avons présenté dans ce chapitre plusieurs types de réseaux de neurones artificiels. Nous nous sommes concentrés sur les réseaux supervisés. Nous avons décrit au début la méthode de la BP qui est très souvent utilisée dans les réseaux de neurones et quelques unes de ses variantes les plus connues.

Chapitre II.

Introduction aux Algorithmes Génétiques

II.1. INTRODUCTION

Dans ce chapitre, nous donnerons le vocabulaire nécessaire à la compréhension des algorithmes génétiques, et nous identifions les points importants de ces algorithmes.

II.2. HISTORIQUE

En 1866, Mendel publie l'article retraçant dix années d'expériences d'hybridation chez les végétaux (recombinaison des gènes) et l'adresse aux sociétés scientifiques des quatre coins du monde. Les réactions sont mitigées, voire inexistantes. Le monde scientifique n'est pas prêt à reconnaître la qualité de ses résultats. C'est seulement qu'en 1900, que la publication de trois nouveaux articles signés Hugo De Vries, Carl Correns et Erich Von-Tschermak révélant des résultats similaires à ceux de Mendel, et feront que ces premiers seront reconnus [15] [10].

C'est alors à partir du 20^{ème} siècle que la mutation génétique a été mise en évidence.

Les problèmes de traitement de l'information sont résolus de manière figée : lors de sa phase de conception, le système reçoit toutes les caractéristiques nécessaires pour les conditions d'exploitations connues au moment de sa conception, ce qui empêche une adaptation à des conditions d'environnement inconnues, variables ou évolutives. Les chercheurs en informatique étudient donc des méthodes pour permettre aux systèmes d'évoluer spontanément en fonction de nouvelles conditions : c'est l'émergence de la programmation évolutionnaire (cf. Figure II.1).

Dans les années 1960, John Holland étudie les systèmes évolutifs et, en 1975, il introduit le premier modèle formel des algorithmes génétiques (*the canonical genetic algorithm AGC*) dans son livre *Adaptation in Natural and Artificial Systems* [10]. Il explique comment ajouter de l'intelligence dans un programme informatique avec les croisements (échangeant le matériel génétique) et la mutation (source de la diversité génétique). Ce modèle servira de base aux recherches ultérieures et sera plus particulièrement repris par Goldberg qui publiera en 1989, un ouvrage de vulgarisation

des algorithmes génétiques, et ajouta à la théorie des algorithmes génétiques les idées suivantes [15] [10] :

- un individu est lié à un environnement par son code d'ADN.
- une solution est liée à un problème par son indice de qualité.

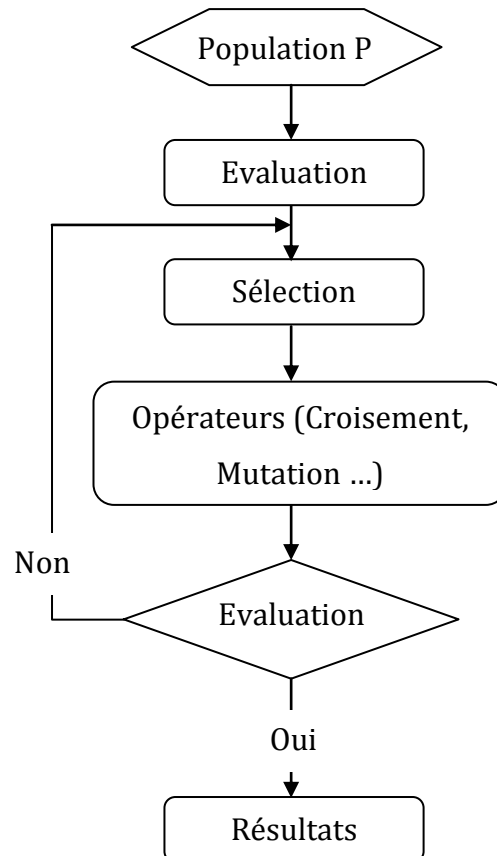


Figure I.1. Organigramme d'un algorithme évolutionnaire.

Ci-dessus est présenté l'organigramme d'un algorithme évolutionnaire. Il s'agit de simuler l'évolution d'une population d'individus divers (généralement tirés aléatoirement au départ) à laquelle on applique différents opérateurs (recombinaisons, mutations...) et que l'on soumet à une sélection, à chaque génération. Si la sélection s'opère à partir de la fonction d'adaptation, alors la population tend à s'améliorer.

Un tel algorithme ne nécessite aucune connaissance du problème : on peut représenter celui-ci par une boîte noire comportant des entrées (les variables) et des sorties (les fonctions objectifs). L'algorithme ne fait que "manipuler" les entrées, lire les sorties, "manipuler" à nouveau les entrées de façon à améliorer les sorties, etc. C'est ainsi qu'ont procédé les éleveurs pendant des millénaires : ils ont réussi à modifier, selon leurs

désirs, de nombreuses espèces animales sans connaissance en génétique ou biologie moléculaire [15] [10].

Les algorithmes évolutionnaires constituent une approche originale : il ne s'agit pas de trouver une solution analytique exacte, ou une bonne approximation numérique, mais de trouver des solutions satisfaisant au mieux à différents critères, souvent contradictoires. S'ils ne permettent pas de trouver à coup sûr la solution optimale de l'espace de recherche, du moins, peut-on constater que les solutions fournies sont généralement meilleures que celles obtenues par des méthodes plus classiques, pour un même temps de calcul [15] [10].

Le domaine de l'évolution artificielle n'a connu une réelle expansion qu'à partir de ces 20 dernières années. Pourtant, l'idée de simuler sur ordinateur des phénomènes évolutionnaires remonte aux années 50. Des concepts tels que la représentation des chromosomes par des chaînes binaires étaient déjà présents.

L'essor de l'évolution artificielle, depuis les années 80, peut s'expliquer par deux phénomènes concurrents. Premièrement, cet essor est principalement dû à l'accroissement exponentiel des moyens de calculs mis à la disposition des chercheurs, ce qui leur permet d'afficher des résultats expérimentaux pertinents et prometteurs. Le deuxième point est l'abandon du biologiquement plausible.

Trois types d'algorithmes évolutionnaires ont été développés isolément et à peu près simultanément, par différents scientifiques : la programmation évolutionniste (L. Fogel en 1966), les Stratégies d'évolution (J. Rechenberg en 1973) et les Algorithmes Génétiques (J. Holland en 1975). [15]

Dans les années 90, ces trois champs d'investigation ont commencé à sortir de leur isolement et ont été regroupés sous le terme anglo-saxon d'Evolutionary Computation. Nous traiterons seulement ici les algorithmes génétiques fondés sur l'union de la théorie de l'évolution et de la génétique moderne. Ils s'appuient sur différentes techniques dérivées : croisements, mutation, sélection... [15] [10].

II.3. LES CINQ ELEMENTS PRINCIPAUX DES ALGORITHMES GENETIQUES

Un algorithme génétique recherche le ou les extrema d'une fonction définie sur un espace de données. Pour l'utiliser, on doit disposer des cinq éléments suivants [16] [17]:

1) Un principe de codage de l'élément de population. Cette étape associe à chacun des points de l'espace d'état une structure de données. Elle se place généralement après une phase de modélisation mathématique du problème traité. La qualité du codage des données conditionne le succès des algorithmes génétiques. Les codages binaires ont été très utilisés à l'origine. Les codages réels sont désormais largement utilisés, notamment dans les domaines applicatifs pour l'optimisation de problèmes à variables réelles [16] [17].

2) Un mécanisme de génération de la population initiale. Ce mécanisme doit être capable de produire une population d'individus non homogènes qui servira de base pour les générations futures. [16] [17]

Le choix de la population initiale est important, car il peut rendre plus ou moins rapide la convergence vers l'optimum global. Dans le cas où l'on ne connaît rien du problème à résoudre, il est essentiel que la population initiale soit répartie sur tout le domaine de recherche.

3) Une fonction à optimiser. Celle-ci retourne une valeur appelée fitness ou fonction d'évaluation de l'individu.

4) Des opérateurs permettent de diversifier la population au cours des générations et d'explorer l'espace d'état. L'opérateur de croisement recompose les gènes d'individus existants dans la population ; l'opérateur de mutation a pour but de garantir l'exploration de l'espace d'états.

5) Des paramètres de dimensionnement : la taille de la population, nombre total de générations ou critère d'arrêt, probabilités d'application des opérateurs de croisement et de mutation.

Tout cela, constitue les éléments sur lesquels s'appuient des algorithmes génétiques.

Dans ce qui suit, on approfondit les mécanismes de sélection de populations et la notion de diversité qui en découle. Nous tacherons également de définir les opérateurs évoqués dans l'organigramme de l'algorithme évolutionnaire (cf. figure II.1).

Donner une image à la fois globale et précise des outils principaux des algorithmes génétiques, tel sera notre objectif majeur au cours de notre seconde partie [16] [17].

II.4. FONCTIONNEMENT DES ALGORITHMES GENETIQUES

Les algorithmes génétiques fournissent des solutions aux problèmes n'ayant pas de solutions calculables en temps raisonnable de façon analytique ou algorithmique.

D'une manière générale, un algorithme génétique manipule les solutions potentielles d'un problème donné afin d'atteindre la solution optimale ou une solution jugée satisfaisante.

Un algorithme génétique fait évoluer une population de gènes en utilisant les mécanismes de sélection naturelle. Il utilise une fonction coût basée sur un critère de performance pour calculer une « qualité d'adéquation » (fitness). Les individus les plus forts seront à même de se reproduire et auront plus de descendants que les autres.

Chaque chromosome est constitué d'un ensemble d'éléments appelés caractéristiques ou **gènes**. Le but est de trouver la combinaison optimale de ces éléments qui donne une **fitness** maximale. A chaque itération (génération de population), une nouvelle population est créée à partir de la population précédente.

Le mécanisme d'évolution et de sélection est indépendant du problème à résoudre : seules changent trois fonctions [14] [15] :

- La fonction qui représente le problème en codant chaque information caractérisant une solution possible selon un codage bien particulier. Chaque information représente alors un gène et toutes les valeurs que peuvent prendre cette caractéristique représentent les allèles possibles pour ce gène. La concaténation de tous les gènes, on obtient un chromosome qui lui représente une solution dans son intégralité
- La fonction inverse, qui à partir d'un chromosome permet d'obtenir une solution par décodage du génome.
- La fonction qui évalue l'adaptation d'une solution à un problème, sa pertinence.

Cette technique est d'application générale.

En effet, quand on utilise les algorithmes génétiques, aucune connaissance de la manière dont on résout le problème n'est requise. Il est seulement nécessaire de fournir une fonction permettant de coder une solution sous forme de gènes (et donc de faire le travail inverse) ainsi que de fournir une fonction permettant d'évaluer la pertinence d'une solution au problème donné.

Cela en fait donc un modèle minimal et canonique pour n'importe quel système évolutionnaire et pour n'importe quel problème pouvant être abordé sous cet angle, sous ce paradigme.

Cette représentation nous permet donc d'étudier des propriétés quasiment impossibles à étudier dans leur milieu naturel, ainsi que de résoudre des problèmes n'ayant pas de solutions calculables en temps raisonnables si on les aborde sous d'autres paradigmes, avec des performances quantifiables, facilement mesurables et qu'on peut confronter aux autres stratégies de résolution.

Les algorithmes génétiques peuvent être particulièrement utiles dans les domaines suivants [14] [11] :

- ✓ Optimisation : optimisation de fonctions, planification, etc.
- ✓ Apprentissage : classification, prédiction, robotique, etc.
- ✓ Programmation automatique : programmes LISP, automates cellulaires, etc.
- ✓ Etude du vivant, du monde réel : marchés économiques, comportements sociaux, systèmes immunitaires, etc.

Les principales différences des algorithmes génétiques par rapport aux autres paradigmes sont les suivantes [18] :

- On utilise un codage des informations : on représente toutes les caractéristiques d'une solution par un ensemble de gènes, c'est-à-dire un chromosome, sous un certain codage (binaire, réel, code de Gray, etc.), valeurs qu'on concatène pour obtenir une chaîne de caractères qui est spécifique à une solution bien particulière (il y a une bijection entre la solution et sa représentation codée)
- On traite une population "d'individus", de solutions : cela introduit donc du parallélisme.
- L'évaluation de l'optimalité du système n'est pas dépendante vis-à-vis du domaine.
- On utilise des règles probabilistes : il n'y a pas d'énumération de l'espace de recherche, on en explore une certaine partie en étant guidé par un semi-hasard : en effet des opérateurs comme la fonction d'évaluation permet de choisir de s'intéresser à une solution qui semble représenter un optimum local, on fait donc un choix délibéré, puis de la croiser avec une autre solution optimale localement, en général la solution obtenue par croisement est meilleure ou du même niveau que ses parents, mais ce n'est pas assuré, cela dépend des aléas du hasard, et cela et d'autant plus vrai pour

l'opérateur de mutation qui ne s'applique qu'avec une certaine probabilité et dans le cas où il s'applique choisit aléatoirement sur quel(s) locus(loi) introduire des modifications.

Un algorithme génétique générique à la forme suivante :

- 1) Initialiser la population initiale P.
- 2) Evaluer P.
- 3) Tant Que (Pas Convergence) faire :
 - a) P ' = Sélection des Parents dans P
 - b) P ' = Appliquer Opérateur de Croisement sur P '
 - c) P ' = Appliquer Opérateur de Mutation sur P '
 - d) P = Remplacer les Anciens de P par leurs Descendants de P '
 - e) Evaluer P

Fin Tant Que

Le critère de convergence peut être de nature diverse, par exemple :

- Un taux minimum qu'on désire atteindre d'adaptation de la population au problème,
- Un certain temps de calcul à ne pas dépasser,
- Une combinaison de ces deux points.

Les différents points introduits ci-avant vont maintenant être étudiés en détail dans la section VII « L'opérateur de croisement ou crossover ».

II.4.1. LE CODAGE

Chaque paramètre d'une solution est assimilé à un gène, toutes les valeurs qu'il peut prendre sont les allèles de ce gène, on doit trouver une manière de coder chaque allèle différent de façon unique (établir une bijection entre l'allèle "réel" et sa représentation codée).

Un chromosome est une suite de gène, on peut par exemple choisir de regrouper les paramètres similaires dans un même chromosome (chromosome à un seul brin) et chaque gène sera repérable par sa position : son locus sur le chromosome en question.

Chaque individu est représenté par un ensemble de chromosomes, et une population est un ensemble d'individus [10] [18].

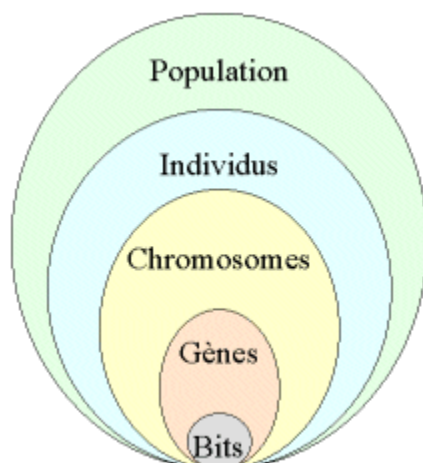


Figure II.2: les cinq niveaux d'organisation d'un algorithme génétique

Il y a trois principaux types de codage utilisables, et on peut passer de l'un à l'autre relativement facilement :

- Le codage binaire : c'est le plus utilisé.

Chaque gène dispose du même alphabet binaire {0, 1}

Un gène est alors représenté par un entier long (32 bits), les chromosomes qui sont des suites de gènes sont représentés par des tableaux de gènes et les individus de notre espace de recherche sont représentés par des tableaux de chromosomes.

Ce cas peut être généralisé à tout alphabet allélique n-aire permettant un codage plus intuitif, par exemple pour le problème du voyageur de commerce on peut préférer utiliser l'alphabet allélique {c1, c2, c3, ..., cn} où ci représente la ville de numéro i.

- Le codage réel : cela peut-être utile notamment dans le cas où l'on recherche le maximum d'une fonction réelle.

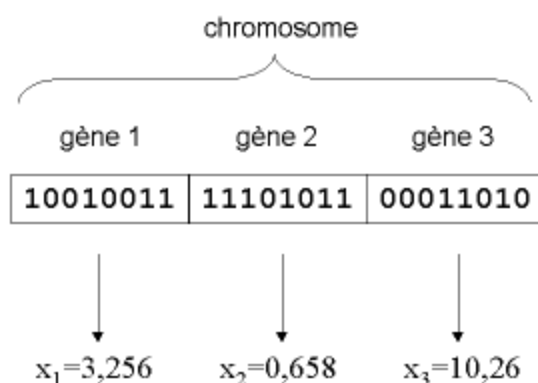


Figure II.3 : illustration schématique du codage des variables réelles

- Le codage de Gray : dans le cas d'un codage binaire on utilise souvent la "distance de Hamming" comme mesure de la dissimilarité entre deux éléments de population, cette mesure compte les différences de bits de même rang de ces deux séquences.

Et c'est là que le codage binaire commence à montrer ses limites. En effet, deux éléments voisins en terme de distance de Hamming ne codent pas nécessairement deux éléments proches dans l'espace de recherche. Cet inconvénient peut être évité en utilisant un "codage de Gray" : le codage de Gray est un codage qui a comme propriété qu'entre un élément n et un élément $n + 1$, donc voisin dans l'espace de recherche, un seul bit diffère [10] [18].

II.4.2. L'OPERATEUR DE SELECTION

Cet opérateur est chargé de définir quels seront les individus de P qui vont être dupliqués dans la nouvelle population P' et vont servir de parents (application de l'opérateur de croisement).

Soit n le nombre d'individus de P , on doit en sélectionner $n/2$ (l'opérateur de croisement nous permet de repasser à n individus).

Cet opérateur est peut-être le plus important puisqu'il permet aux individus d'une population de survivre, de se reproduire ou de mourir. En règle générale, la probabilité de survie d'un individu sera directement reliée à son efficacité relative au sein de la population.

On trouve essentiellement quatre types de méthodes de sélection différentes :

- La méthode de la "loterie biaisée" (roulette wheel) de Goldberg,
- La méthode "élitiste",
- La sélection par tournois,
- La sélection universelle stochastique.

a) La loterie biaisée ou roulette wheel :

Cette méthode est la plus connue et la plus utilisée.

Avec cette méthode chaque individu a une chance d'être sélectionné proportionnelle à sa performance, donc plus les individus sont adaptés au problème, plus ils ont de chances d'être sélectionnés.

Pour utiliser l'image de la "roue du forain", chaque individu se voit attribué un secteur dont l'angle est proportionnel à son adaptation, sa "fitness".

On fait tourner la roue et quand elle cesse de tourner on sélectionne l'individu correspondant au secteur désigné par une sorte de "curseur", curseur qui pointe sur un secteur particulier de celle-ci après qu'elle se soit arrêté de tourner [18].

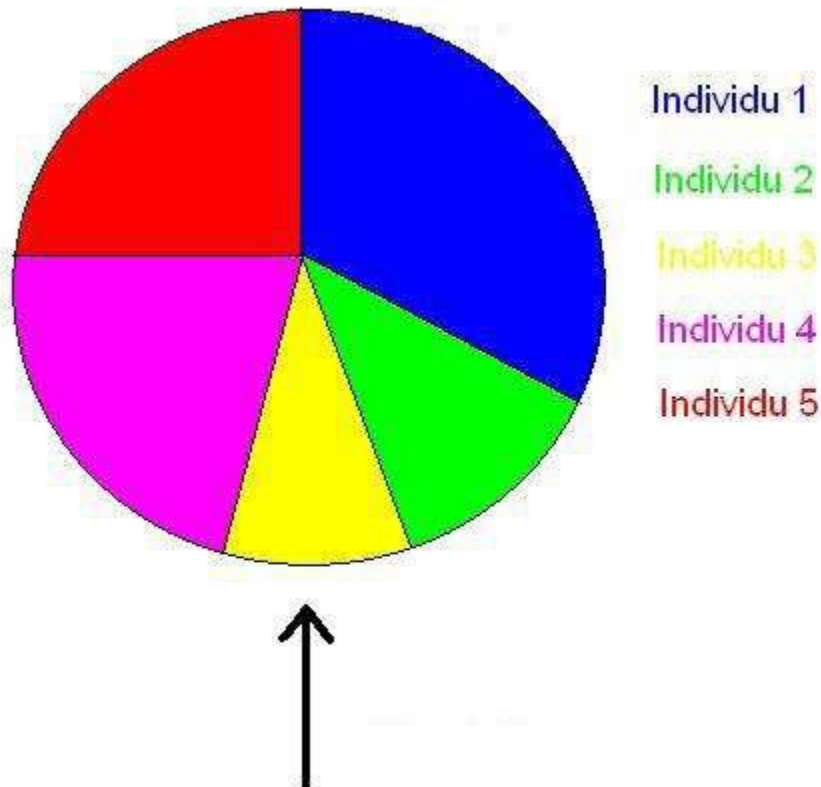


Figure II.4 : la méthode de sélection de la loterie biaisée

Cette méthode, bien que largement répandue, a pas mal d'inconvénients :

- En effet, elle a une forte variance. Il n'est pas impossible que sur n sélections successives destinées à désigner les parents de la nouvelle génération P' , la quasi-totalité, voire pire la totalité des n individus sélectionnés soient des individus ayant une fitness vraiment mauvaise et donc que pratiquement aucun individu voire aucun individu a forte fitness ne fasse partie des parents de la nouvelle génération. Ce phénomène est bien sûr très dommageable car cela va complètement à l'encontre du principe des algorithmes génétiques qui veut que les meilleurs individus soient sélectionnés de manière à converger vers une solution la plus optimale possible.
- A l'inverse, on peut arriver à une domination écrasante d'un individu "localement supérieur".

Ceci entraînant une grave perte de diversité. Imaginons par exemple qu'on ait un individu ayant une fitness très élevée par rapport au reste de la population, disons dix fois supérieure, il n'est pas impossible qu'après quelques générations successives on se retrouve avec une population ne contenant que des copies de cet individu. Le problème est que cet individu avait une fitness très élevée, mais que cette fitness était toute

relative, elle était très élevée mais seulement en comparaison des autres individus. On se retrouve donc face à problème connu sous le nom de "convergence prématurée"; l'évolution se met donc à stagner et on atteindra alors jamais l'optimum, on restera bloqué sur un optimum local.

Il existe certaines techniques pour essayer de limiter ce phénomène, comme par exemple le "scaling", qui consiste à effectuer un changement d'échelle de manière à augmenter ou diminuer de manière forcée la fitness d'un individu par rapport à un autre selon leur écart de fitness.

b) La méthode élitiste.

Cette méthode consiste à sélectionner les n individus dont on a besoin pour la nouvelle génération P' en prenant les n meilleurs individus de la population P après l'avoir triée de manière décroissante selon la fitness de ses individus.

Il est inutile de préciser que cette méthode est encore pire que celle de la loterie biaisée dans le sens où elle amènera à une convergence prématurée encore plus rapidement et surtout de manière encore plus sûre que la méthode de sélection de la loterie biaisée ; en effet, la pression de la sélection est trop forte, la variance nulle et la diversité inexistante, du moins le peu de diversité qu'il pourrait y avoir ne résultera pas de la sélection mais plutôt du croisement et des mutations.

Là aussi il faut opter pour une autre méthode de sélection [10] [18].

c) La sélection par tournois :

Cette méthode est celle avec laquelle on obtient les résultats les plus satisfaisants.

Le principe de cette méthode est le suivant : on effectue un tirage avec remise de deux individus de P , et on les fait "combattre". Celui qui a la fitness la plus élevée l'emporte avec une probabilité p comprise entre 0.5 et 1. On répète ce processus n fois de manière à obtenir les n individus de P' qui serviront de parents.

La variance de cette méthode est élevée et le fait d'augmenter ou de diminuer la valeur de p permet respectivement de diminuer ou d'augmenter la pression de la sélection [10] [18].

d) La sélection universelle stochastique :

Cette méthode semble être très peu utilisée et qui plus est possède une variance faible, donc introduit peu de diversité, nous n'entrerons donc pas dans les détails, on se contentera d'exposer sa mise en œuvre :

On prend l'image d'un segment découpé en autant de sous-segments qu'il y a d'individus. Les individus sélectionnés sont désignés par un ensemble de points équidistants [10] [18].

II.4.3. L'OPERATEUR DE CROISEMENT OU CROSSOVER

Le crossover utilisé par les algorithmes génétiques est la transposition informatique du mécanisme qui permet, dans la nature, la production de chromosomes qui héritent partiellement des caractéristiques des parents.

Son rôle fondamental est de permettre la *recombinaison* des informations présentes dans le patrimoine génétique de la population.

Cet opérateur est appliqué après avoir appliqué l'opérateur de sélection sur la population P; on se retrouve donc avec une population P' de $n/2$ individus et on doit doubler ce nombre pour que notre nouvelle génération soit complète.

On va donc créer de manière aléatoire $n/4$ couples et on les fait se "reproduire".

Les chromosomes (ensembles de paramètres) des parents sont alors copiés et recombinaison de façon à former deux descendants possédant des caractéristiques issues des deux parents.

Détaillons ce qui se passe pour chaque couple au niveau de chacun de leurs chromosomes :

Un, deux, voire jusqu'à $lg - 1$ (où lg est la longueur du chromosome) points de croisements (loci) sont tirés au hasard, chaque chromosome se retrouve donc séparé en "segments". Puis chaque segment du parent 1 est échangé avec son "homologue" du parent 2 selon une probabilité de croisement pc . De ce processus résulte 2 fils pour chaque couple et notre population P' contient donc bien maintenant n individus.

On peut noter que le nombre de points de croisements ainsi que la probabilité de croisement pc permettent d'introduire plus ou moins de diversité.

En effet, plus le nombre de points de croisements sera grand et plus la probabilité de croisement sera élevée plus il y aura d'échange de segments, donc d'échange de

paramètres, d'information, et plus le nombre de points de croisements sera petit et plus la probabilité de croisement sera faible, moins le croisement apportera de diversité.

Ci-dessous, un schéma illustrant un croisement en un point, un autre pour un croisement en deux points, et enfin un schéma représentant un croisement avec $lg - 1$ points de croisements (on notera d'ailleurs sur ce schéma que l'échange d'un segment avec son homologue ne se fait pas toujours) :

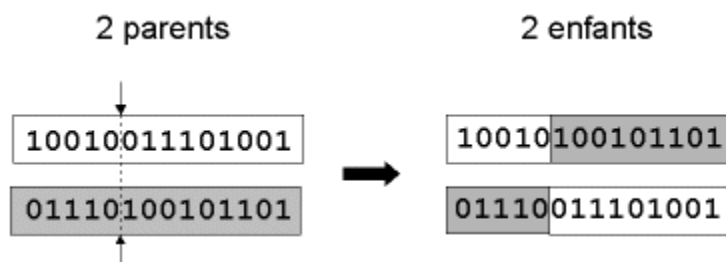


Figure 5: croisement avec un point de crossover

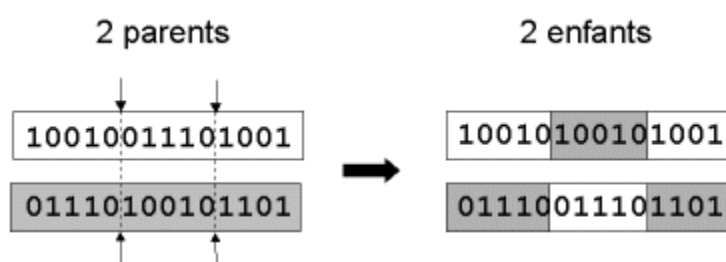


Figure 6: croisement avec 2 points de crossover

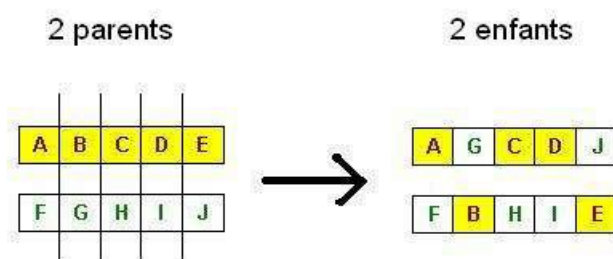


Figure II.7: croisement uniforme

On peut citer aussi une autre méthode très utilisée dans le cas des problèmes modélisés par un codage binaire, il s'agit du **croisement uniforme**. La mise en œuvre de ce procédé est fort simple, elle consiste à définir de manière aléatoire un "masque", c'est-à-dire une chaîne de bits de même longueur que les chromosomes des parents sur

lesquels il sera appliqué. Ce masque est destiné à savoir, pour chaque locus, de quel parent le premier fils devra hériter du gène s'y trouvant; si face à un locus le masque présente un 0, le fils héritera le gène s'y trouvant du parent n° 1, s'il présente un 1 il en héritera du parent n° 2. La création du fils n° 2 se fait de manière symétrique : si pour un gène donné le masque indique que le fils n° 1 devra recevoir celui-ci du parent n° 1 alors le fils n° 2 le recevra du parent n°2, et si le fils n° 1 le reçoit du parent n° 2 alors le fils 2 le recevra du parent n° 1.

L'opérateur de croisement favorise l'exploration de l'espace de recherche. En effet, considérons deux gènes A et B pouvant être améliorés par mutation. Il est peu probable que les deux gènes améliorés A' et B' apparaissent par mutation dans un même individu. Mais si un parent porte le gène mutant A' et l'autre le gène mutant B', l'opérateur de croisement permettra de combiner rapidement A' et B' et donc de créer un nouvel individu possédant cette combinaison, combinaison grâce à laquelle il est possible qu'il soit encore plus adapté que ses parents.

L'opérateur de croisement assure donc le brassage du matériel génétique et l'accumulation des mutations favorables. En termes plus concrets, cet opérateur permet de créer de nouvelles combinaisons des paramètres des composants.

Malgré tout, il est possible que l'action conjointe de la sélection et du croisement ne permette pas de converger vers la solution optimale du problème.

En effet, imaginons que nous avons une population d'individus possédant un seul chromosome.

Considérons un gène particulier de ce chromosome, on l'appellera G, gène ayant 2 allèles possibles :

0 et 1; si aucun individu de la population initiale ne possède l'allèle 1 pour ce gène, aucun croisement possible ne permettra d'introduire cet allèle pour notre gène G. Si la solution optimale au problème est telle que notre gène G possède l'allèle 1, il nous sera impossible d'atteindre cette solution optimale simplement par sélection et croisement.

C'est pour remédier entre autre à ce problème que l'opérateur de mutation est utilisé [10] [18].

II.4.4. L'OPERATEUR DE MUTATION

Cet opérateur consiste à changer la valeur allélique d'un gène avec une probabilité pm très faible, généralement comprise entre 0.01 et 0.001 [10] [18].

On peut aussi prendre $pm = 1 / lg$ où lg est la longueur de la chaîne de bits codant notre chromosome.

Une mutation consiste simplement en l'inversion d'un bit (ou de plusieurs bits, mais vu la probabilité de mutation c'est extrêmement rare) se trouvant en un locus bien particulier et lui aussi déterminé de manière aléatoire; on peut donc résumer la mutation de la façon suivante :

On utilise une fonction censée nous retourner *true* avec une probabilité pm .

Pour chaque locus **faire**

Faire appel à la fonction

Si cette fonction nous renvoie *true* **alors**

on inverse le bit se trouvant à ce locus

FinSi

FinPour

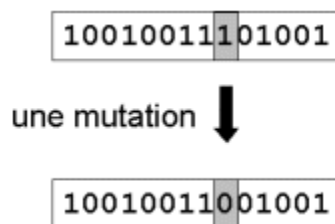


Figure 8 : une mutation

L'opérateur de mutation modifie donc de manière complètement aléatoire les caractéristiques d'une solution, ce qui permet d'introduire et de maintenir la diversité au sein de notre population de solutions. Cet opérateur joue le rôle d'un "élément perturbateur", il introduit du "bruit" au sein de la population.

Cet opérateur dispose de 4 grands avantages :

- Il garantit la diversité de la population, ce qui est primordial pour les algorithmes génétiques.
- Il permet d'éviter un phénomène connu sous le nom de **dérive génétique**. On parle de dérive génétique quand certains gènes favorisés par le hasard se répandent au détriment des autres et sont ainsi présents au même endroit sur tous les chromosomes. Le fait que l'opérateur de mutation puisse entraîner de manière aléatoire des changements au niveau de n'importe quel locus permet d'éviter l'installation de cette situation défavorable.

- Il permet de limiter les risques d'une convergence prématurée causée par exemple par une méthode de sélection élitiste imposant à la population une pression sélective trop forte. En effet, dans le cas d'une convergence prématurée on se retrouve avec une population dont tous les individus sont identiques mais ne sont que des optimums locaux. Tous les individus étant identiques, le croisement ne changera rien à la situation. En effet, l'échange d'informations par crossover entre des individus strictement identiques est bien sûr totalement sans conséquences; on aura beau choisir la méthode de croisement qu'on veut on se retrouvera toujours à échanger des portions de chromosomes identiques et la population n'évoluera pas. L'évolution se retrouvant bloquée on n'attendra jamais l'optimum global.

La mutation entraînant des inversions de bits de manière aléatoire permet de réintroduire des différences entre les individus et donc de nous extirper de cette situation.

Il est quand même utile de garder à l'esprit que ceci n'est pas une solution "miracle" et qu'il est bien entendu plus intelligent de ne pas utiliser de méthodes de sélection connues pour entraîner ce type de problème.

- La mutation permet d'atteindre la propriété d'**ergodicité**.

L'ergodicité est une propriété garantissant que chaque point de l'espace de recherche puisse être atteint.

En effet, une mutation pouvant intervenir de manière aléatoire au niveau de n'importe quel locus, on a la certitude mathématique que n'importe quel permutation de notre chaîne de bits peut apparaître au sein de la population et donc que tout point de l'espace de recherche peut être atteint.

Grâce à cette propriété on est donc sûr de pouvoir atteindre l'optimum global.

On notera que la mutation règle donc le problème exposé à la fin du Section sur le croisement [10] [18].

II.4.5. L'OPERATEUR DE REMPLACEMENT

Cet opérateur est le plus simple, son travail consiste à réintroduire les descendants obtenus par application successive des opérateurs de sélection, de croisement et de mutation (la population P') dans la population de leurs parents (la population P) [10] [18].

Ce faisant il vont remplacer une certaine proportion de ceux-ci, proportion pouvant bien sûr être choisie. Le rapport entre le nombre d'individus nouveaux allant être introduits dans la population P et le nombre d'individus de cette population est connu sous le nom de **génération gap**.

On trouve essentiellement 2 méthodes de remplacement différentes :

- Le **remplacement stationnaire** : dans ce cas, les enfants remplacent automatiquement les parents sans tenir compte de leurs performances respectives, et le nombre d'individus de la population ne varie pas tout au long du cycle d'évolution simulé, ce qui implique donc d'initialiser la population initiale avec un nombre suffisant d'individus. Cette méthode peut être mise en œuvre de 2 façons différentes :
 - ✓ La première se contente de remplacer la totalité de la population P par la population P', cette méthode est connue sous le nom de **remplacement générationnel** et on a donc un génération gap qui vaut 1.
 - ✓ La deuxième méthode consiste à choisir une certaine proportion d'individus de P' qui remplaceront leurs parents dans P (proportion égale à 100 % dans le cas du remplacement générationnel).

Ce type de remplacement engendre une population ayant une grande variation et de ce fait favorise la dérive génétique qui se manifeste d'autant plus que la population est de petite taille.

De plus dans bien des cas, étant donné que même un enfant ayant une faible performance remplace forcément un parent, on n'atteint pas la meilleure solution mais on s'en approche seulement.

- Le **remplacement élitiste** : dans ce cas, on garde au moins l'individu possédant les meilleures performances d'une génération à la suivante. En général, on peut partir du principe qu'un nouvel individu (enfant) prend place au sein de la population que s'il remplit le critère d'être plus performant que le moins performant des individus de la population précédente. Donc les enfants d'une génération ne remplaceront pas nécessairement leurs parents comme dans le remplacement stationnaire et par la même la taille de la population n'est pas figée au cours du temps.

Ce type de stratégie améliore les performances des algorithmes évolutionnaire dans certains cas. Mais présente aussi un désavantage en augmentant le taux de convergence prématuré.

Néanmoins, des implémentations plus fines procèdent de manière différente. Dans ce cas là, le taux de remplacement n'est pas de 100 %, la taille de la population augmente donc au cours des générations successives, on dit qu'il y a **overcrowding**. Il faut donc trouver un moyen pour sélectionner les parents qui seront supprimés, qui vont mourir. DeJong [12] a proposé la solution suivante : imaginons qu'on veuille remplacer 30 % des parents, soit np le nombre de parents correspondants à ce pourcentage, on remplacera les np parents les plus proches de leurs descendants de P' . Cette méthode permet donc premièrement de maintenir la diversité et deuxièmement d'améliorer la fitness globale de la population [10] [18].

II.5. CONCLUSION

On sait que les applications des algorithmes génétiques sont multiples : optimisation de fonctions numériques difficiles, traitement d'image, optimisation d'emplois du temps, contrôle de systèmes industriels, cryptographie, apprentissage des réseaux de neurones ...etc.

Les algorithmes génétiques seuls ne sont pas très efficaces dans la résolution d'un problème. Ils apportent cependant assez rapidement une solution acceptable. Néanmoins, il est possible d'améliorer cette technique assez efficacement en l'a combinant avec une technique intelligente comme par exemple les réseaux de neurones artificiels [13].

Chapitre III.

Vieillessement Thermique du PRC

III.1. INTRODUCTION

Soumis à la contrainte électrique en service, les câbles isolés au PRC peuvent être également soumis au vieillissement thermique. La contrainte thermique est l'une des contraintes de service les plus sévères pouvant affecter sérieusement les propriétés du matériau constituant l'enveloppe isolante en les dégradant et réduisant ainsi la durée de vie des câbles [19] [20].

Le vieillissement thermique peut se manifester sous deux modes différents : continu ou cyclique. Dans le premier mode, le câble est soumis à l'action de la température normale de service alors que dans le second, le câble subit des chocs thermiques suite aux régimes critiques transitoires (court-circuit, surcharge) caractérisés par de hautes températures. Les conditions de vieillissement thermique appliquées aux câbles en service constituent des facteurs de premières importances dans la connaissance de leur comportement et leur fiabilité [19] [20].

Dans ce chapitre, nous présentons les différentes caractéristiques du PRC étudiées par les chercheurs. Nous présentons également, l'influence du vieillissement thermique sur ces caractéristiques. Les résultats donnés ci-dessous sont des extraits des recherches effectuées au sein du laboratoire de haute tension de l'Ecole Nationale Polytechnique d'Alger.

III.2. PRESENTATION DU MATERIAU

Le matériau "Polyéthylène Réticulé Chimiquement (PRC)" qui a fait l'objet de notre étude, est un produit Union Carbide Référence 4201, contenant un antioxydant, le santinox R à raison de 0,2% en poids et un peroxyde organique, le peroxyde de dicumyle (2% en teneur massique). Le PRC est obtenu dans l'industrie par la réticulation du polyéthylène basse densité (PEBD) [1]. Ce matériau est constitué de chaînes macromoléculaires comprenant un certain nombre de ramifications latérales de longueurs variables. La présence des ramifications dans la chaîne de polyéthylène influe

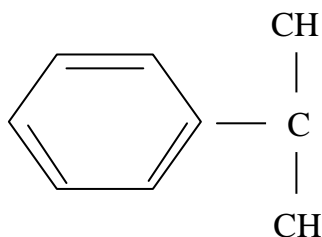
sur les propriétés physiques (densité, flexibilité, viscosité,...), et favorisent l'initiation de l'oxydation [20].

La polymérisation du PEBD se fait à une pression comprise entre 1000 et 3000 atmosphères et à une température comprise entre 100°C et 300°C. Différents catalyseurs sont employés, dont l'oxygène en premier lieu. La concentration de l'oxygène est très importante et les rendements optimaux sont obtenus avec de l'oxygène à environ 0.05% à 0.06%. Les autres catalyseurs sont les peroxydes organiques (tels que le peroxyde de benzoyle et le peroxyde bitertiaire de butyle), les alkyles métalliques et les composés azoïques. Ensuite, le PEBD est extrudé avec incorporation éventuelle de quelques additifs du type antioxydant et mis sous forme de granulés utilisables directement dans l'industrie. Le PEBD est un polymère semi-cristallin, de densité située entre 0.91 et 0.92. Le taux de cristallinité est d'environ 43% à 50% et la température de fusion cristalline est située entre 105°C et 115°C. En général, il présente une bonne résistance chimique et de bonnes propriétés mécaniques. Ce matériau non polaire est caractérisé par une permittivité relative faible (≈ 2 à 3) limitant ainsi le courant de déplacement, un facteur de pertes diélectriques relativement faible, et une rigidité diélectrique élevée [20].

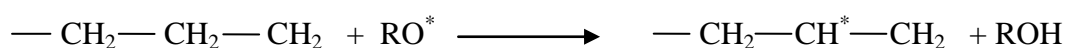
La réticulation par voie thermo-chimique sous gaz inerte chauffé s'effectue dans un tube rempli du gaz d'azote porté à une température de l'ordre de 300°C à 400°C (selon la vitesse du câble) sous une pression de 10 bars. Le refroidissement se fait dans un autre tube rempli d'eau également sous la même pression. Le procédé consiste en l'extrusion d'un mélange de polyéthylène basse densité (PEBD) et de peroxyde de dicumyle (DCP). Sous l'effet de la température de l'ordre de 180°C, le peroxyde se scinde au niveau des deux atomes d'oxygène donnant lieu à deux radicaux cumyle qui, d'une part libèrent des sites sur la chaîne macromoléculaire, d'autre part, forment des produits de décomposition suivant la réaction suivante:



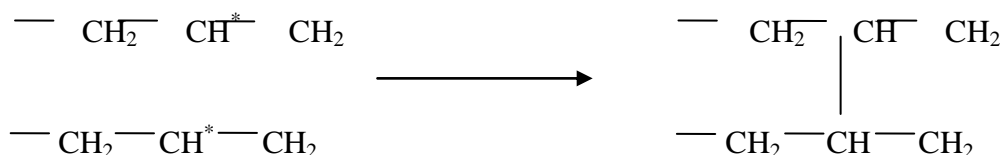
Dans le cas du peroxyde de dicumyle, le radical R présente la formule semi-développée suivante :



Les radicaux issus de la réaction sont instables chimiquement. Par conséquent, ils vont se combiner avec les hydrogènes de chaînes du polymère pour former des sites réactifs et une molécule d'alcool cumylique :



La chaîne se trouve maintenant avec un atome de carbone insatisfait. Le processus de réticulation est ainsi amorcé. Lorsque deux chaînes contenant des atomes de carbone, et dont les liaisons ne sont pas satisfaites, se trouvent assez proches, une liaison de réticulation est créée entre ces deux atomes :



Les réactions de réticulation sont complexes et font apparaître des résidus : alcool cumylique, acétophénone, éthane et méthane. Pour améliorer la stabilité thermique du PRC, l'ajout d'antioxydant (AO) pendant la phase d'extrusion est nécessaire. En présence d'antioxydant et après un vieillissement thermique, le taux de carbonyles reste constant à la périphérie du matériau. Par ailleurs, l'antioxydant engendre lui aussi des produits dérivés. En effet, il s'oxyde à la place de la matière à protéger. De là naissent de nouveaux composés de nature complexe.

La réticulation confère au PEBD des propriétés mécaniques améliorées et en particulier une meilleure résistance au fluage (déformation à chaud). Celles-ci peuvent également être améliorées par l'incorporation de charges renforçantes au polymère. Le

produit obtenu est un thermoplastique dont la tenue au vieillissement thermique est améliorée par rapport à celle du PEBD ; du fait de la réticulation, la température en régime de service est de 90°C et la température tolérée en court-circuit est de 250°C. Le PRC se carbonise sans fondre à des températures comprises entre 250°C et 300°C. Le PRC résiste bien aux solvants aliphatiques, aromatiques et au stress-cracking. Les propriétés électriques du PRC ne sont pas très différentes de celles du PEBD. Il semblerait même qu'un PRC résiste mieux à la propagation des arborescences électriques qu'un polyéthylène. Le polyéthylène réticulé chimiquement connaît actuellement un grand développement. Il est cependant très important que le mélange polymère-antioxydant-péroxyde utilisé soit le plus pur possible et que les conditions technologiques de réalisation du câble, extrusion et refroidissement soient, bien définies et contrôlées de manière à éviter la formation de vacuoles et de fissures qui pourrait nuire à la qualité du câble [20].

III.3. Travaux effectués sur la variation des caractéristiques du PRC utilisé dans les câbles MT

A fin de prédire les variations des propriétés du PRC utilisé dans l'isolation des câbles MT, vieilli chimiquement, en utilisant les réseaux de neurones artificiels, nous allons faire une synthèse bibliographique des travaux expérimentaux sur le vieillissement thermique du PRC.

Les principaux travaux traitants ce sujet sont : la thèse de doctorat du M. MECHERI, et la thèse de doctorat du M. BOUKEZZI, faites au sein du laboratoire de Haute tension de l'ENP sous la direction du M. BOUBAKEUR.

III.4. POLARISATION DANS LES POLYMERES

Dans un diélectrique imparfait, la polarisation est due à des mécanismes de déformation de la répartition des charges électriques sous l'influence du champ électrique. Dans un champ électrique alternatif, cette polarisation ne suit pas le champ d'une façon instantanée. Cette relaxation diélectrique induit une perte d'énergie. Un tel condensateur est équivalent à un circuit comportant une résistance en parallèle avec une capacité **Figure III-1** [33].

Le courant alternatif complexe I résultant de l'application de la tension alternative U aux bornes d'un condensateur plan rempli de diélectrique s'écrit :

$$I = I_C + I_P$$

Avec I_C : courant de charge.

I_P : courant de pertes d'énergie dans le diélectrique (dus à la polarisation).

Les courants I_C et I_P sont donnés par les relations suivantes :

$$I_C = j\omega C_S U \quad \text{et} \quad I_P = \frac{U}{R}$$

avec ω : pulsation du courant alternatif.

C_S : capacité d'un condensateur parfait.

R : résistance du diélectrique réel.

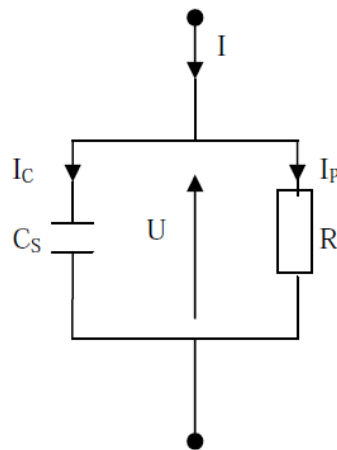


Figure III-1 Circuit équivalent parallèle représentant un diélectrique.

En pratique, I est proportionnel à $C_0 U$ et peut être écrit :

$$I = j\omega(\epsilon'_r - j\epsilon''_r)C_0 U$$

Avec C_0 : capacité du condensateur sous vide.

La polarisation du point de vue macroscopique est donc caractérisée par l'apparition de deux minces couches de charges électriques de signes opposés, chacune sur une face du matériau [33]. La relaxation est due au fait, au déphasage entre le déplacement électrique D et le champ électrique E qui fait, qu'en tension alternative, la polarisation par orientation des dipôles est plus difficile et entraîne des pertes d'énergie active, d'où [33] :

$$D = \epsilon E$$

où : ϵ est la permittivité complexe absolue du diélectrique définie comme :

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

avec ϵ_0 : la permittivité du vide ($\epsilon_0 = 8,85 \cdot 10^{-12}$ F/m).

ϵ_r : la permittivité relative complexe du diélectrique définie comme :

$$\epsilon_r = \epsilon'_r + j\epsilon''_r$$

ϵ'_r : Permittivité relative du diélectrique ou la constante diélectrique en courant alternatif.

ϵ''_r : Indice de pertes diélectriques.

La permittivité absolue ϵ définit la possibilité d'être traversé par un courant électrique, sous l'action du champ électrique. Ceci résulte de diverses contributions de déplacement de charges : déplacement des électrons, déplacements des atomes et des ions, orientation des dipôles électriques et déplacement des charges d'espace.

Dans le cas d'un diélectrique parfait, $\epsilon_r = \epsilon'_r$ est le quotient de la capacité C_s entre deux électrodes, supposées noyées dans ce diélectrique, par la capacité C_0 de la même configuration d'électrodes dans le vide :

$$\epsilon_r = \frac{C_s}{C_0}$$

La séparation des charges positives et négatives ainsi produites entraîne l'apparition de dipôles électriques qui tendent à s'aligner dans la direction du champ électrique [33]. Toutefois, un déphasage subsiste entre la direction du champ électrique et l'établissement des dipôles, entraînant une dissipation d'énergie responsable des pertes diélectriques caractérisées par $\tan \delta$, représentant le facteur de dissipation ou tangente de l'angle de pertes ou facteur de pertes diélectriques (**Figure III-2**).

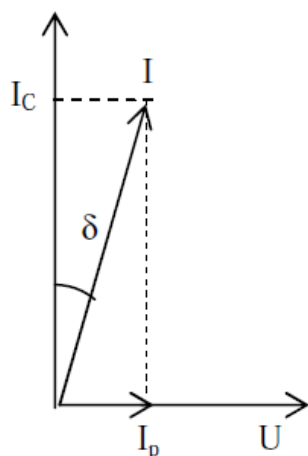


Figure III-2 Diagramme de Fesnel pour un condensateur contenant un diélectrique imparfait.

On calcule ce facteur en faisant le rapport entre la partie réelle et la partie imaginaire de l'impédance équivalente du diélectrique :

$$tg\delta = \frac{I_P}{I_C} = \frac{\epsilon''_r}{\epsilon'_r} = \frac{1}{RC_s\omega}$$

Il existe une deuxième configuration où un isolant peut être modélisé par un schéma équivalent série (**Figure III-3**).

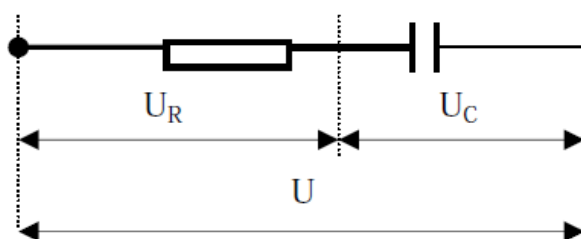


Figure III-3: Circuit équivalent série représentant un diélectrique

III.5. LES PRINCIPALES CARACTERISTIQUES ETUDIEES

Ces travaux étudient principalement les caractéristiques suivantes :

III.5.1. VARIATION DE L'INDICE DE PERTES DIELECTRIQUES EN FONCTION DU TEMPS DE VIEILLISSEMENT

L'indice de pertes diélectriques ($\epsilon_r tg\delta$) est le produit de la permittivité relative ϵ_r par le facteur de dissipation diélectrique $tg\delta$. L'utilisation de cette propriété présente l'avantage de réunir en un seul terme deux particularités du matériau, ainsi elle nous renseigne sur la puissance perdue dans le matériau polymère permettant de prévenir le risque de l'emballement thermique [20].

Sur les figures III.1, à III.4, nous donnons les résultats obtenus par M. Mecheri [20]. Ces graphes représentent la variation de l'indice de pertes diélectriques en fonction de la durée d'exposition à la chaleur pour les températures de vieillissement de 80°C et 140°C, sous les deux modes appliqués (vieillessement continu et vieillissement cyclique) et ce, pour les mêmes températures et tensions d'essai explorées dans la mesure de la permittivité relative. Pour une température de vieillissement donnée, l'indice de pertes diélectriques présente des variations en fonction du temps de

vieillessement caractérisées par la présence de pics. Les valeurs obtenues sous les deux modes précités ont le même sens de variations du point de vue qualitatif avec le temps de vieillissement et ce, pour les différents paliers de tension et à température d'essai constante. Du point de vue quantitatif, l'indice de pertes diélectriques est d'autant plus important que la valeur de la tension est plus élevée, Cette élévation des pertes diélectriques est expliquée par l'énergie dissipée par les décharges partielles [20].

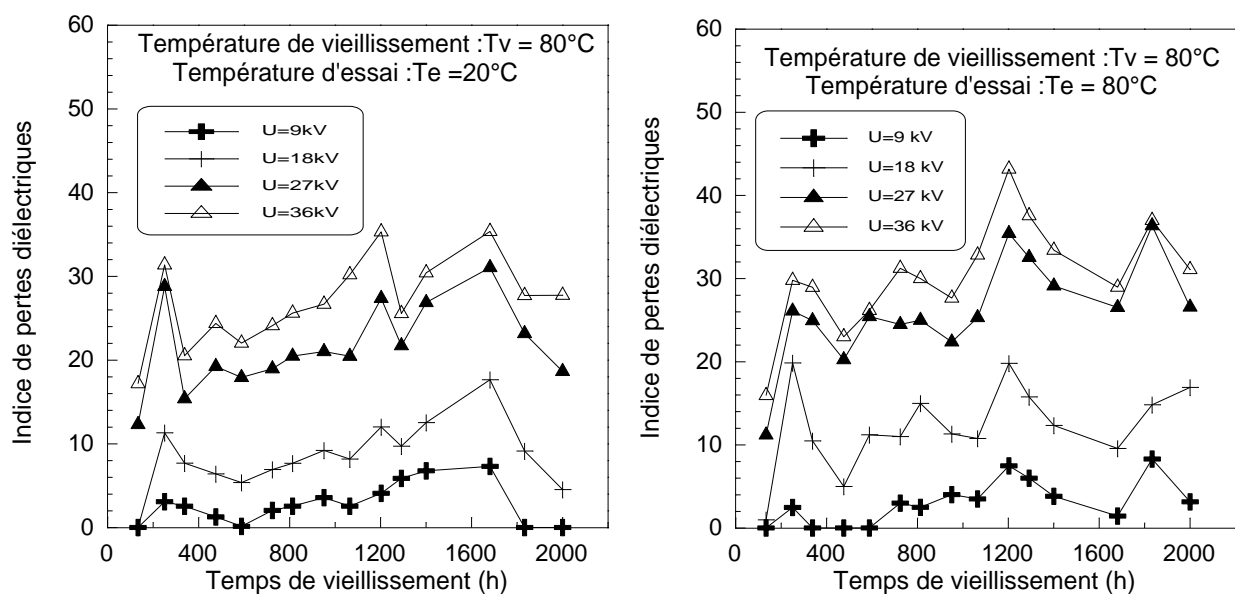


Figure III.1 Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement continu à la température de 80°C

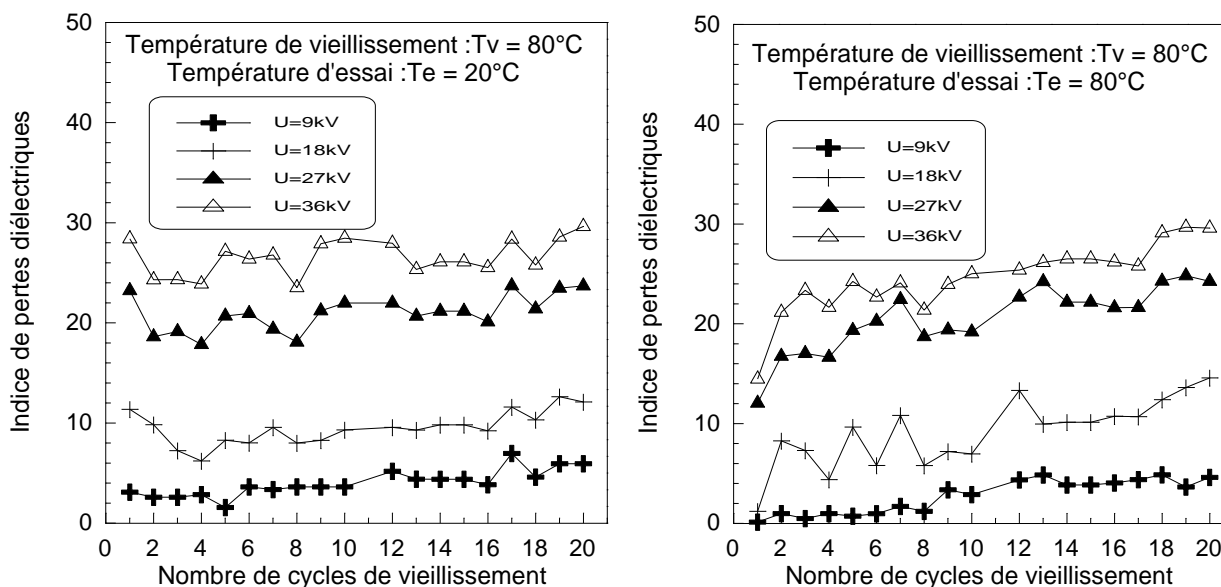


Figure III.2 Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement cyclique à la température de 80°C.

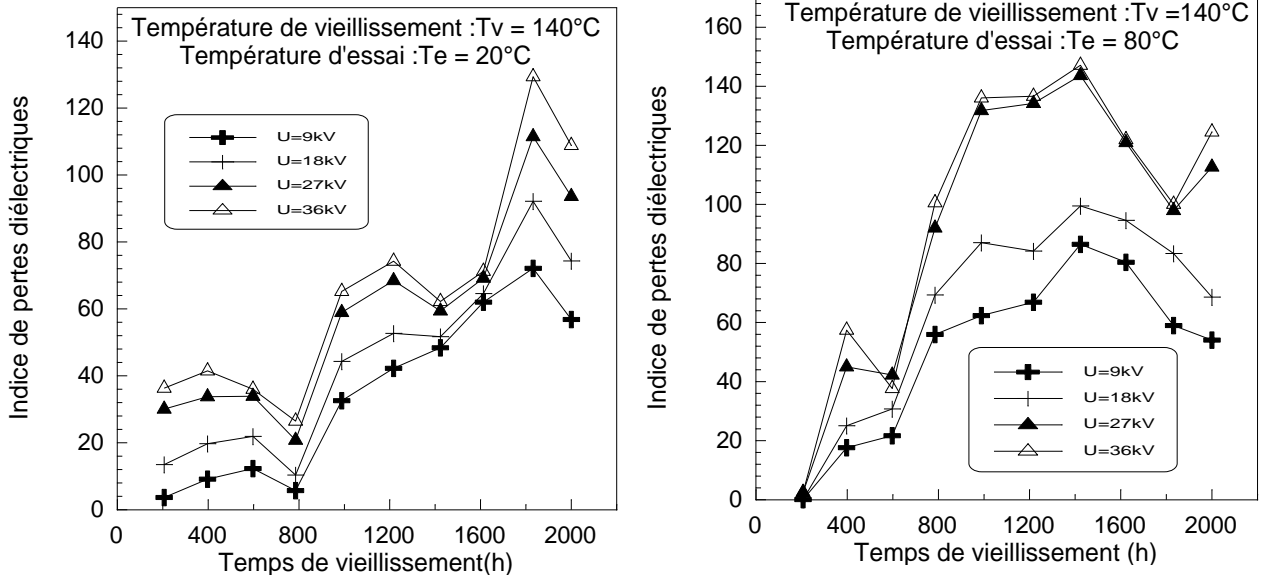


Figure III.3 Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement continu à la température de 140°C.

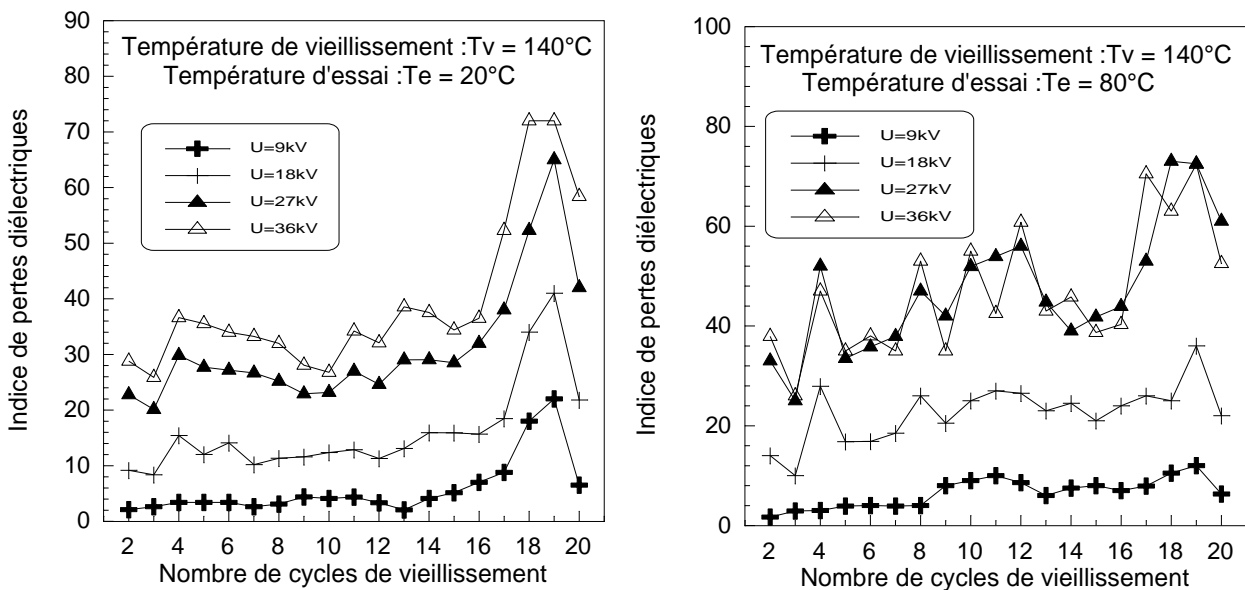


Figure III.4 Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement cyclique à la température de 140°C.

III.5.2.VARIATION DE LA CONSTANTE DIELECTRIQUE EN FONCTION DE LA TEMPERATURE D'ESSAI

L'effet des conditions du vieillissement thermique sur la variation de la constante diélectrique ϵ_r (partie réelle de la permittivité complexe) en fonction de la température a également fait l'objet d'une étude dans le travail du M. Mecheri [20]. Les essais ont été

effectués avant et après vieillissement thermique des câbles et ce, à la tension spécifiée U_0 soit 18 kV, 50 Hz et sur une gamme de température variant de 20°C à 140°C. Les résultats de mesure de la permittivité en fonction de la température d'essai pour différentes températures de vieillissement pour les deux modes sont présentés aux figures III.5 et III.6 [20].

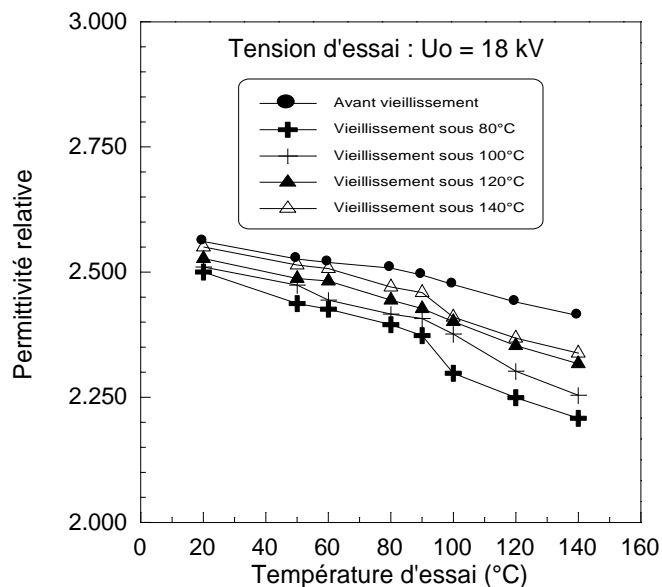


Figure III.5 Variations de ϵ_r en fonction de la température avant et après 2000 heures.

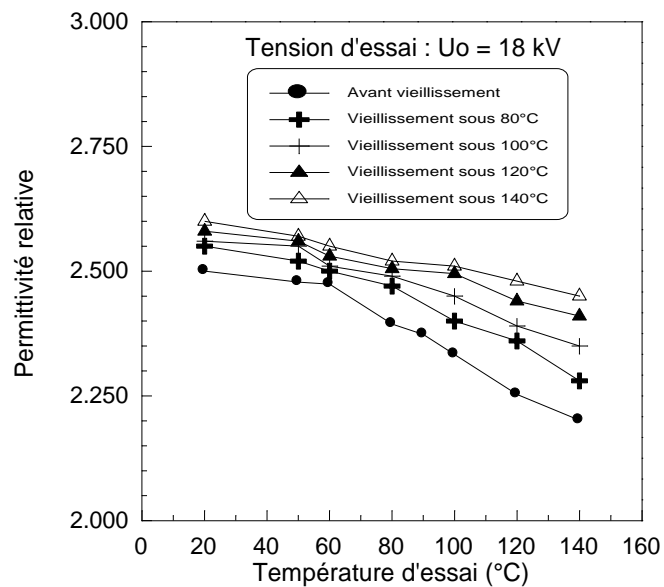


Figure III.6 Variations de ϵ_r en fonction de la température avant et après 20 cycles

Pour les deux modes de vieillissement thermique, la permittivité présente pratiquement la même variation du point de vue qualitatif. En effet, en se basant sur les résultats expérimentaux, nous remarquons que ϵ_r décroît faiblement à mesure que la température d'essai augmente (figures II.5 et II.6). Notons que les valeurs de ϵ_r obtenues après vieillissement thermique sous le mode cyclique sont supérieures à celles obtenues avant vieillissement. Un comportement contraire a été obtenu sous le vieillissement continu.

Ce résultat présente un grand intérêt d'un point de vue pratique, dans la mesure où il pourrait dans certains cas permettre de substituer un contrôle électrique à un contrôle mécanique des matériaux.

III.5.3. INFLUENCE DU VIEILLESSEMENT THERMIQUE SUR LA TENSION

SEUIL D'APPARITION DES DECHARGES PARTIELLES

Une décharge partielle est une décharge électrique dont le trajet se développe sur une partie seulement de l'intervalle isolant séparant des conducteurs. Les décharges partielles (D.P) se présentent généralement sous la forme d'impulsions individuelles qui peuvent être détectées en tant qu'impulsions électriques dans le circuit extérieur (détecteur de décharges) relié à l'objet en essai.

La mesure de la tension seuil d'apparition des décharges partielles est cruciale, car elle est révélatrice de l'état diélectrique de l'isolation et de la qualité de fabrication des câbles électriques. Dans les investigations effectuées par M. Mecheri, l'évolution de la tension seuil d'apparition des décharges partielles en fonction du temps d'exposition à la chaleur sous les deux modes différents précités a été étudiée, en prenant la température de vieillissement comme paramètre [20].

L'étude de cette caractéristique a fait apparaître une décroissance de la tension seuil d'apparition des décharges partielles en fonction de la durée et de la température de vieillissement (figures III.7 et III.8).

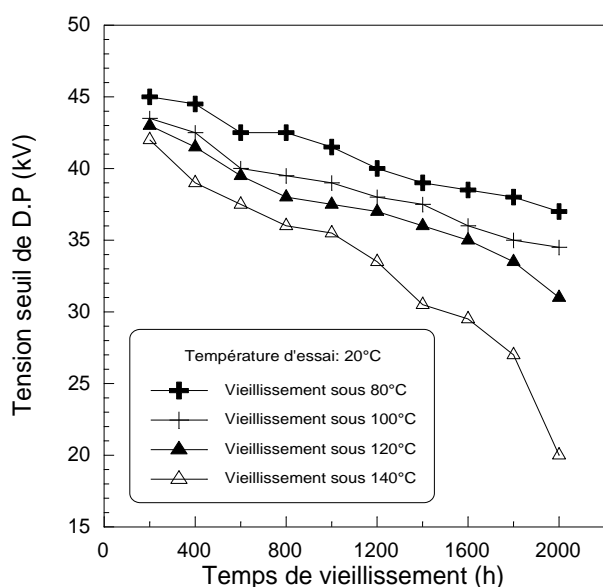


Figure III.7 Évolution de la tension seuil de D.P en fonction de la durée de vieillissement continu.

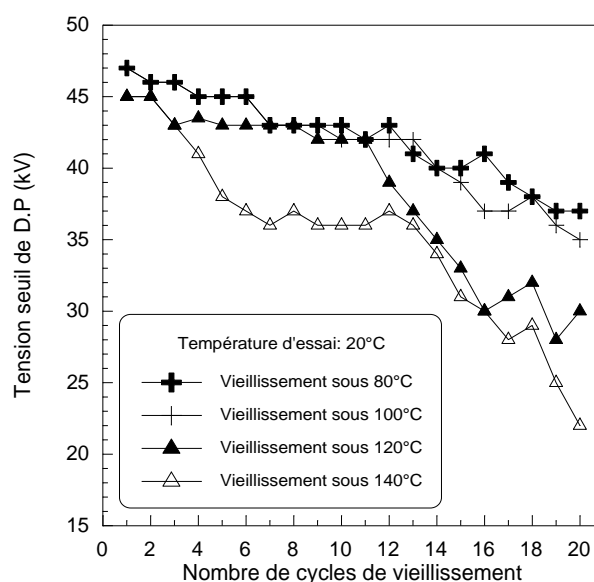


Figure III.8 Évolution de la tension seuil de D.P en fonction du nombre de cycles.

Ces défauts ou facteurs de vieillissement peuvent, plus ou moins conjointement, contribuer, en présence des contraintes électriques et thermiques, à l'altération des propriétés diélectriques et physico-chimiques du matériau et causer ainsi la défaillance du câble. La diminution du niveau de la tension seuil d'apparition des décharges partielles pourrait être expliquée par la décroissance dans le temps des teneurs des différents produits tels que les stabilisants, les antioxydants ou les produits de décomposition du peroxyde résultant du processus de réticulation sous l'effet combiné de la tension et de la température [20].

III.5.4. VARIATION DE LA PERTE DE MASSE EN FONCTION DU TEMPS DE VIEILLESSEMENT

La figure II.9 représente les variations de la perte de masse du PRC en fonction du temps de vieillissement à 80°C et 100°C. Une légère perte de masse est observée pour les deux températures de vieillissement. Cette perte de masse est due à la perte des constituants volatiles et des produits de décompositions formés durant le processus de vieillissement [19].

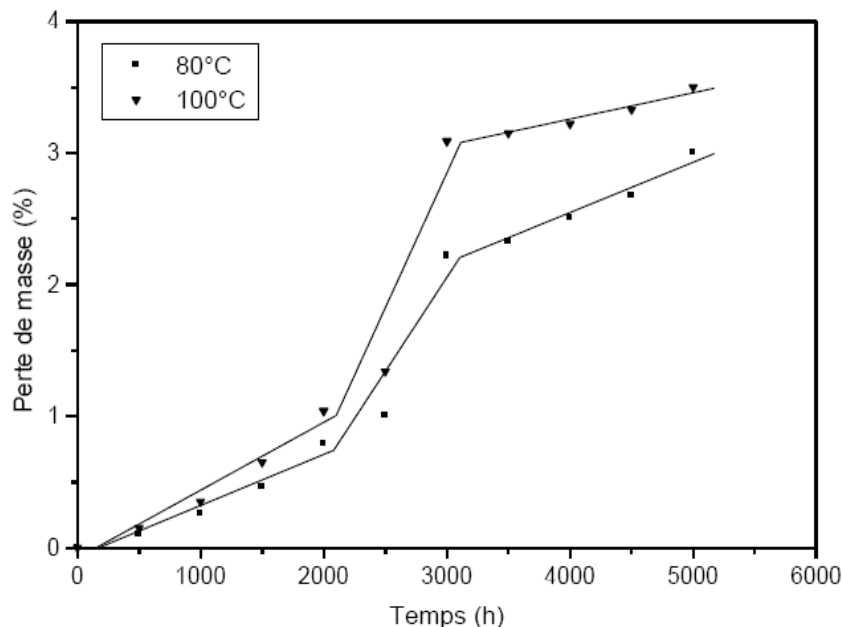


Figure III.9: Variation de la perte de masse en fonction du temps de vieillissement

D'après les résultats présentés à la Figure II.9, on constate que, pour les deux températures, la dégradation se fait selon trois phases, avec différentes vitesses. La vitesse de dégradation augmente au cours de la deuxième phase puis diminue pendant la troisième [19].

Chapitre IV.

Approche Suivie, Résultats et Discussions

IV.1. INTRODUCTION

Les réseaux de neurones (RN) utilisés dans ce chapitre sont appliqués pour prédire les variations non linéaires des caractéristiques du PRC vieilli thermiquement, utilisé dans les câbles MT.

Ces réseaux de neurones sont optimisés à l'aide des algorithmes génétiques. L'optimisation est appliquée principalement pour obtenir les meilleurs paramètres structurels (nombre de couches, nombre de neurones par couche), ce qui nous a donné les meilleurs résultats.

Pour pouvoir comparer et ensuite évaluer les types des RN utilisés dans la prédiction, nous ferons référence aux expériences effectuées dans ce domaine, et plus précisément celles effectuées par M. Mecheri (thèse de doctorat -2008-), et M. Boukezzi (thèse de doctorat -2007-), au laboratoire de haute tension de CABEL, et la Compagnie Nationale des Câbles (ex. ENICAB). Les expériences sur éprouvette ont été effectuées sous différentes températures s'étendant de 80°C à 140°C, avec un temps de vieillissement maximum de 5000 heures.

Les essais de vieillissement thermique ont été effectués sur des échantillons de câbles moyenne tension (18/30kV), isolés au polyéthylène réticulé chimiquement (PRC), d'une longueur de 12 m, fournis par l'Entreprise Algérienne des Industries des Câbles (ENICA. Biskra)

Les principales caractéristiques qui feront objet de notre étude sont les suivantes :

- Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement.
- Variation de la constante diélectrique en fonction de la température d'essai
- Influence du vieillissement thermique sur la tension seuil d'apparition des décharges partielles
- Variation de la perte de masse en fonction du temps de vieillissement

Le but de cette partie du travail est donc de prédire les caractéristiques du PRC vieilli thermiquement, utilisé dans les câbles MT, en utilisant les réseaux de neurones artificiels. Cela va réduire la période expérimentale du vieillissement.

IV.2. LA STRATEGIE SUIVIE DANS CE TRAVAIL

Le développement d'un modèle d'un réseau de neurones feed-forward pose certains problèmes. Le plus important, c'est qu'il n'y a aucune garantie antérieure pour que le modèle se comportera de façon similaire au problème actuel. Plusieurs études ont été développées pour trouver des méthodes qui guideront aux modèles les plus performants, parce qu'il n'y a aucun fond théorique quant à la façon dont cette architecture sera trouvée. La méthode la plus typique suivie est le processus « essais et erreurs », pour lequel un grand nombre de différentes architectures sont examinées et comparées. Par conséquent, ce processus est très long et est principalement basé sur l'expérience antérieure et l'intuition de l'expert humain. De ce fait, cela impliquera un niveau important d'incertitude [31].

Dans ce travail, nous proposons une méthodologie pour déterminer la meilleure architecture. Elle est basée sur l'utilisation d'un algorithme génétique (AG) et du développement des critères originaux qui mesurent les performances des RNA (apprentissage et généralisation) aussi bien que sa complexité [21] [26] [28] [29] [30] [32].

Afin de comprendre le problème il est nécessaire d'examiner de manière approfondie les caractéristiques d'un RNA feed-forward. Il y a quatre éléments qui composent les architectures d'un feed-forward RNA [31] :

- Le nombre de couches
- Le nombre de neurones dans chaque couche.
- La fonction d'activation de chaque couche.
- L'algorithme d'apprentissage.

IV.2.1. LE NOMBRE DE COUCHES CACHEES :

Le nombre de couches cachées (nc) joue un rôle crucial dans le contrôle de la capacité du réseau de neurones. Si la valeur de nc est trop petite, alors le réseau possède trop peu de paramètres et ne peut capter toutes les dépendances qui servent à modéliser et prédire les valeurs du processus observé. À l'inverse, si l'on choisit une valeur trop grande pour nc , alors le nombre de paramètres du modèle augmente et il devient possible, pendant la phase d'optimisation des paramètres, de modéliser certaines relations qui ne sont que le fruit de fluctuations statistiques propres à l'ensemble d'entraînement utilisé, plutôt que des relations fondamentales de dépendance entre les variables. Il faut comprendre que les réseaux de neurones sont des approximateurs universels, c.à.d. qu'ils peuvent modéliser n'importe quelle fonction, si le nombre d'unités cachées est suffisant [31] [23] [27].

Autrement dit, un réseau de neurones peut apprendre par cœur un ensemble d'entraînement. Afin de s'assurer que le réseau de neurones s'en tient aux relations fondamentales de dépendance, on utilise, en plus de l'ensemble d'entraînement, un second ensemble appelé *ensemble de validation* : à la fin de chaque étape d'entraînement, on mesure non seulement l'erreur d'entraînement mais aussi l'erreur de validation, c.à.d. l'erreur totale commise sur tous les exemples de l'ensemble de validation. Cette erreur de validation est calculée une fois que la phase d'optimisation des paramètres est terminée.

Après avoir entraîné quelques modèles, chacun avec un nombre différent de couches cachées, on peut comparer les erreurs d'entraînement et de

validation. On obtient généralement le résultat suivant : l'erreur d'entraînement diminue au fur et à mesure que le nombre de couches cachées augmente. L'erreur de validation, quant à elle, est élevée lorsque le nombre de couches cachées est faible, décroît avec l'augmentation du nombre de couches cachées, atteint un minimum pour un certain nombre optimal de couches cachées, puis croît lorsque le nombre de couches devient trop grand. C'est donc l'utilisation d'un ensemble de validation, distinct de l'ensemble d'entraînement, qui nous permet de choisir le nombre optimal d'unités cachées ou neurones.

Dans le reste de ce travail, nous allons fixer le nombre de couches cachées à une seule [31] [23] [27].

IV.2.2. LE NOMBRE DE NEURONES DES COUCHES CACHEES :

D'après plusieurs recherches [31], le nombre de couches cachées n'a pas une grande influence sur la fiabilité et le pouvoir de modélisation des réseaux de neurones. Les chercheurs utilisent alors généralement des réseaux avec une seule couche cachée, comme par exemple les réseaux RBF dont les paramètres les plus importants sont le nombre de neurones de la couche cachée et la nature de la fonction de transfert (ou d'activation). De ce qui précède, nous pouvons conclure que le nombre de neurones par couche cachée est d'une importance cruciale.

Plusieurs recherches et publications [23] [27] ont démontré que les performances des réseaux de neurones augmentent avec le nombre de neurones de la couche cachée, puis atteignent un maximum, avant de diminuer quand le nombre de neurones devient trop élevé.

Nous testons l'influence du nombre des neurones dans la couche cachée sur le niveau de performance des réseaux de neurones utilisés, c.à.d. nous allons fixer tous les autres paramètres (nombre de couches cachées, nature de la fonction d'activation, algorithme d'apprentissage), et faire varier le nombre de neurones de la couche cachée, puis essayer de déterminer l'évolution des performances des réseaux utilisés en fonction du nombre de neurones dans la couche cachée.

Remarque :

On parle souvent du nombre de neurones de la couche cachée, et non pas de ceux de la couche d'entrée ou de ceux de la couches de sortie, car ces derniers sont généralement déterminés à partir des exemples d'apprentissage, c.à.d. le nombre de neurones d'entrée est déterminé à partir de la dimension du vecteur d'entrée (facteurs influençant sur le phénomène), et le nombre de neurones de sortie est déterminé à partir du vecteur de sortie (vecteur cible) [31][23] [27].

IV.3. L'UTILISATION DES ALGORITHMES GENETIQUES DANS CE TRAVAIL

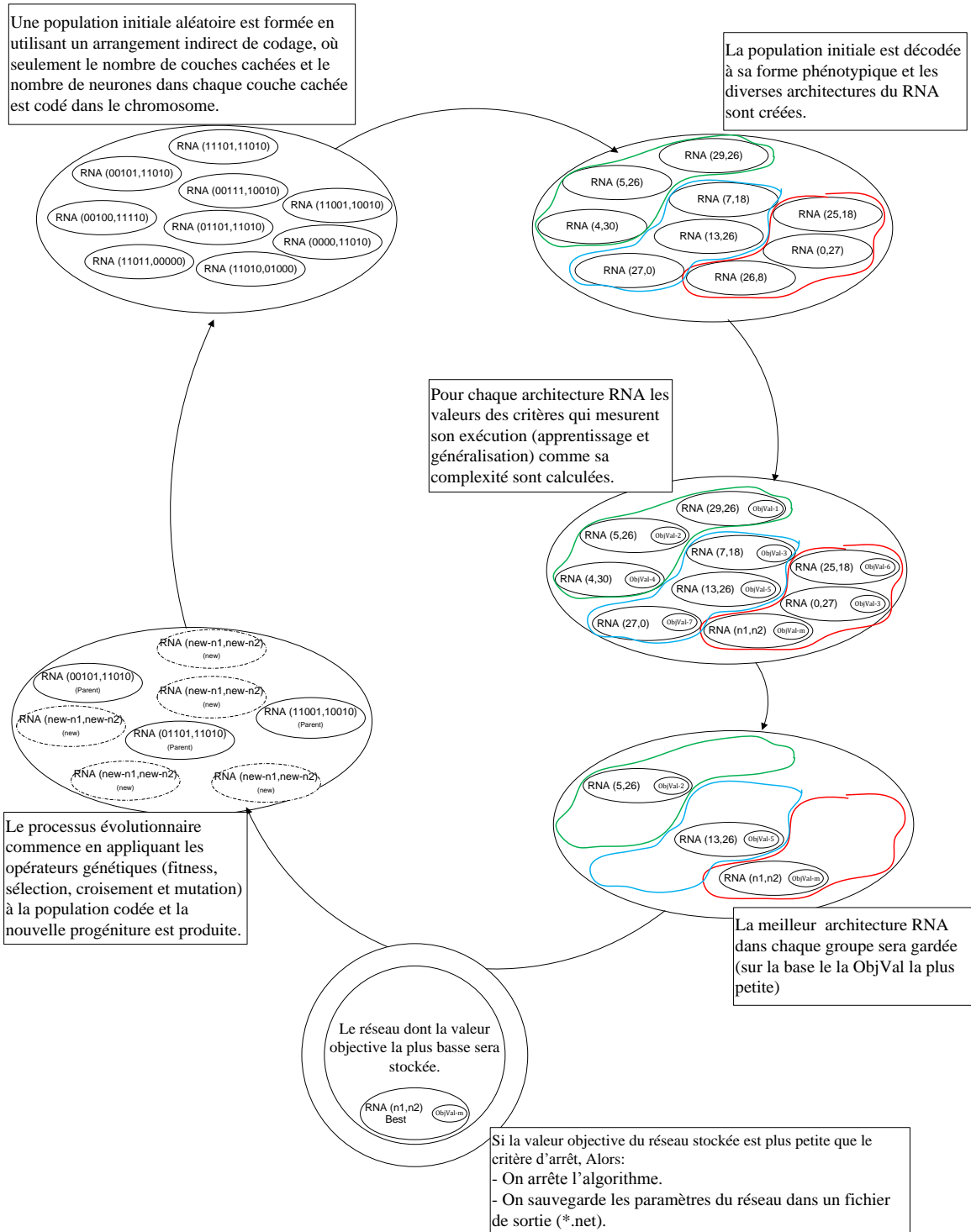
L'utilisation de la technique intelligente connue sous le nom des 'Algorithmes Génétiques', découle principalement de la problématique elle-même, car notre objectif est de trouver le réseau de neurones dont les paramètres (nombre de couches, et nombre de couche cachées) sont inconnus au début. C'est un réseau qui va nous donner une meilleur approximation de notre fonction (notre fonction ici, c'est la variation non linéaire des caractéristiques du PRC vieilli Thermiquement).

L'approche présentée dans ce travail est de considérer le problème en tant que problème d'optimisation à objectifs multiples. L'espace de solutions se compose de toutes les différentes combinaisons des couches cachées et des neurones cachés, c.-à-d. de toutes les architectures [21] [26] [28] [29] [30] [32].

Etant donné la nature complexe du problème, l'algorithme génétique est utilisé pour rechercher dans l'espace des solutions, la « meilleure » architecture, définie selon un ensemble de critères. L'idée fondamentale des algorithmes génétiques (AG) est que la solution optimale sera trouvée dans les secteurs de l'espace de solution qui contiennent de bonnes solutions et que ces secteurs peuvent être identifiés par des prélèvements robustes. Pratiquement, ceci signifie que l'espace de solution est divisé en sous-ensembles qui sont évalués afin de trouver la meilleure solution. Le procédé de prélèvement est effectué répétitivement, tout

en maintenant les meilleures solutions de chaque sous-ensemble jusqu'à ce que la solution optimale soit trouvée [22] [24] [25].

IV.4. L'ORGANIGRAMME DE L'APPROCHE SUIVIE



L'ORGANIGRAMME DE L'APPROCHE SUIVIE

IV.5. RESULTATS ET DISCUSSION

IV.5.1. LE DOMAINE DE CONFIANCE DE LA PREDICTION

La prédiction, en utilisant les réseaux de neurones artificiels, repose sur le principe de construire un model (RNA) qui porte les mêmes comportements que le phénomène à être prédit, mais dans un intervalle plus large que l'intervalle initial (d'apprentissage). Pour cela, le problème suivant se pose : « quelle est la limite du domaine de confiance, au-delà-duquel, le model ne restera plus fidèle au phénomène ».

D'après plusieurs recherches [23] [27], et d'après la confirmation dans notre travail, le model (RNA) peut rester fidèle dans un intervalle de 10 à 15 % au-delà du domaine initial (d'apprentissage). Au-delà de cet intervalle, nous ne pouvons pas assurer la fidélité de notre réseau de neurones artificiels.

Dans notre travail, nous avons opté pour une prédiction à 15 % au-delà de l'intervalle initiale (d'apprentissage).

IV.5.2. LA TAILLE DE LA BASE DE DONNEES D'APPRENTISSAGE

Parmi les points essentiels que nous pouvons signaler, c'est l'importance crucial de la taille de la base de données d'apprentissage. Malheureusement dans notre cas, nous sommes très limités en matière de données expérimentales. Par exemple pour la prédiction de la constance diélectrique, nous avons que huit (08) points expérimentaux, ce qui est très insuffisant pour avoir un bon apprentissage [31].

IV.5.3. PREDICTION DE LA VARIATION DE L'INDICE DE PERTES DIELECTRIQUES

L'indice de pertes diélectriques ($\epsilon_r \text{tg}\delta$) est le produit de la permittivité relative ϵ_r par le facteur de dissipation diélectrique $\text{tg}\delta$. L'utilisation de cette propriété présente l'avantage de réunir en un seul terme deux particularités du matériau. Ainsi, elle nous renseigne sur la puissance perdue dans le matériau polymère permettant de prévenir le risque de l'emballement thermique.

Dans les graphes qui suivent (**Figure IV-1 à Figure IV-7**), nous présentons les résultats de la prédiction de cette propriété (l'indice de pertes diélectriques) en fonction du temps de vieillissement, pour le mode de vieillissement continu, et en fonction du nombre de cycle de vieillissement, pour le mode de vieillissement cyclique.

En première étape, nous avons appliqué la technique des RNA optimisés par les AG pour but d'interpolation.

Les résultats sont représentés dans la **Figure IV-1**.

Dans les figures qui suivent, la technique des RNA optimisé par AG est utilisée pour la prédiction de la variation non linéaire des propriétés du PRC vieilli thermiquement.

Remarque Importante :

Dans tous les graphes qui suivent, les résultats représentés en "**points discrets**" (triangle, carré...) sont ceux obtenus par Boukezzi [19] et Mecheri [20].

Par contre, les résultats obtenus dans le cadre de ce travail sont représentés en "**traits continues**" (prédiction).

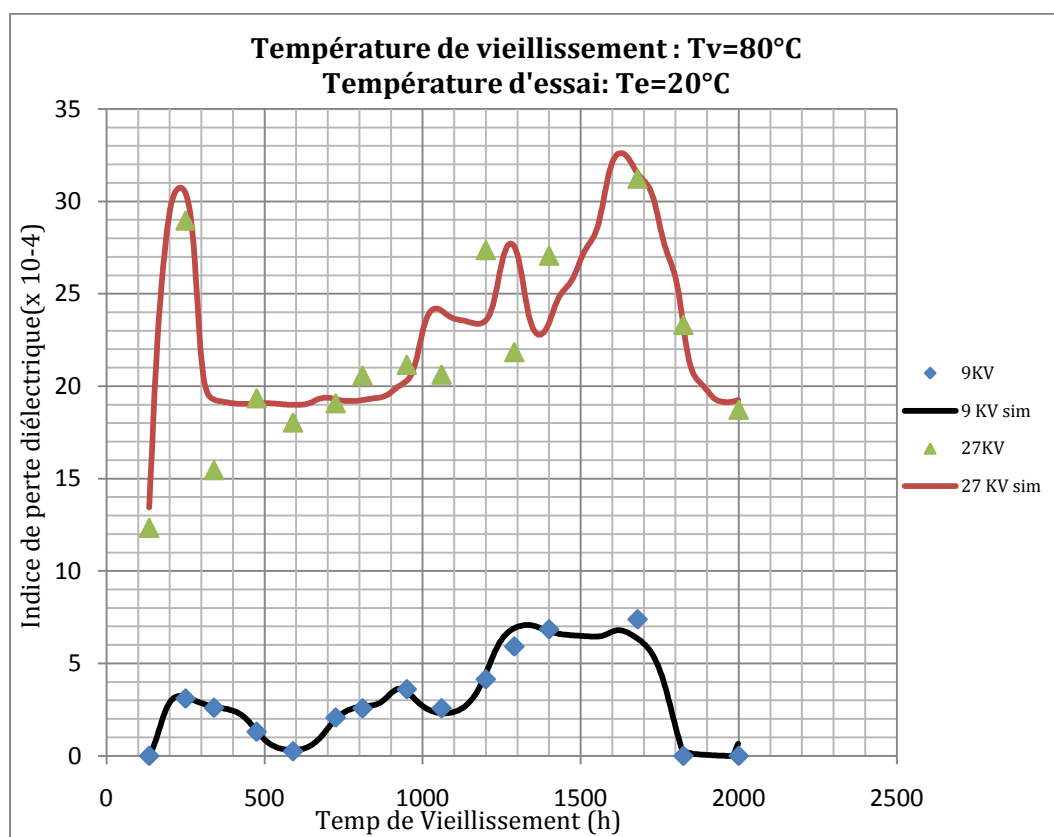


Figure IV-1 : Interpolation de la Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement continu $T_v : 80^{\circ}\text{C}$, $T_e : 20^{\circ}\text{C}$ (9KV, 27KV)

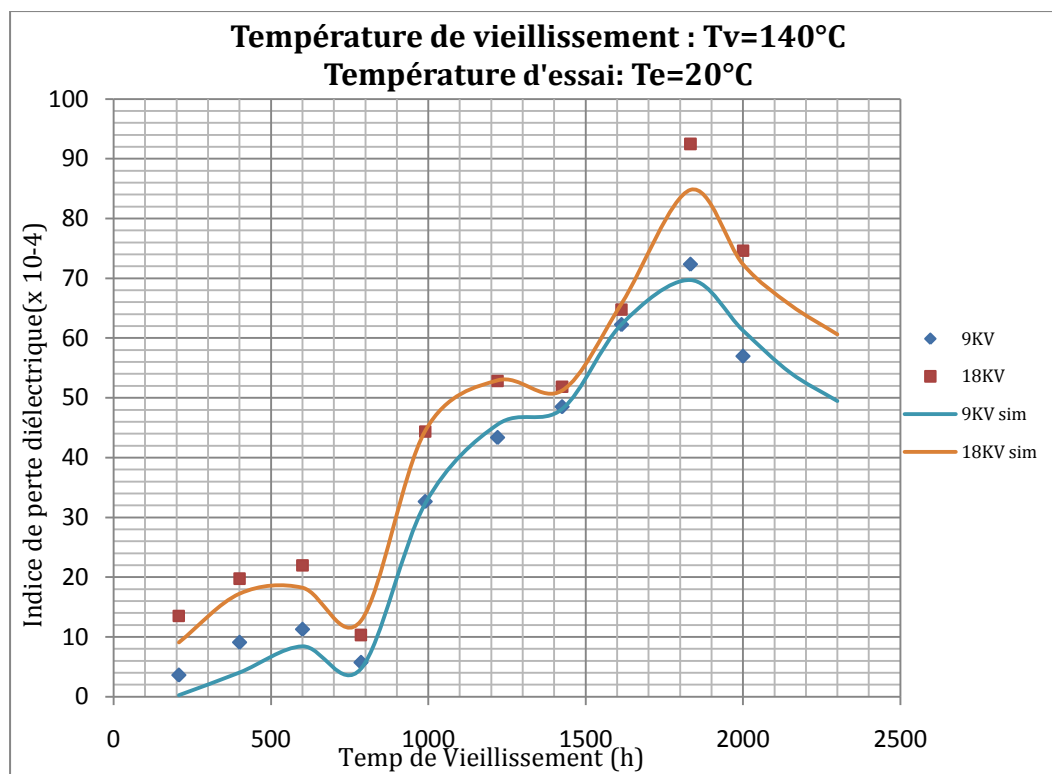


Figure IV-2 : Prédiction de la Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement continu. T_v : 140°C , T_e : 20°C (9KV, 18KV).

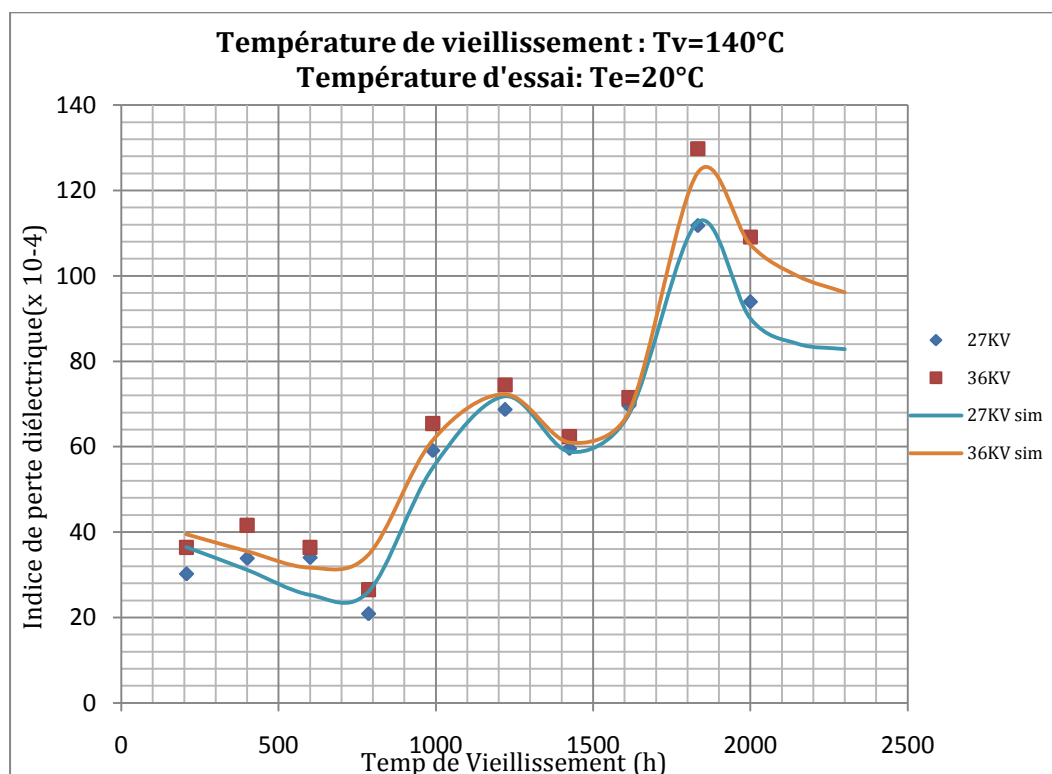


Figure IV-3 : Prédiction de la Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement continu. T_v : 140°C , T_e : 20°C (27KV, 36KV).

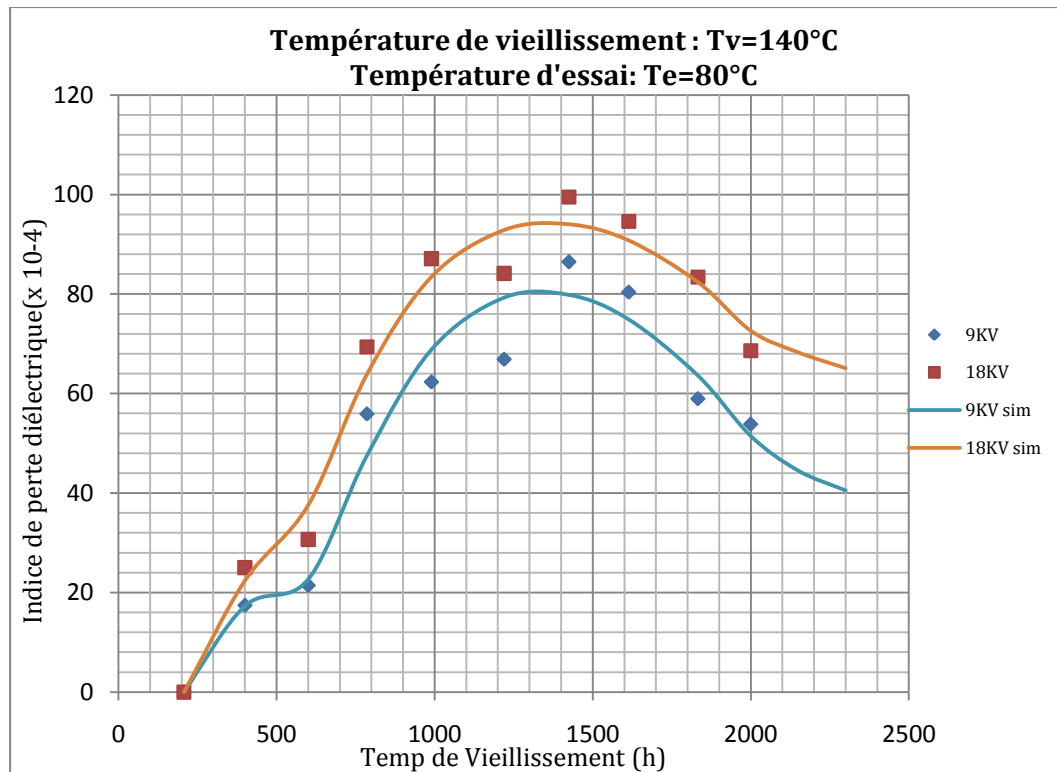


Figure IV-4 : Prédiction de la Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement continu. $T_v : 140^{\circ}\text{C}$, $T_e : 80^{\circ}\text{C}$ (9KV, 18KV).

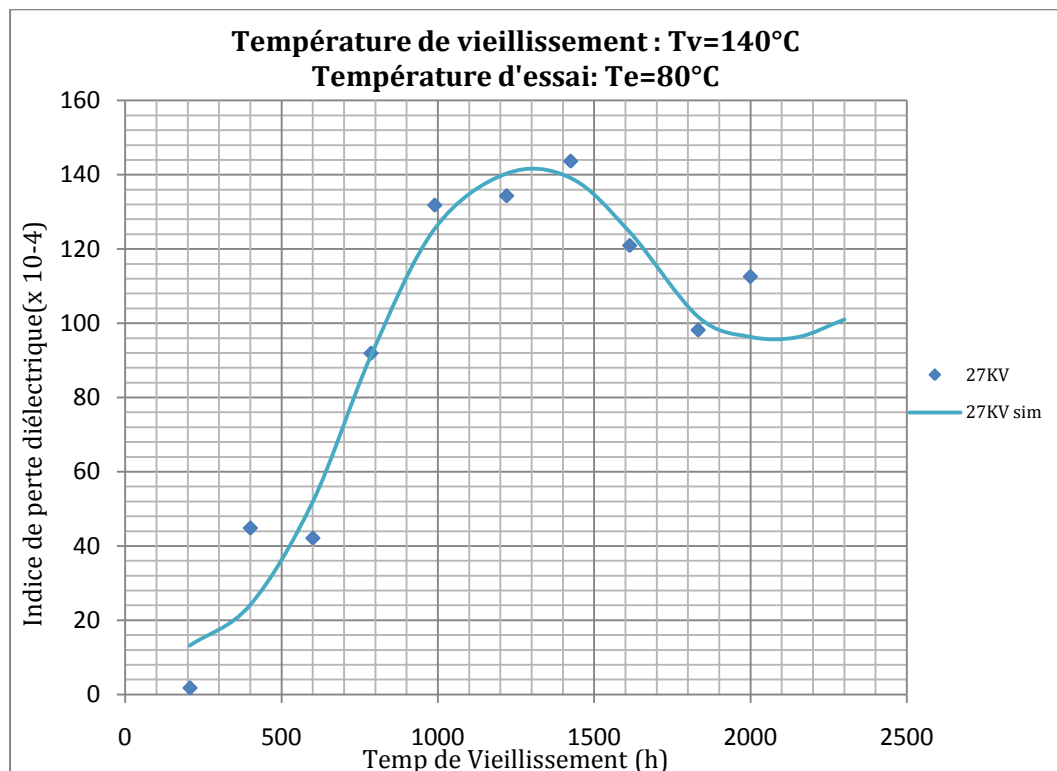


Figure IV-5 : Prédiction de la Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement continu. $T_v : 140^{\circ}\text{C}$, $T_e : 80^{\circ}\text{C}$ (27KV).

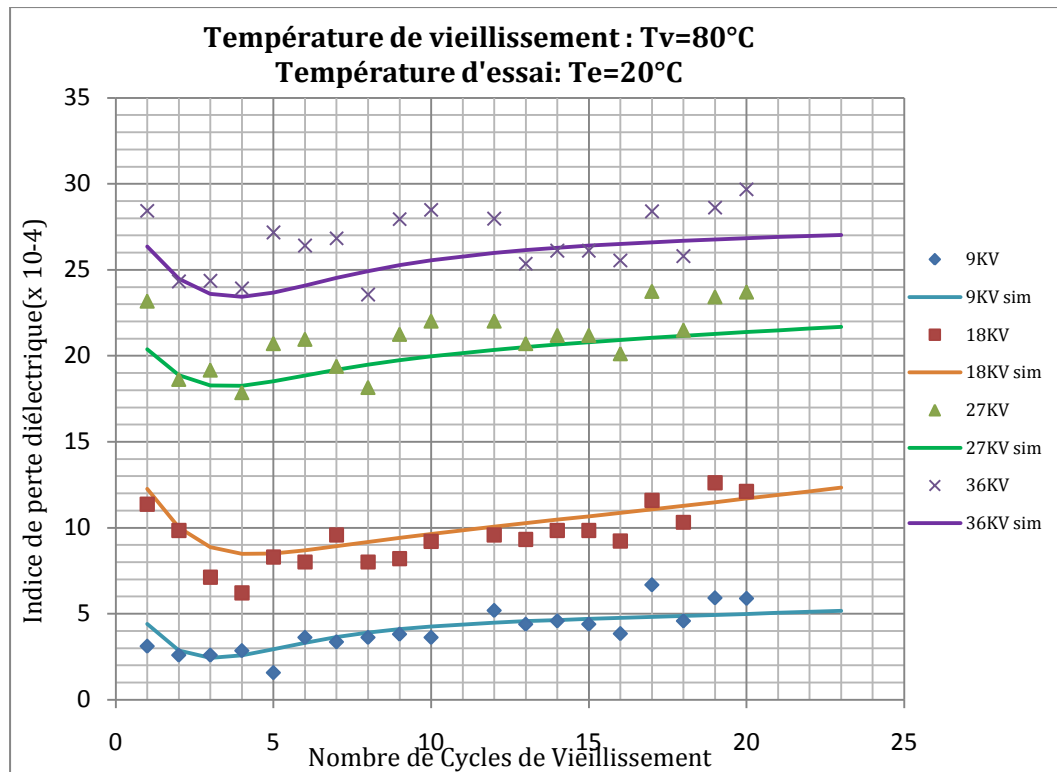


Figure IV-6 : Prédiction de la Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement cyclique. $T_v : 80^\circ\text{C}$, $T_e : 20^\circ\text{C}$ (9KV, 18KV, 27KV, 36KV).

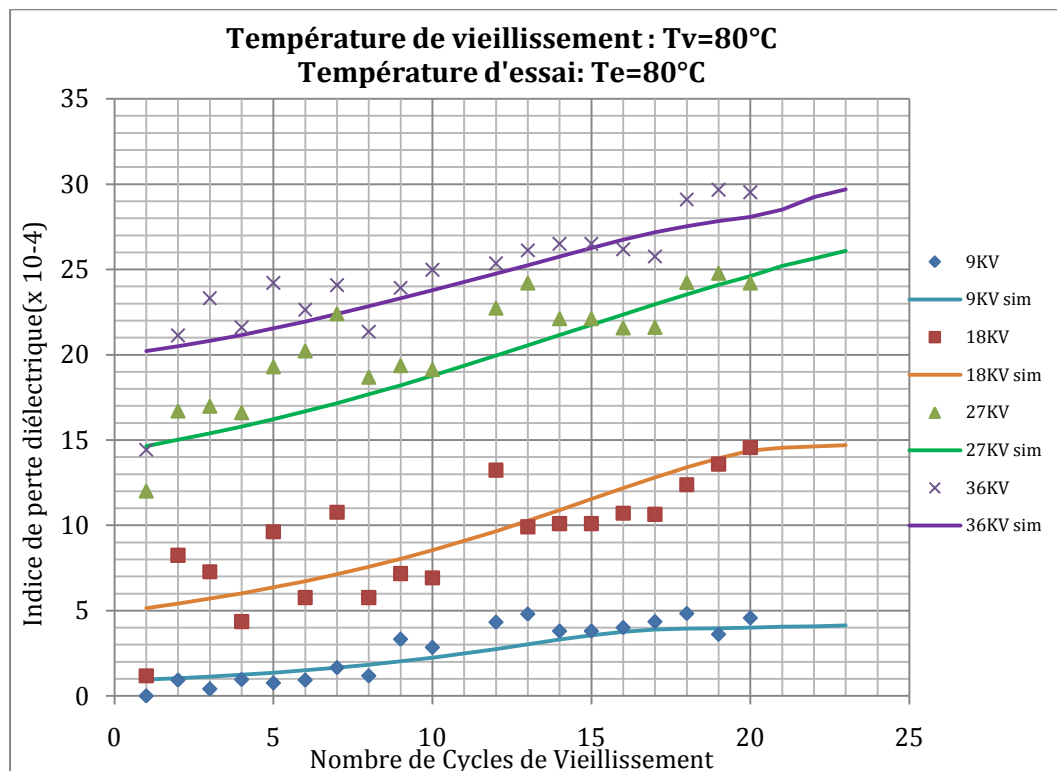


Figure IV-7 : Prédiction de la Variation de l'indice de pertes diélectriques en fonction du temps de vieillissement cyclique. $T_v : 80^\circ\text{C}$, $T_e : 80^\circ\text{C}$ (9KV, 18KV, 27KV, 36KV).

Pour l'interprétation des résultats obtenus, nos commentaires portent uniquement sur la capacité des réseaux de neurones artificiels à donner des réponses prédictives pertinentes. Les phénomènes et interactions chimiques à l'intérieur du PRC ne sont pas abordés.

En premier lieu, nous allons diviser l'intervalle de réponse en deux parties :

- A l'intérieur du domaine initial (d'apprentissage) « La fonction d'INTEROLATION » :

Nous observons que la réponse du RNA suit la tendance générale de la variation non linéaire, et ne tient pas compte des perturbations des entrées.

Ceci peut être applicable aux graphes où les points expérimentaux sont moins dispersés par rapport à la tendance générale de la variation, ce qui rend la sortie simulée, une représentation plus au moins fidèle (Figure IV-6/ 9KV et Figure IV-7/ 9KV, 36KV).

Par contre, nous voyons clairement que pour les graphes où les points expérimentaux sont plus dispersés par rapport à la tendance générale de la variation, nous ne pouvons pas dire que la sortie simulée peut être considérée comme une représentation fidele. (Figure IV-6/ 27KV, 36KV et Figure IV-7/ 18KV, 27KV).

Nous observons aussi que pour les graphes où nous avons un nombre de points relativement important, la sortie simulée suit les points expérimentaux, malgré la dispersion par rapport à la tendance générale. C'est le cas de la Figure IV-6.

- Au-delà du domaine initial (d'apprentissage) « La fonction de PREDICTION » :

La première observation que nous pouvons faire, c'est qu'il y a une continuité entre la phase interpolation et la phase prédiction, ce qui nous

montre que la prédiction des variations dans nos graphes est une extension logique.

Nous observons aussi que les variations non linéaires de l'indice de perte diélectrique continuent à suivre la même tendance de la variation dans la partie interpolation.

D'après les résultats de la prédiction de la variation non linéaire de l'indice de perte diélectrique, nous pouvons dire que l'approche suivie peut être considérée comme satisfaisante.

IV.5.4. PREDICTION DE LA VARIATION DE LA CONSTANTE DIELECTRIQUE

L'étude de la variation de la permittivité avec la température permet de suivre le comportement diélectrique du matériau (partie réelle de la permittivité complexe). En outre, elle constitue un complément précieux aux observations des propriétés mécaniques

Les essais expérimentaux ont été effectués par M. Mechri dans le cadre de son travail sur l'influence du mode de vieillissement thermique sur les propriétés du PRC.

Les essais sur la variation de la constante diélectrique ont été effectués avant et après vieillissement thermique des câbles, et ce à la tension spécifiée U_0 , soit 18 kV, 50 Hz et pour une température de vieillissement de 80°C.

Les résultats de la prédiction de la variation de la constante diélectrique en fonction de la température d'essai sont représentés sur les Figure IV-8 et Figure IV-9.

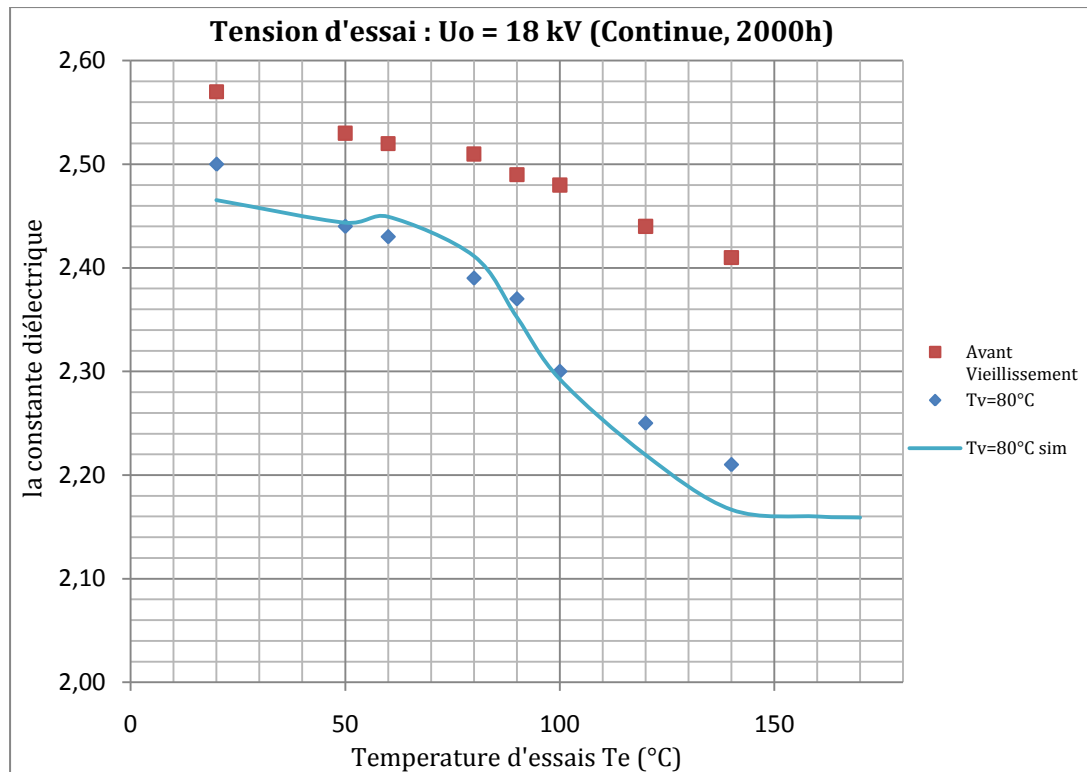


Figure IV-8 : Prédiction de la variation de la constante diélectrique en fonction de la température avant et après vieillissement continu.

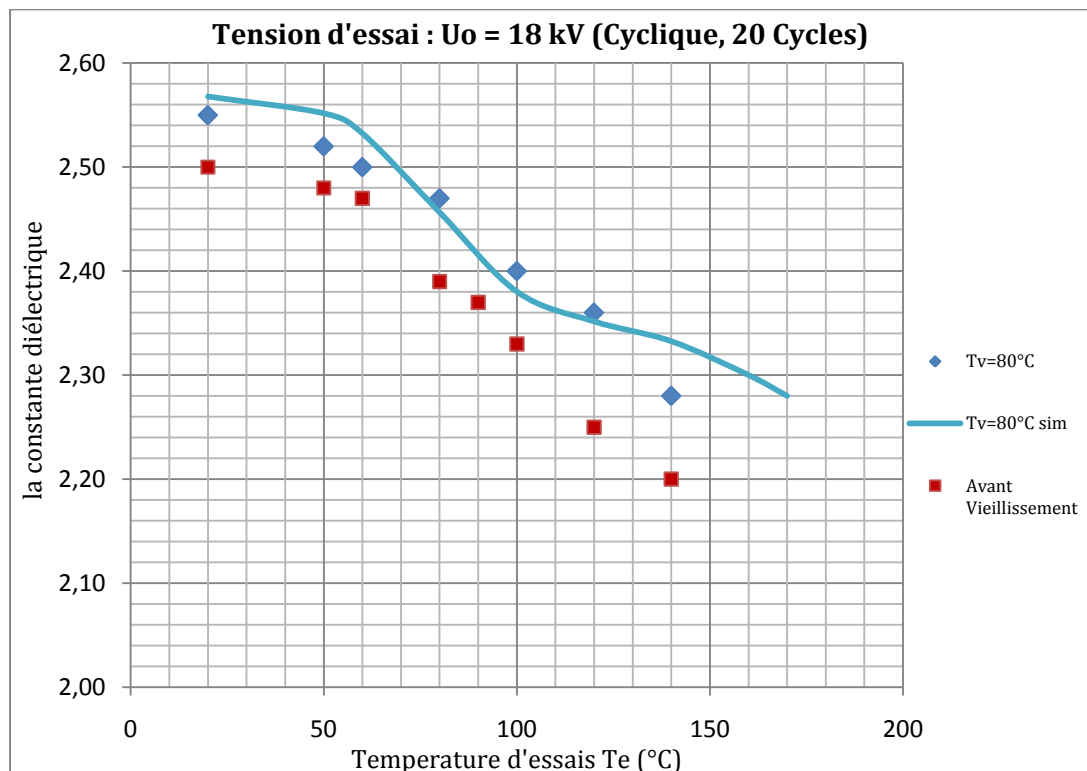


Figure IV-9 : Prédiction de la variation de la constante diélectrique en fonction de la température avant et après vieillissement cyclique.

Nous observons que la sortie simulée suit la tendance générale de la variation de la constante diélectrique en fonction de la température. Cependant, elle ne passe pas par tous les points. Ceci veut dire qu'elle ne représente pas fidèlement la base de données d'apprentissage. Cela est dû principalement à la taille de la base de données, car avec une petite taille de base de données, le réseau de neurones ne peut pas apprendre le modèle à cause de l'insuffisance des données qui ne couvre pas en détail le phénomène à être prédit.

En général, nous pouvons dire que le modèle appris a pu suivre la tendance générale de la variation, ce qui peut nous donner une idée plus au moins claire (mais pas précise) sur la variation non linéaire hors du domaine d'apprentissage.

IV.5.5. PREDICTION DE LA VARIATION DE LA TENSION SEUIL D'APPARITION DES DECHARGES PARTIELLES

Une décharge partielle est une décharge électrique dont le trajet se développe sur une partie seulement de l'intervalle isolant séparant des conducteurs. Les décharges partielles (D.P) se présentent généralement sous la forme d'impulsions individuelles qui peuvent être détectées en tant qu'impulsions électriques dans le circuit extérieur (détecteur de décharges) relié à l'objet en essai.

Il est généralement admis que les décharges partielles dans les câbles sont un des principaux facteurs de vieillissement. De ce fait, et pour prévoir la dégradation à long terme de la tenue électrique des câbles, la mesure des caractéristiques des décharges partielles (tension d'apparition et d'extinction, charge maximale apparente,...etc.) est recommandée.

Dans les graphes ci-dessous (**Figure IV-10 et Figure IV-11**), nous observons pour le vieillissement continu, que les courbes simulées (traits continues) suivent parfaitement les résultats expérimentaux (spécialement pour $T_v=120^{\circ}\text{C}$) (points discrets),

Dans la partie prédiction (au-delà de l'intervalle d'apprentissage), nous remarquons que la tension seuil de décharge partielle atteint la limite inférieure après 2000 heures pour le vieillissement continu, et après 21 Cycles pour le vieillissement cyclique. Ce qui nous montre que le PRC –après une durée très longue- sous vieillissement, atteint une limite de dégradation. Ceci provoque le changement des propriétés aussi bien mécaniques, électriques que physico-chimiques du matériau.

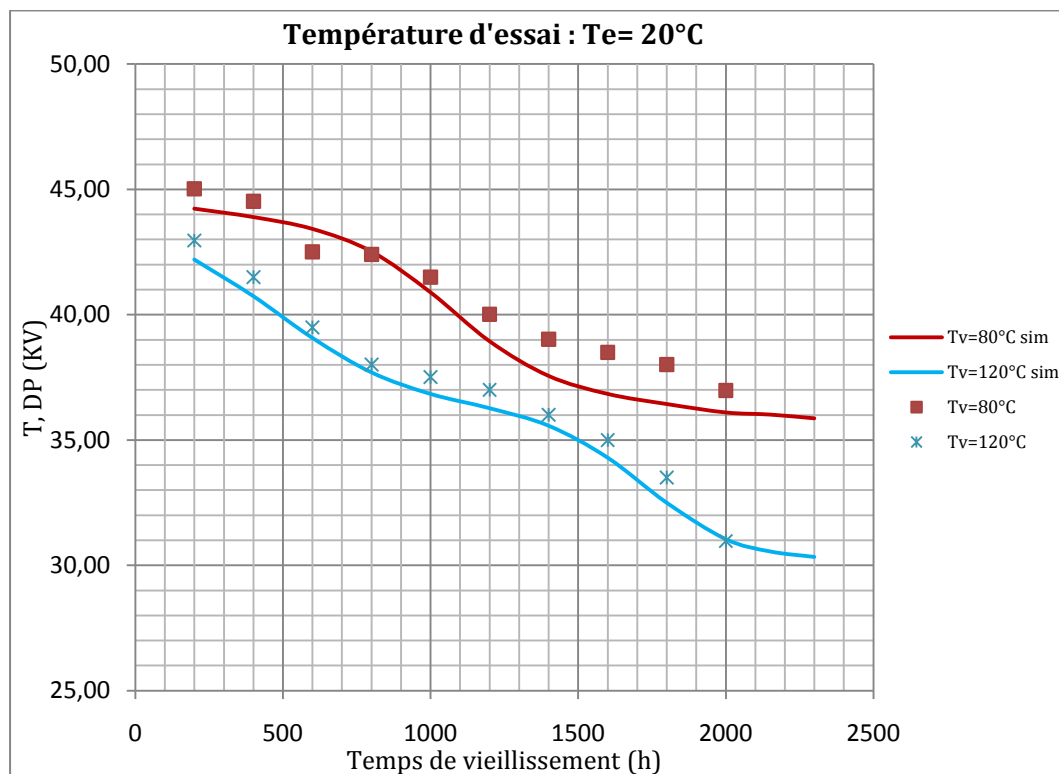


Figure IV-10 : Prédiction de l'évolution de la tension seuil de D.P en fonction de la durée de vieillissement continu.

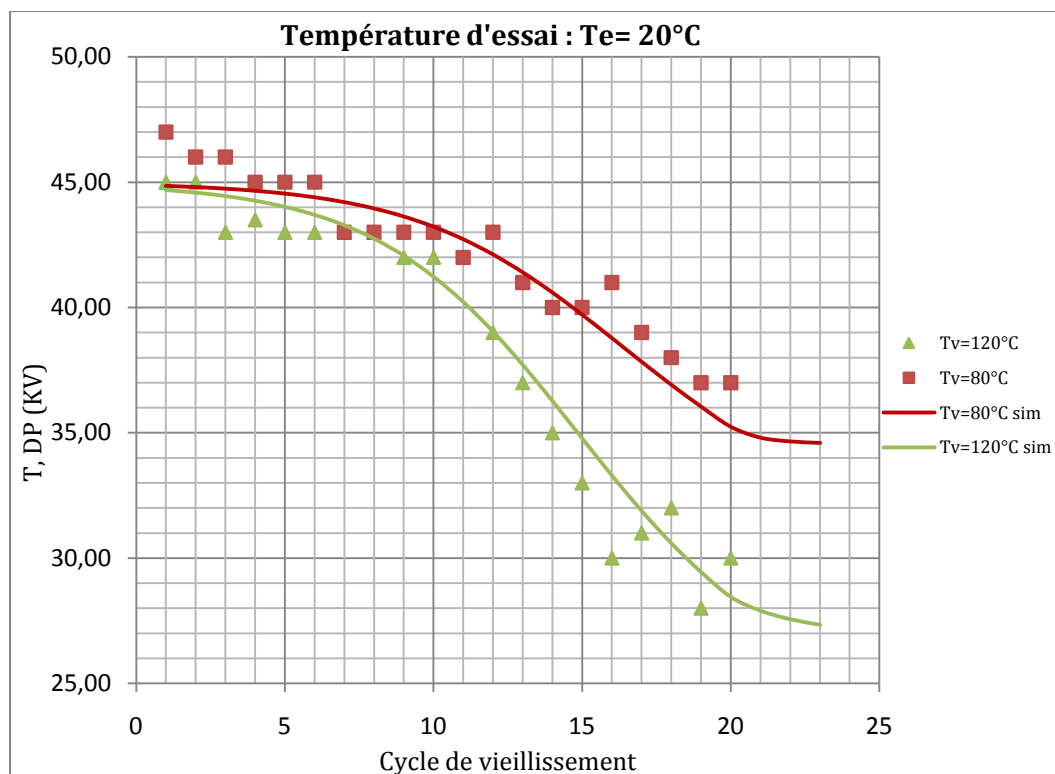


Figure IV-11 : Prédiction de l'évolution de la tension seuil de D.P en fonction de la durée de vieillissement cyclique.

IV.5.6. PREDICTION DE LA VARIATION DE LA PERTE DE MASSE

Les Figure IV-12 et IV-13 représentent les variations de la perte de masse du PRC en fonction du temps de vieillissement. Une légère perte de masse est observée pour les deux températures de vieillissement.

Pour les deux courbes ci-dessous Figure IV-12 et IV-13, nous observons que la sortie simulée est proche des résultats expérimentaux.

Dans la partie interpolation, nous observons que la perte de masse varie en plusieurs phases ; chaque phase ayant une vitesse caractéristique. M. Boukezzi a divisé l'intervalle de variation en 3 parties. Cependant, d'après les résultats de simulation par l'utilisation des RNA optimisés par les AG (Figures IV-12 et IV-13), nous voyons clairement qu'il y a plusieurs phases. Les résultats obtenus par la simulation nécessitent une validation expérimentale, en minimisant l'intervalle d'échantillonnage (chaque 100 heures au lieu de 500 heures).

Dans la partie prédiction, il est clair que la caractéristique de la perte de masse en fonction du temps de vieillissement continu à augmenter avec le temps, ce qui engendre une dégradation sévère du matériau (PRC).

En comparant les deux modes de vieillissement, nous observons qu'il y a une similitude de la variation. Ceci veut dire que le mode de vieillissement n'a pas d'influence sur cette caractéristique.

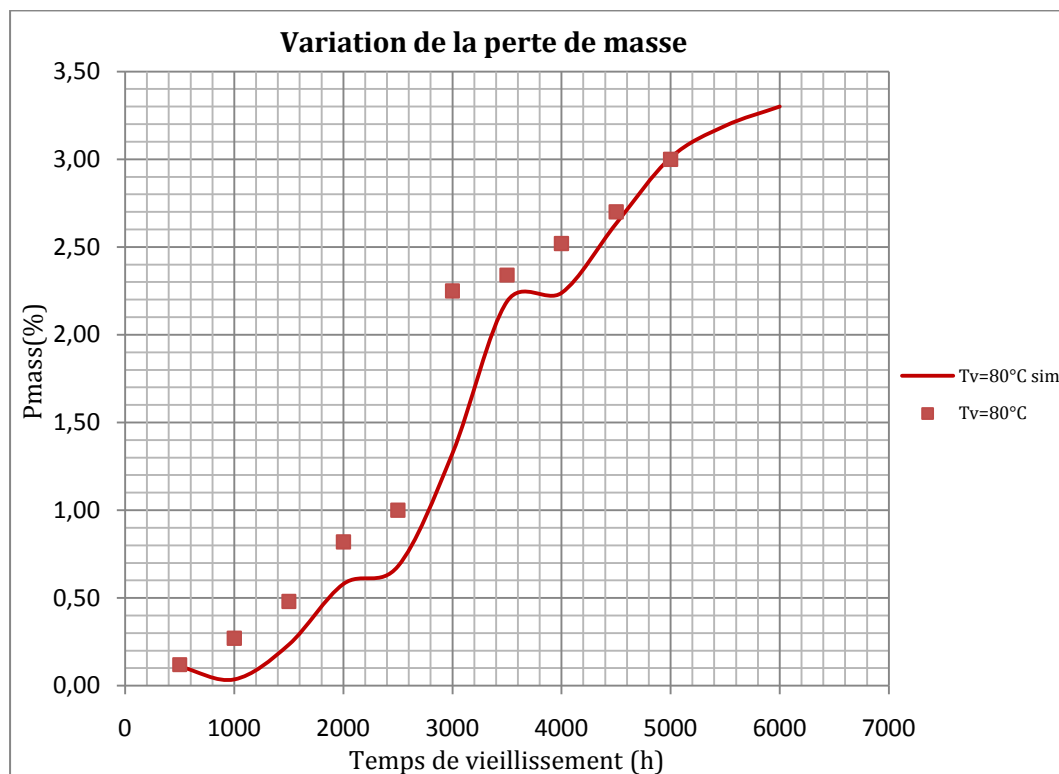


Figure IV-12 : Prédiction de la variation de la perte de masse en fonction du temps de vieillissement continu (Tv : 80°C)

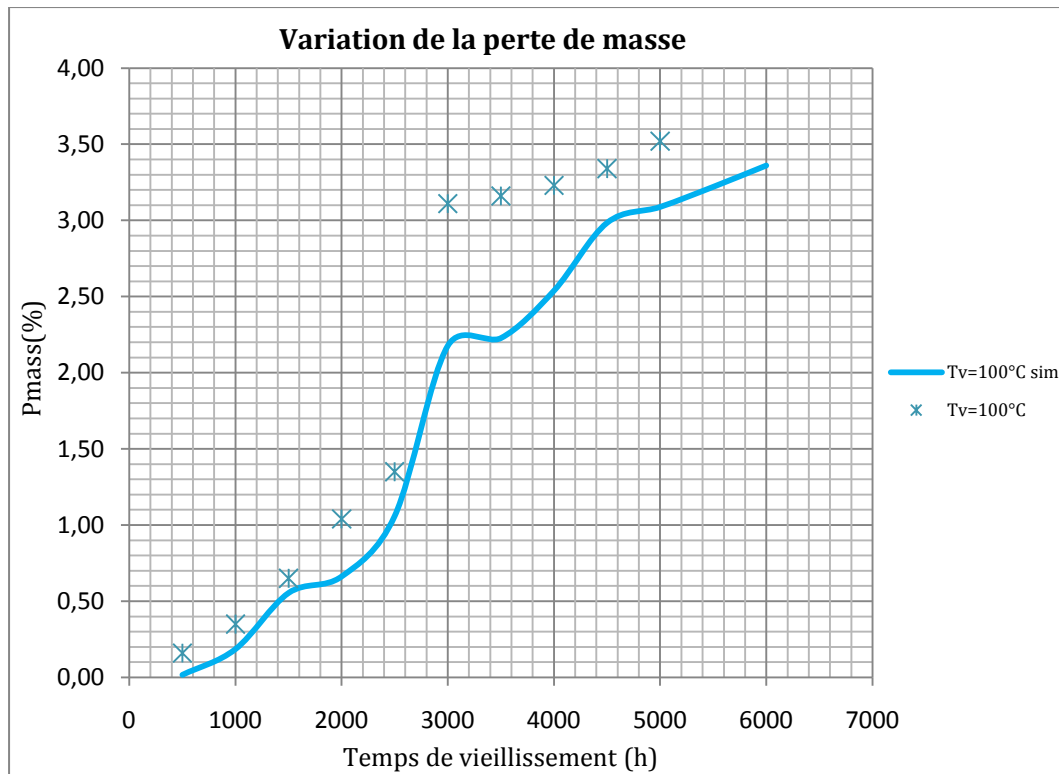


Figure IV-13 : Prédiction de la variation de la perte de masse en fonction du temps de vieillissement continu (Tv : 100°C)

Conclusion Générale

Les organes et les fonctionnalités de l'être humain ont fait l'objet de beaucoup de recherches dans différents domaines pour but de comprendre cette mystérieuse créature, pour comprendre la complexité de la morphologie microscopique du corps humain. Cette compréhension était la base principale qui a permis à une nouvelle science de naître. Il s'agit de « l'intelligence artificielle ».

L'intelligence artificielle repose sur le principe d'**inspiration**, c.à.d. observer la manière avec laquelle le corps humain fonctionne et réagit aux différents stimuli, et comprendre les différentes étapes, et à la fin créer des systèmes et algorithmes qui possèdent la même logique de fonctionnement.

Parmi les techniques les plus répandues dans la recherche scientifique et plus précisément dans le domaine de l'intelligence artificielle, on trouve : les Réseaux de Neurones Artificiels (RNA), et les Algorithmes Génétiques (AG).

Les Réseaux de Neurones Artificiels (RNA) peuvent être –grâce à son principe- utilisés dans différentes applications. L'interpolation et la prédiction sont l'une de ces applications.

Et comme pour toutes les techniques de l'intelligence artificielle, les RNA possèdent des insuffisances comme le choix des paramètres structurels essentiels (nombre de couches cachées, et nombre de neurones dans chaque couche),

Pour surmonter ce problème (choix des paramètres structurels), nous avons fait appel à une autre technique, il s'agit des Algorithmes Génétiques. Cette dernière est utilisée pour choisir la meilleure structure parmi plusieurs structures (individus) et après plusieurs générations.

Ce travail est une contribution aux recherches sur les techniques hybrides pour but de surmonter les limites et les insuffisances des techniques utilisées d'une manière individuelle.

L'approche suivie est appliquée à la prédiction des variations non linéaires du PRC vieillissant thermiquement.

Dans les deux premiers chapitres, nous avons présenté une introduction aux techniques utilisées (RNA et AG), et cela pour donner au lecteur une idée de base sur ces techniques.

Dans le troisième chapitre, nous avons présenté les travaux de recherche sur le vieillissement thermique du PRC utilisé dans les câbles moyenne tension. Il s'agit des expériences effectuées par M. Mecheri (thèse de doctorat -2008-), et M. Boukezzi (thèse de doctorat -2007-), au laboratoire de haute tension de CABEL, la Compagnie Nationale des Câbles (ex. ENICAB), et ça pour pouvoir comparer et ensuite évaluer les RNA utilisés dans la prédiction.

Dans le quatrième chapitre, nous avons présenté l'approche suivie, et les résultats obtenus de l'application de la technique hybride (RNA & AG).

D'une manière générale, nous pouvons dire que les Réseaux de Neurones Artificiels optimisés par Algorithmes Génétiques sont applicables à ce type de problème (la prédiction des fonctions non linéaires), mais ceci n'empêche pas de dire qu'il y a des conditions et des circonstances qui contribuent à l'amélioration des résultats de la prédiction.

On peut citer :

- La taille de la base de données : elle joue un rôle primordial sur la capacité d'apprentissage des réseaux de neurones. Lorsque la base de données est de taille petite, le réseau de neurones ne peut pas apprendre le modèle sur les résultats expérimentaux, et le modèle appris devient incapable de simuler le phénomène à prédire.
- La fréquence d'échantillonnage des résultats expérimentaux : quand on a une base de données qui ne représente pas parfaitement le phénomène à être prédit, le modèle réseau de neurones ne reflète pas exactement le phénomène, car il y a des changements et des réponses qui ne sont pas représentés dans la base de données.

Ce travail peut être considéré comme une contribution à l'étude des techniques hybrides, pour but de contourner les insuffisances et les besoins des techniques utilisées d'une manière individuelle.

Ce travail est évidemment non exhaustif. Il Pourra donner lieu à des études complémentaires dans les directions suivantes :

- L'optimisation des paramètres non structurels des réseaux de neurones artificiels, comme les fonctions d'activation, et les algorithmes d'apprentissage.
- L'utilisation d'autres techniques en matière d'optimisation.
- Valider les résultats obtenus par prédiction, à l'aide des recherches expérimentales.

Bibliographie

- [1] Marc Parizeau, "Réseaux de Neurones", Cour d'université, université de Laval, Canada, Hiver 2006.
- [2] Claude Touzet, "Les Reseaux de Neurones Artificiels Introduction au Connexionnisme", Cour d'université, Université de Provence, 1992.
- [3] Leïla MOKHNACHE, ' Application des réseaux de neurones dans le diagnostic et la prédiction des isolations de haute tension'. Thèse de Doctorat d'état, ENP, Décembre 2003.
- [4] YU HEN HU, "Handbook of Neural Network Signal Processing", Livre, CRC PRESS, 2002
- [5] Simon Haykin, "Neural Networks A Comprehensive Foundation 2nd Edition", Livre, Pearson Prentice Hall, 2005.
- [6] Madan M. Gupta, "Static and Dynamic Neural Networks", Livre, John Wiley & Sons Publication, 2003.
- [7] Dave Anderson, "Artificial Neural Network Technologies", Livre, Kaman Sciences Corporation, 1992.
- [8] E.DAVALO, P.NAÏM, "Des réseaux de neurones ", Livre, Edition Eyrolles, 1990.
- [9] Daniel Eggert, "Neural Network Control", Livre, Technical University of Denmark, 2003.
- [10] S.N.Sivanandam et S.N.Deepa , "Introduction to Genetic Algorithms", Livre ,Springer- Verlag Berlin, Heidelberg , 2008.
- [11] Gang Zhao et Wenjuan Luo, Huiping Nie, Chen Li, "A Genetic Algorithm Balancing Exploration and Exploitation for the Travelling Salesman Problem", Article, Fourth international Conference on Natural Computation , volume1 , 2008.
- [12] Randy L. Haupt et Douglas H. Werner, "Genetic Algorithms in Electromagnetics", Livre, WILEY-INTERSCIENCE,2007.
- [13] Abhinav Saxena et Ashraf Saad , "Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems", Article, ELSIVER, applied soft computing, Science Direct 2005.
- [14] Randy L. Haupt et Sue Ellen Haupt, "Practical Genetic Algorithms", seconde edition , A Wiley-Interscience publication,2004.
- [15] Jean-Marc Alliot, Nicolas Durand, "Algorithmes génétiques", Livre, Cepadues, 1992.
- [16] M. Mitchell, S. Forrest, "Genetic Algorithms and Artificial Life", Article, Santa Fe Institute Working Paper, 1993.
- [17] Mehrdad Dianati, Insop Song, and Mark Treiber "An Introduction to Genetic Algorithms and Evolution Strategies", Article, University of Waterloo, Ontario, Canada.
- [18] Marc Schoenauer, "Les algorithmes évolutionnaires : état de l'art et enjeux", Article, Algorithms Seminar 2002.
- [19] L. Boukezzi, "Influence du vieillissement thermique sur les propriétés du Polyéthylène Réticulé Chimiquement utilisé dans l'isolation des câbles de haute tension", Thèse de Doctorat es Sciences, ENP, 2007.
- [20] Y. Mechri, "Influence du Mode de Vieillessement Thermique sur les Propriétés du Polyéthylène Réticulé Chimiquement Utilisé dans l'isolation des Câbles Moyenne Tension", Thèse de Doctorat es Sciences, ENP, 2008.
- [21] S. H. Ling, Frank H. F. Leung, H. K. Lam, Yim-Shu Lee, and Peter K. S. Tam, "A Novel Genetic-Algorithm-Based Neural Networkfor Short-Term Load Forecasting",IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 50, NO. 4, AUGUST 2003.
- [22] S.H. Ling , H.K. Lam and F.H.F. Leung, "A variable-parameter neural network trained by improved genetic algorithm and its application", Proceedings of International Joint Conference on Neural Networks, Montreal, Canada, August, 2005.

- [23] S.H. Ling, F.H.F. Leung and H.K. Lam "A Variable Node-to-Node-Link Neural Network and Its Application to Hand-Written Recognition", International Joint Conference on Neural Networks, Vancouver, BC, Canada, 2006.
- [24] Frank H. F. Leung, K. Lam, S. H. Ling, and Peter K. S. Tam, "Tuning of the Structure and Parameters of a Neural Network Using an Improved Genetic Algorithm", IEEE Transactions On Neural Networks, Vol. 14, NO. 1, January 2003.
- [25] S.H. Ling, H.K. Lam, F.H.F. Leung, and Y.S. Lee, "A Genetic Algorithm Based Variable Structure Neural Network", IEEE Transactions On Neural Networks, 2003.
- [26] M. Zanaganeh, S. Jamshid Mousavi, A.F. Etemad Shahidi, "A hybrid genetic algorithm-adaptive network-based fuzzy inference system in prediction of wave parameters", ELSEVIER, Engineering Applications of Artificial Intelligence 22 Pages: 1194-1202, 2009.
- [27] Yoram Reicha and S.V. Baraib, "A methodology for building neural networks models from empirical engineering data", ELSEVIER, Engineering Applications of Artificial Intelligence 13, Pages: 685-694, 2000.
- [28] Alireza Khotanzad, Hassan Elragal and Tsun-Liang Lu "Combination of Artificial Neural-Network Forecasters for Prediction of Natural Gas Consumption", IEEE Transactions On Neural Networks, Vol. 11, NO. 2, March, 2000.
- [29] K.P.M. Madhugeeth, H.L. Premaratna "Forecasting Power Demand Using Artificial Neural Networks For Sri Lankan Electricity Power System", IEEE Region 10 Colloquium and the Third International Conference on Industrial and Information Systems, Kharagpur, India December, 2008.
- [30] Liu Hong, Cui Wenhua, Zhang Qingling, "Nonlinear Combination Forecasting Model and Application Based on Radial Basis Function Neural Networks", IITA International Conference on Control, Automation and Systems Engineering, 2009.
- [31] A. Azadeh and Behshtipour, "The effect of neural network parameters on the performance of neural network forecasting", IEEE International Conference on Industrial Informatics, Deajeon, Korea, 2008.
- [32] Li-Hua Feng, Jia Lu, "The practical research on flood forecasting based on artificial neural networks", ELSEVIER, Expert Systems with Applications 37, Pages: 2974-2977, 2010.
- [33] D. Bouguedad, "Influence du vieillissement thermique sur les propriétés de l'Ethylene Propylene Diene Monomere (EPDM)", Thèse de Doctorat d'état, UMM Tizi Ouzou, 2010.

ملخص

تدخل هذه الأطروحة ضمن موضوع البحث المتعلق باستخدام تقنيات الذكاء الاصطناعي بهدف التنبؤ، في إطار التغلب على نقائص الشبكات العصبية الاصطناعية، خصوصاً تلك المتعلقة باختيار الخصائص البنيوية (عدد الطبقات المخفية، و عدد العصبونات في كل طبقة)، إرتأينا إستعمال تقنية معروفة بامتازاتها في مجال التحسين، ألا و هي الخوارزميات الجينية، التقنية الهجينة المطبقة استعملت من أجل التنبؤ بالتغيرات غير الخطية لخصائص البوليتيلان المعرق كيميائياً (ب.م.ك) (PRC) والمستعمل في كوابل الجهد المتوسط.

كلمات مفاتيح: تنبؤ، الشبكات العصبية الاصطناعية، الخوارزميات الجينية، البوليتيلان المعرق كيميائياً (ب.م.ك) (PRC)

Résumé

Notre mémoire de magistère entre dans la thématique de recherche sur l'utilisation des techniques de l'intelligence artificielle (IA) dans la prédiction. Dans le but de surmonter les limites des réseaux de neurone artificiel (RNA), et plus précisément, celle reliée au choix des paramètres structurels (nombre de couches, et nombre de neurone par couche), nous avons fait appel à une technique connue par sa capacité d'optimisation, il s'agit des algorithmes génétiques (AG). La technique hybride est appliquée à la prédiction de la variation non linéaire des propriétés du PRC vieilli thermiquement utilisé dans les câbles MT.

Mots Clés: Prédiction, Réseaux de neurones artificiels, Algorithmes génétiques, Polyéthylène Réticulé chimiquement (PRC),

Abstract:

Our thesis comes into the thematic of research on the use of the artificial intelligence (AI) techniques for forecasting. In order to overcome the limits of the artificial neural networks (ANN), especially that related to the choice of the structural parameters (the number of hidden layers, and the number of neurones per layer), we have used an optimization specialized technique, it is the genetic algorithms (GA). The hybrid technique is used to forecast the non linear variation of the properties of the Cross-Linked Polyethylene (XLPE) used in medium voltage cables.

Key words: Forecasting, Artificial neural networks, genetic algorithms, Cross-Linked Polyethylene (XLPE)