

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

ECOLE NATIONALE POLYTECHNIQUE

DEPARTEMENT ELECTROTECHNIQUE

Mémoire de Projet de Fin d'Études
Pour l'obtention du Diplôme d'Ingénieur d'État en Électrotechnique

Thème

*Utilisation des réseaux de neurones artificiels pour la
caractérisation du comportement d'un isolateur pollué*

Proposé par

- Pr. A.MEKHALDI

Étudié par

- Mehdi MAADA

- Adel ZAITRI

Promotion JUIN 2011

ENSP 10, Avenue Hassen Badi, 16200 El Harrach, ALGER

Résumé:

Dans ce travail nous nous intéressons à la prédiction de la tension de contournement et du courant de fuite pour modèle plan d'un isolateur de haute tension soumis à une pollution uniforme en utilisant les réseaux de neurones artificiels. Nous avons utilisé les résultats des essais effectués en laboratoire comme base de données afin d'entraîner deux types de réseaux de neurones (RNA) suivant des stratégies d'apprentissage différentes, tout en essayant d'interpréter et de tirer des conclusions sur les résultats obtenus. Le premier RNA est le FeedForward Back-Propagation représentant les RNA à apprentissage supervisé. Le deuxième type de RNA est les Self Organized Maps (SOM) représentant les RNA à apprentissage non-supervisé.

Mots clés:

Réseaux de neurones artificiels, tension de contournement, courant de fuite, pollution uniforme, prédiction, Feed Forward-Back Propagation (FFBP), Organized Maps (SOM).

Abstract:

In this work we have been interested in predicting the flashover voltage and leakage current on the surface of a plane model of high voltage insulator exposed to uniform pollution using artificial neural networks. For the database, we have used the results of laboratory tests to train two types of neural networks (ANN) according to different learning strategies: the Feed Forward Back-Propagation (FFBP) as supervised networks and the Self Organized Maps (SOM) as unsupervised networks.

Key-words:

Artificial neural networks, flashover voltage, leakage current, uniform pollution, prediction, Feed Forward Back-Propagation (FFBP), Self Organized Maps (SOM).

ملخص:

تناول هذا العمل في الأساس تنبؤ توتر الإحاطة الكهربائية و شدة التيار المتسرب على سطح نموذج مسطح للعوازل المستعملة في شبكات الجهد العالي. هذا النموذج تم تعريضه لتلوث موحّد على كامل سطحه. للتنبؤ قمنا باستعمال الشبكات العصبية الاصطناعية. قاعدة المعلومات هي عبارة عن نتائج تجارب مخبرية تم استعمالها لتدريب الشبكات العصبية الموجهة بإستراتيجيتين إثننتين : طريقة Forward Back-Propagation (FFBP) و طريقة Self Organized Maps (SOM) .

كلمات مفتاحية :

الشبكات العصبية الاصطناعية , توتر الإحاطة الكهربائية, شدة التيار المتسرب, تلوث موحّد, تنبؤ , طريقة Forward Back-Propagation (FFBP) و طريقة Self Organized Maps (SOM) .

Dédicace

Voici ironiquement, la partie la plus délicate de ce présent rapport, car contrairement à ce qui va suivre, aucun logiciel n'est en mesure d'indiquer un message d'erreur pour signaler l'absence d'une personne ayant joué un important rôle, direct ou indirect dans l'élaboration de ce document qui vient mettre un terme à de nombreuses années d'étude.

Tout d'abord, et comme nous sommes algériens, on remerciera la mère, sans qui, rien de ce qui figure ici n'aurait pu être, ou en tous les cas, pas dans cette forme ordonnée, soignée et frisant la perfection...

Une autre pensée pour la famille, proche ou lointaine, frères et sœurs, oncles et tantes, et mêmes les cousins et cousines éloignés que l'on voit une fois par vie.

Une autre pensée bien sûr pour le père, qui a réussi, via quelques gifles, à inculquer les valeurs qui ont fait de nous ce que l'on est aujourd'hui.

Au cours d'une vie, on fait des rencontres, et grâce à Dieu, des gens extraordinaires ont croisé notre route, depuis le primaire jusqu'à ces derniers mois, des potes, qui ont réussi, l'espace d'un match de PES, de football, d'une partie de coinche, d'une soirée, à faire revivre le film « la vita è bella » de Roberto Begnini.

Parmi ces gens extra ordinaires, comment ne pas citer Nassim Salhi ? Un colloc d'exception, un binôme d'exception, et qui a su rester humble et modeste avec le temps. Son départ a laissé un vide, mais on n'a pas cessé de penser à lui, il était pour nous ce que la tour Eiffel était pour Paris ...

Un grand hommage à Ouaichouche nadia «dieu ait son âme», une mère, une sœur, une tante, une artiste et tant d'autres qualificatifs pour décrire cette personne d'exception qui a tout au long de sa vie encouragé et poussé ses proches à être des meilleurs personnes.

En espérant n'avoir oublié personne, et grâce à ces quelques lignes, on repart avec le sentiment du devoir accompli et la conscience tranquille.

Hier est le passé, aujourd'hui est un cadeau, c'est pour ça que ça s'appelle le présent ...

Remerciements

Nous tenons à remercier notre promoteur Mr. A. Mekhaldi pour sa constante présence au sein du laboratoire de Haute Tension ainsi que ses précieux conseils dans le domaine de la recherche scientifique.

Nous tenons également à remercier tous les enseignants du département du Génie Electrique ainsi que les enseignants de l'École Nationale Préparatoire aux Études d'Ingéniorat ayant contribué à notre formation et ce depuis le tronc commun jusqu'à la dernière année de graduation.

Un grand merci également à nos familles et amis respectifs qui ont partagé nos moments de joie et nous ont soutenus dans les moments difficiles.

Table des matières

Introduction générale.....	1
<i>CHAPITRE I: Introduction aux réseaux de neurones artificiels.</i>	
I.1. Introduction.....	3
I.2. Objectif.....	3
I.3. Historique.....	4
I.4. Applications :	5
I.5. Les neurones artificiels	6
I.5.1. Le neurone formel.....	6
I.5.2. Fonction discriminant	7
I.5.2.1. Fonction dynamique linéaire.....	7
I.5.2.2. Fonction non dynamique non linéaire.....	7
I.5.3. Structure d'interconnexion	8
I.5.3.1. Réseau multicouche	8
I.5.3.3. Réseau à connexions récurrentes	9
I.5.3.4. Réseau à connexions complexes	9
I.5.4. Apprentissage des réseaux de neurones	10
I.5.4.1 Procédure d'apprentissage	10
I.5.4.2 Types d'apprentissage.....	11
I.5.5. LES RESEAUX DE NEURONES FORMELS	11
I.5.5.1. Les réseaux de neurones non bouclés	12
I.5.5.2. Les réseaux de neurones bouclés	13
I.5.6. Mise en œuvre des réseaux de neurones et domaines d'applications.....	15
I.5.7. L'apprentissage des réseaux de neurones formels.....	17
I.5.8. Modélisation de phénomènes non linéaires statiques	18
I.5.9. Modélisation de processus non linéaires dynamiques.....	19
I.5.9.1. Les modèles dynamiques "boîtes noires"	19
I.5.9.2. Les modèles neuronaux de connaissances	19
I.5.10. Commande de processus.....	20

CHAPITRE II: *Application de la méthode des réseaux de neurones artificiels à la haute tension*

II.1. Les décharges partielles.....	23
II.2. Les huiles de transformateurs.....	24
II.3. Les isolateurs de haute tension.....	24
II.4. Autres.....	27

CHAPITRE III: *Présentation des méthodes utilisées*

III.1. Introduction	29
III.2. Méthode 1: Feed Forward-Back Propagation.....	30
III.2.1. La couche d'entrée:.....	30
III.2.2. La couche de sortie:.....	30
III.2.3. Les couches cachées:.....	30
III.2.4. Le nombre de couches cachées:	30
III.2.5. Le nombre de neurones dans les couches cachées:	31
III.2.6. Méthode de Levenberg-Marquardt (L.M)	32
III.3. Méthode 2: Les cartes auto-organisées de Kohonen (Self-Organization Maps)	34
III.3.1. Réseau de neurones SOM et règle d'apprentissage.....	35
III.3.2. Phase d'apprentissage du réseau de prédiction de Kohonen	37

CHAPITRE IV: *Présentation des résultats.*

IV.1. Introduction	40
IV.2. Méthode 1 : La Back-propagation.....	40
IV.2.1. Prédiction de la tension de contournement :	41
IV.2.2. Prédiction du courant de fuite :	51
IV.3. Méthode 2 : Réseau de neurones SOM :	54
IV.3.1. Prédiction de la tension de contournement :	54
IV.3.2. Prédiction du courant de fuite :	55
IV.4. Observations :	57
IV.5. Conclusion :	58
Conclusion générale	60

Introduction générale

Les enjeux économiques relatifs aux dysfonctionnements des réseaux électriques sont très importants et se comptent par millions de dollars. Pour cela les nouvelles politiques de régulation des réseaux électriques prennent les pannes accidentelles très au sérieux, en faisant des tests préventifs pour prédire les défauts et élaborer un programme de maintenance et de remplacement des équipements.

La dégradation des systèmes d'isolation peut diminuer la vie utile des équipements de puissance et provoquer des arrêts de systèmes coûteux. Pour prévenir de telles situations et pour aider les industries et les réseaux indépendants à assurer à coût concurrentiel la maintenance préventive de leurs équipements en charge, des systèmes de surveillance-diagnostic et maintenance doivent être installés.

La maintenance préventive et prévisionnelle inclut des tâches telles que la surveillance, le diagnostic, le contrôle et la prédiction. Pour la maintenance préventive du matériel électrique, la détection et le diagnostic des défauts sont réalisés à l'aide de l'interprétation de signatures complexes tandis que la prédiction des propriétés est réalisée à l'aide d'une assez large base de données de mesures.

Les isolateurs de haute tension forment un élément essentiel des systèmes de transmission de la puissance électrique. Le moindre défaut ou dysfonctionnement de ces installations entraîne d'énormes pertes de capitaux, puisque plusieurs industries ont besoin d'un approvisionnement constant et ininterrompu en énergie. La recherche travaille dans le but d'arriver à des méthodes et plans de prévention, afin d'éviter les situations d'urgences et faire de la maintenance plutôt que de la réparation qui coûte beaucoup plus chère.

Ce travail fait partie d'une large stratégie qui vise à faire le diagnostic du réseau de transport d'énergie algérien et cela en trouvant de nouvelles méthodes efficaces pour la caractérisation de ce réseau et de ces éléments. Dans ce travail en particulier, nous nous intéressons à la prédiction du comportement des chaînes d'isolateurs de haute tension soumises à une pollution uniforme. Pour cela nous nous basons sur les essais effectués au laboratoire de haute tension de l'École Nationale Polytechnique. Pour le traitement des résultats, nous avons utilisé deux stratégies de réseaux de neurones artificiels (RNA), la FFBP (Feed Forward Back Propagation) et les SOM (Self-Organization Maps).

Notre travail est réparti en quatre chapitres. Dans le premier chapitre nous introduisons les réseaux de neurones artificiels en donnant quelques principes de base qui aideront à comprendre la démarche que nous avons suivie. Le deuxième chapitre est consacré à l'histoire de l'application des RNA dans le domaine de la haute tension. Dans le troisième chapitre nous détaillerons les deux méthodes réseaux de neurones utilisées, en parlant de la procédure d'apprentissage, du choix des paramètres,... Dans le chapitre quatre nous exposerons les résultats obtenus après traitement par RNA. Pour terminer par une conclusion générale.

CHAPITRE I

Introduction aux réseaux de neurones artificiels

I.1. Introduction

Les réseaux de neurones, constitués par des structures cellulaires artificielles, constituent une approche permettant d'aborder sous des angles nouveaux les problèmes de perception, de mémoire, d'apprentissage et de raisonnement. Ils constituent aussi des alternatives très prometteuses pour contourner certaines limitations des ordinateurs classiques. Grâce à leur traitement parallèle de l'information et à leurs mécanismes inspirés des cellules nerveuses (neurones), ils infèrent des propriétés émergentes permettant de solutionner des problèmes jadis qualifiés de complexes.

Dans les années quarante, un certain nombre de chercheurs s'est intéressée à modéliser simplement le cerveau humain à partir de ses éléments essentiels, les neurones et les synapses. Les modèles de neurones et de structures de réseaux qui sont nés de ces recherches sont d'une simplicité extrême mais leur but n'était pas d'essayer de représenter fidèlement le modèle biologique (ce qui semble impossible). Il consiste plutôt à prendre en compte les caractéristiques principales les plus simplifiées possibles du modèle biologique et à observer les propriétés qui résultent de leur association. De même que dans le cas du cerveau, ces réseaux utilisent le principe de la mémoire distribuée stockant les informations dans l'ensemble des neurones et synapses. A partir d'éléments bâtis sur une biologie simplifiée à l'extrême, est-il possible pour ces réseaux de neurones formels d'apprendre à compter, de reconnaître un objet ou encore de résoudre un problème jamais étudié auparavant ? [1].

Dans ce chapitre nous donnerons l'historique d'apparition des réseaux de neurones artificiels. Nous y exposerons sommairement certaines architectures que l'on trouve dans la littérature. Il ne s'agit pas de les étudier toutes, car elles sont trop nombreuses, mais plutôt d'en comprendre les mécanismes internes fondamentaux et de savoir comment et quand les utiliser.

I.2. Objectif

Le cerveau humain contient environ 100 milliards de neurones. Ces neurones vous permettent d'exécuter plusieurs tâches à la fois sans même s'en rendre compte comme maintenir une respiration régulière permettant d'oxygéner le sang, en actionnant le cœur qui assure une circulation efficace de ce sang pour nourrir vos cellules, etc.

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 genre de théorie qui a inspiré la plupart des architectures de réseaux de neurones artificiels que nous aborderons dans ce cours. L'apprentissage consiste alors soit à établir de nouvelles connexions, soit à en modifier des 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? Eh bien, la réponse courte est «oui», même si les réseaux que nous allons développer ne possèdent qu'une infime fraction de la puissance du cerveau humain.

I.3. Historique

De nombreux ouvrages ont permis de documenter l'histoire des recherches en réseaux de neurones. En particulier, le livre intitulé «Neurocomputing: Foundations of Research» édité par John Anderson et Edward Rosenfeld est une compilation de 43 articles qui ont marqué le domaine sur le plan historique. Chacun d'entre eux est d'ailleurs précédé d'une introduction qui permet de situer l'article dans son contexte.

L'histoire de la théorie des réseaux de neurones artificiels est passée pas plusieurs étapes:

Les premiers pas:

- **1890** : W. James, célèbre psychologue américain introduit le concept de mémoire associative, et propose ce qui deviendra une loi de fonctionnement pour l'apprentissage sur les réseaux de neurones connue plus tard sous le nom de loi de Hebb.
- **1943** : J. Mc Culloch et W. Pitts laissent leurs noms à une modélisation du neurone biologique (Un neurone au comportement binaire). Ce sont les premiers à montrer que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes (tout au moins au niveau théorique).
- **1949** : D. Hebb, physiologiste américain explique le conditionnement chez l'animal par les propriétés des neurones eux-mêmes. Ainsi, un conditionnement de type pavlovien, tel que nourrir tous les jours à la même heure un chien, entraîne chez cet animal la sécrétion de salive à cette heure précise même en l'absence de nourriture. La loi de modification des propriétés des connexions entre neurones qu'il propose, explique en partie ce type de résultats expérimentaux [1].

Les premiers succès:

- **1957** : F. Rosenblatt développe le modèle du Perceptron. Il construit le premier neuro-ordinateur basé sur ce modèle et l'applique au domaine de la reconnaissance de formes. Notons qu'à cette époque les moyens à sa disposition sont limités et c'est une prouesse technologique de réussir à faire fonctionner correctement cette machine pendant quelques minutes.
- **1960** : B. Widrow, un automaticien, développe le modèle Adaline (Adaptative Linear Element). Dans sa structure, le modèle ressemble au Perceptron. Cependant la loi d'apprentissage est différente. Celle-ci est à l'origine de l'algorithme de rétro- propagation de gradient très utilisé aujourd'hui avec les Perceptrons multicouches. Les réseaux de type Adaline restent utilisés de nos jours pour certaines applications particulières. B. Widrow a créé dès cette époque une des premières firmes proposant neuro-ordinateurs et neuro-composants, la "Memistor Corporation".

- **1969** : M. Minsky et S. Papert publient un ouvrage qui met en exergue les limitations théoriques du perceptron. Limitations alors connues, notamment concernant l'impossibilité de traiter par ce modèle des problèmes non linéaires. Ils étendent ces limitations à tous les modèles de réseaux de neurones artificiels. Après, les recherches dans le domaine furent abandonnées par manque de financement et les chercheurs se tournèrent principalement vers les systèmes à bases de règles [1].

L'ombre:

- **1967-1982** : Toutes les recherches ne sont, bien sûr, pas interrompues. Elles se poursuivent, mais sous le couvert de divers domaines comme : le traitement adaptatif du signal, la reconnaissance de formes, la modélisation en neurobiologie, etc. De grands noms travaillent durant cette période tels: S.Grosberg, T.Kohonen,...

Le renouveau:

- **1982** : J. J. Hopfield est un physicien reconnu, à qui l'on doit le renouveau d'intérêt pour les réseaux de neurones artificiels [1].

La levée des limitations:

- **1983** : La Machine de Boltzmann est le premier modèle connu apte à traiter de manière satisfaisante les limitations recensées dans le cas du perceptron. Mais l'utilisation pratique s'avère difficile, la convergence de l'algorithme étant extrêmement lente (les temps de calcul sont considérables).

- **1985** : La rétro-propagation de gradient apparaît. C'est un algorithme d'apprentissage adapté aux réseaux de neurones multicouches (aussi appelés perceptrons multicouches). Sa découverte réalisée par trois groupes de chercheurs indépendants indique que "la chose était dans l'air". Dès cette découverte, nous avons la possibilité de réaliser une fonction non linéaire d'entrée/sortie sur un réseau en décomposant cette fonction en une suite d'étapes linéairement séparables. De nos jours, les réseaux multicouches et la rétro-propagation de gradient reste le modèle le plus étudié et le plus productif au niveau des applications [1].

A partir de là, les réseaux de neurones ont connu un essor continu. Ainsi à cet algorithme d'apprentissage, sont venus s'ajouter d'autres, qui lui ont apporté des améliorations. Les architectures de ces réseaux ont eux aussi connu une très grande évolution, donnant naissance à de nouveaux types de réseaux, parfois plus complexes, mais souvent plus adaptés à certains problèmes, tels les DBNN (*Data Based Neural Networks*), les RBF (*Radial Based Networks*), les réseaux Gaussiens, et les réseaux dynamiques [2].

I.4. Applications :

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

Aérospatial : pilotage automatique, simulation du vol...

- Automobile : système de guidage automatique,...
- Défense : guidage de missile, suivi de cible, reconnaissance du visage, radar, sonar, Lidar, traitement du signal, compression de données, suppression du bruit...
- □ Electronique : prédiction de la séquence d'un code, vision machine, synthétiseur vocal, modèle non linéaire,...
- Finance : Prévion du coût de la vie
- Secteur médical : Analyse EEC et ECG
- Télécommunications : Compression de données ...

I.5. Les neurones artificiels

I.5.1. Le neurone formel

Le premier passage entre les observations neurophysiologiques et anatomiques et la notion de neurone formel date de 1943 et a été présenté par Mc Culloch et Pitts. Ils ont proposé le modèle suivant : « un neurone formel fait une somme pondérée des potentiels d'actions qui lui parviennent (chacun de ces potentiels est une valeur numérique qui représente l'état du neurone qui l'a émis), puis s'active suivant la valeur de cette sommation pondérée. Si cette somme dépasse un certain seuil, le neurone est activé et transmet une réponse (sous forme d'un potentiel d'action) dont la valeur est celle de son activation. Si le neurone n'est pas activé, il ne transmet rien ». Par cette définition, le neurone biologique est transformé en une fonction (dite de transfert), les synapses reliant les neurones sont représentés par des poids (réels pondérant les entrées de la fonction) et les signaux électriques (ou potentiels d'action) deviennent des nombres [1].

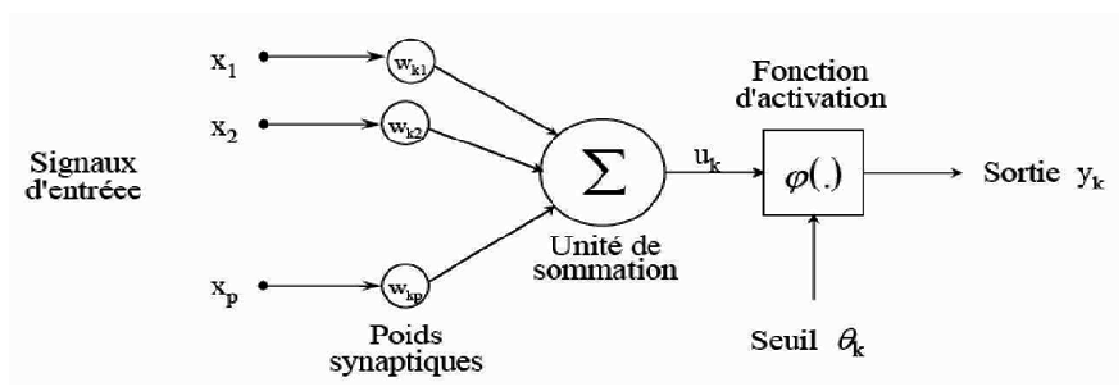


Figure I.1: Le modèle de Mc Culloch et Pitts [1].

I.5.2. Fonction discriminant

Cette fonction appelée aussi fonction de base (Basis Function), définit l'activité du neurone. Dans le cas du perceptron, c'est la fonction linéaire qui est utilisée. Cette fonction est depuis longtemps la plus usitée. C'est le cas pour les réseaux multicouches et d'une manière générale les réseaux **LBF** (Linear Basis Function).

Afin d'améliorer les performances des réseaux, d'autres formes de Fonction discriminant non linéaires ont été élaborées ces dernières années, offrant ainsi aux réseaux de meilleures possibilités dans certains domaines et un large champ d'application. Ces techniques nécessitent néanmoins des calculs plus compliqués pour leur élaboration.

I.5.2.1. Fonction dynamique linéaire

Cette fonction détermine la dynamique du réseau, elle est régie par l'équation différentielle suivante :

$$a_0 u_i(t) + a_1 \dot{u}_i(t) = x_i(t)$$

u_i est l'activation du $i^{\text{ème}}$ neurone et x_i est l'entrée du système dynamique décrit par l'équation précédente et a_i sont des paramètres.

I.5.2.2. Fonction non dynamique non linéaire

Elle a pour rôle de borner l'activité du neurone. Le premier modèle proposé est la fonction seuil qui est une fonction binaire, mais dont le problème majeur est sa non dérivabilité. D'autres fonctions qui sont dérivables, monotones et qui peuvent effectuer le seuillage ont été mises au point. Parmi ces fonctions, on retrouve la sigmoïde, la tangente hyperbolique ou la log-sigmoïde.

Une autre fonction qui est très utilisée malgré qu'elle ne soit pas monotone, est la fonction gaussienne dont le grand avantage est son comportement local.

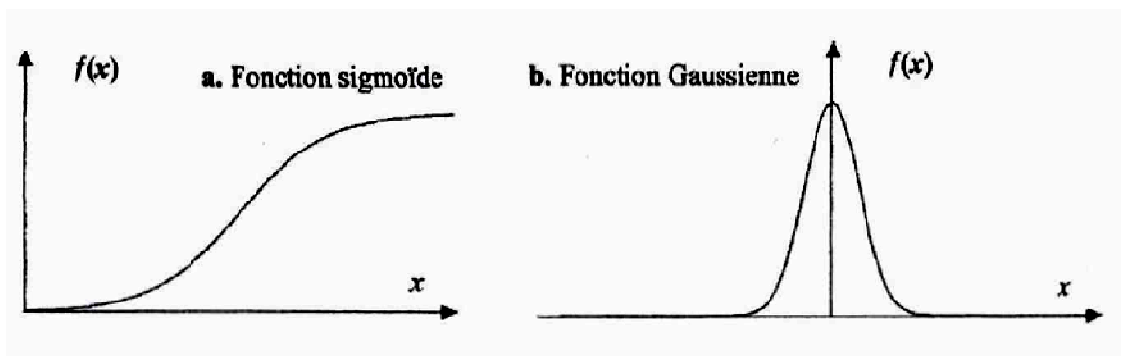


Figure I.2: Allure de quelques fonctions d'activation.

I.5.3. Structure d'interconnexion

Les connexions entre les neurones qui composent le réseau décrivent la « topologie » du modèle. Le plus souvent, cette topologie fait apparaître une certaine régularité de l'arrangement des neurones; cependant, celui-ci peut être quelconque.

I.5.3.1. Réseau multicouche

Les neurones sont arrangés par couche. On place ensuite bout à bout plusieurs couches et l'on connecte les neurones de deux couches adjacentes. Les entrées des neurones de la deuxième couche sont en fait les sorties des neurones de la couche amont. Les neurones de la première couche sont reliés au monde extérieur et reçoivent le vecteur d'entrée. Ils calculent alors leurs sorties qui sont transmises aux neurones de la seconde couche qui calculent eux aussi leurs sorties et ainsi de suite de couche en couche jusqu'à celle de sortie. Il peut y avoir une ou plusieurs sorties à un réseau de neurones.

Dans un réseau multicouche classique, il n'y a pas de connexion entre neurones d'une même couche et les connexions ne se font qu'avec les neurones de la couche aval. Tous les neurones de la couche amont sont connectés à tous les neurones de la couche aval.

Les couches extérieures du réseau sont appelées respectivement couches d'entrée et de sortie ; les couches intermédiaires sont appelées couches cachées [1].

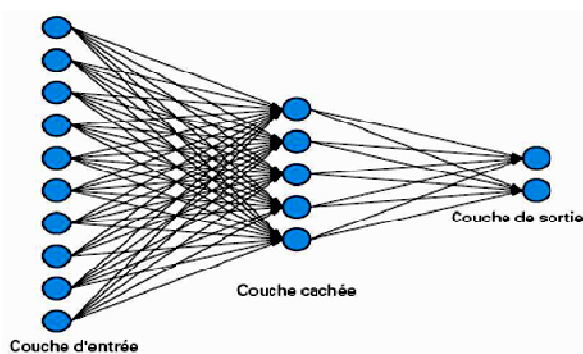


Figure I.3: Réseau multicouche classique.

C'est aussi un réseau multicouche, mais tous les neurones d'une couche amont ne sont pas connectés à tous les neurones de la couche aval. Nous avons donc dans ce type de réseau de neurones un nombre de connexions moins important que dans le cas du réseau de neurones multicouche classique.

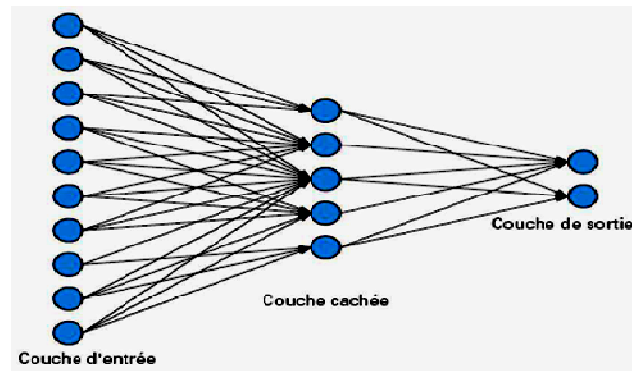


Figure I.4: Réseau multicouche à connexions locales.

I.5.3.3. Réseau à connexions récurrentes

Un réseau de ce type signifie qu'une ou plusieurs sorties de neurones d'une couche aval sont connectées aux entrées des neurones de la couche amont ou de la même couche. Ces connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche.

Les réseaux à connexions récurrentes sont des réseaux plus puissants car ils sont séquentiels plutôt que combinatoires comme l'étaient ceux décrits précédemment. La rétroaction de la sortie vers l'entrée permet à un réseau de ce type de présenter un comportement temporel.

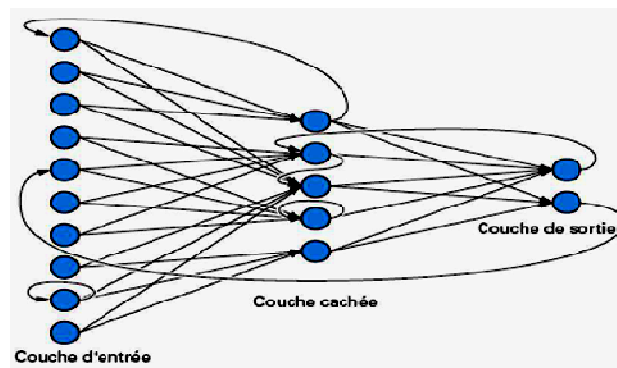


Figure I.5: Réseau multicouche à connexions récurrentes.

I.5.3.4. Réseau à connexions complexes

Chaque neurone est connecté à tous les neurones du réseau y compris lui-même, c'est la structure d'interconnexion la plus générale.

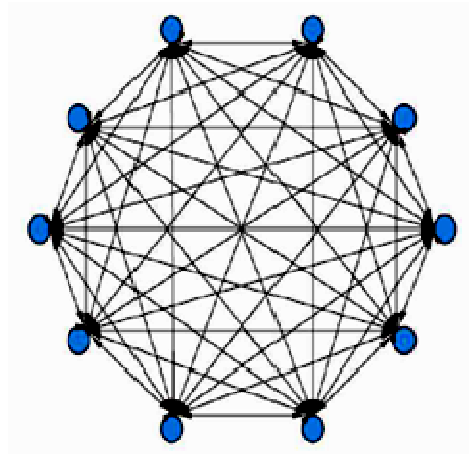


Figure I.6: Réseau à connexions complexes.

I.5.4. Apprentissage des réseaux de neurones

L'idée d'apprentissage dans les neurones biologiques a été avancée par Donald Hebb en 1949. Ce dernier propose l'existence d'un renforcement du lien entre deux neurones qui sont activés en même temps. L'apprentissage est sans doute la propriété la plus intéressante des réseaux neuronaux. On peut définir l'apprentissage d'un réseau de neurones comme un cycle d'évolution d'un réseau de neurones durant lequel le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. Il faut s'avoir aussi que l'apprentissage neuronal s'appuie sur des exemples de comportement.

Dans la plupart des algorithmes modernes, les variables modifiées durant l'apprentissage sont les poids des liens synaptiques. L'apprentissage est la modification des poids du réseau dans le but est d'accorder la réponse du réseau aux exemples et à l'expérience. A la fin de l'apprentissage, les poids sont établis: c'est alors la phase d'utilisation. Certains modèles de réseaux sont improprement désignés à apprentissage permanent. Dans ce cas il est vrai que l'apprentissage ne s'arrête jamais, cependant on peut toujours distinguer une phase d'apprentissage et une phase d'utilisation [1].

I.5.4.1 Procédure d'apprentissage

En général, une procédure d'apprentissage se compose de quatre étapes:

1. Initialisation de façon aléatoire des poids synaptiques,
2. Présentation des conditions et des données qui sert de base d'apprentissage,
3. Calcul de l'erreur commise et l'évaluation des performances,
4. Calcul de la correction à apporter au poids synaptique.

I.5.4.2 Types d'apprentissage

On peut distinguer deux types d'apprentissage:

a) Apprentissage supervisé:

Cet apprentissage se fait en introduisant des couples d'entrées et leur sorties désirées, On détermine par la suite les performances du réseau par l'intermédiaire d'un critère à optimiser.

b) Apprentissage non supervisé:

L'apprentissage non supervisé nécessite la présence des entrées seulement. Il est basé seulement sur l'information locale existant au niveau des neurones.

La figure 7 montre les différentes manières que nous avons de classifier les réseaux de neurones et cela en fonction de leurs architectures, technique d'apprentissage ou leurs fonctions.

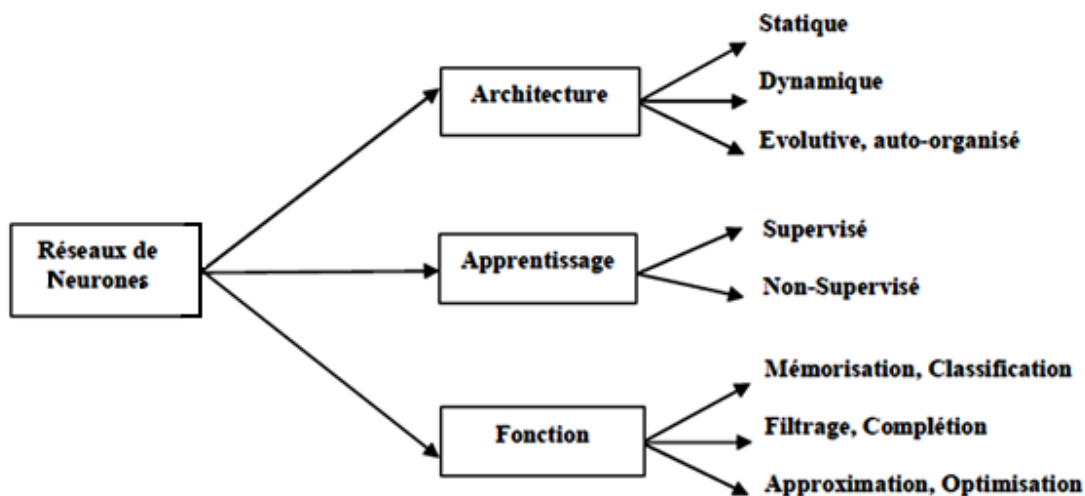


Figure I.7: Les différentes possibilités de classification des réseaux de neurones [1].

I.5.5. LES RESEAUX DE NEURONES FORMELS

On distingue deux grands types d'architectures de réseaux de neurones : les réseaux de neurones non bouclés et les réseaux de neurones bouclés.

I.5.5.1. Les réseaux de neurones non bouclés

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonctions algébriques de ses entrées, par composition des fonctions réalisées par chacun de ses neurones.

Un réseau de neurones non bouclé est représenté graphiquement par un ensemble de neurones "connectés" entre eux, l'information circulant des entrées vers les sorties sans "retour en arrière" ; si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les "connexions" entre ceux-ci, le graphe d'un réseau non bouclé est acyclique. Le terme de "connexions" est une métaphore : dans la très grande majorité des applications, les réseaux de neurones sont des formules algébriques dont les valeurs numériques sont calculées par des programmes d'ordinateurs, non des objets physiques (circuits électroniques spécialisés) ; néanmoins, le terme de connexion, issu des origines biologiques des réseaux de neurones, est passé dans l'usage, car il est commode quoique trompeur. Il a même donné naissance au terme de connexionnisme.

La Figure I.8 représente un réseau de neurones non bouclé qui a une structure particulière, très fréquemment utilisée : il comprend des entrées, une couche de neurones "cachés" et des neurones de sortie. Les neurones de la couche cachée ne sont pas connectés entre eux. Cette structure est appelée Perceptron multicouche.

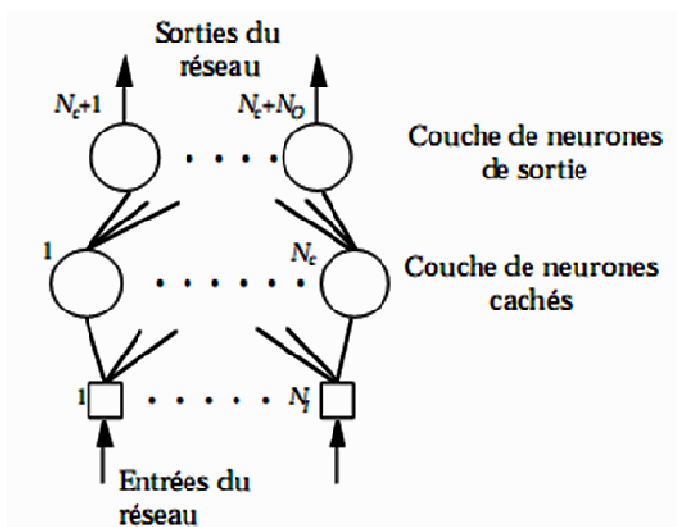


Figure I.8: Un Perceptron multicouches [2].

Les réseaux de neurones non bouclés sont des objets statiques : si les entrées sont indépendantes du temps, les sorties le sont également. Ils sont utilisés principalement pour effectuer des tâches d'approximation de fonction non linéaire, de classification ou de modélisation de processus statiques non linéaires.

I.5.5.2. Les réseaux de neurones bouclés

Contrairement aux réseaux de neurones non bouclés dont le graphe de connexions est acyclique, les réseaux de neurones bouclés peuvent avoir une topologie de connexions quelconque, comprenant notamment des boucles qui ramènent aux entrées la valeur d'une ou plusieurs sorties. Pour qu'un tel système soit causal, il faut évidemment qu'à toute boucle soit associé un retard : un réseau de neurones bouclé est donc un système dynamique, régi par des équations différentielles, comme l'immense majorité des applications sont réalisées par des programmes d'ordinateurs, on se place dans le cadre des systèmes à temps discret, où les équations différentielles sont remplacées par des équations aux différences.

Un réseau de neurones bouclé à temps discret est donc régi par une (ou plusieurs) équations aux différences non linéaires, résultant de la composition des fonctions réalisées par chacun des neurones et des retards associés à chacune des connexions.

La forme la plus générale des équations régissant un réseau de neurones bouclé est appelée forme canonique ;

$$\begin{aligned}x(k+1) &= \phi [x(k), u(k)] \\ y(k) &= \psi [x(k), u(k)]\end{aligned}$$

Où ϕ et ψ sont des fonctions non linéaires réalisées par un réseau de neurones non bouclé (par exemple, mais pas obligatoirement, un Perceptron multicouche), et où k désigne le temps (discret). La forme canonique est représentée sur la Figure 9. Tout réseau de neurones, aussi compliqué soit-il, peut être mis sous cette forme canonique, de manière complètement automatique. Ainsi, le réseau représenté sur la Figure 10 peut être mis sous la forme canonique, strictement équivalente mais beaucoup plus facilement manipulable, représentée sur la Figure I.11.

Les réseaux de neurones bouclés sont utilisés pour effectuer des tâches de modélisation de systèmes dynamiques, de commande de processus, ou de filtrage.

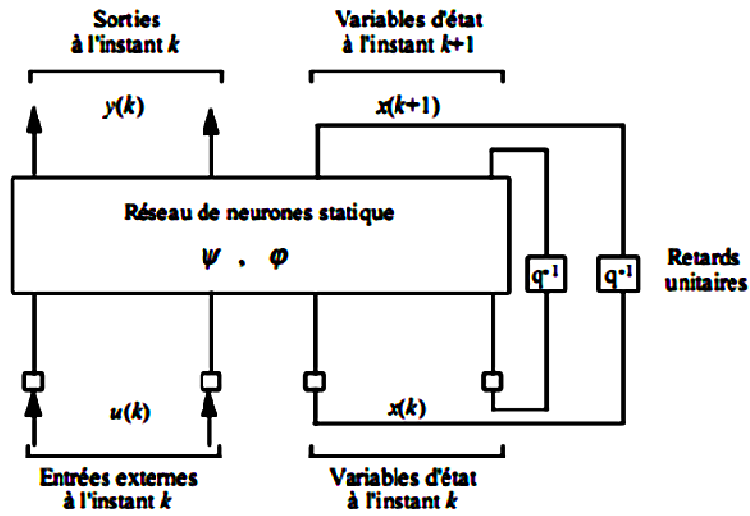


Figure I.9: Forme canonique d'un réseau de neurones bouclé [2].

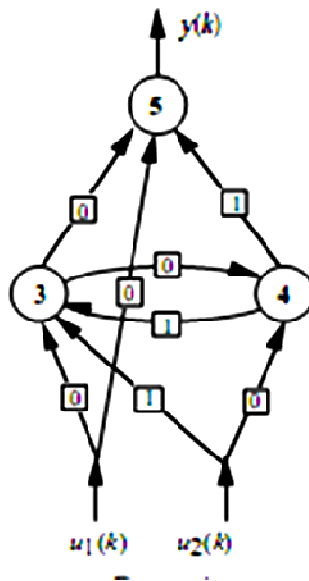


Figure I.10: Un réseau de neurones bouclé. Les nombres dans les carrés sont les retards (exprimés en nombre de périodes d'échantillonnage) associés à chaque connexion [2].

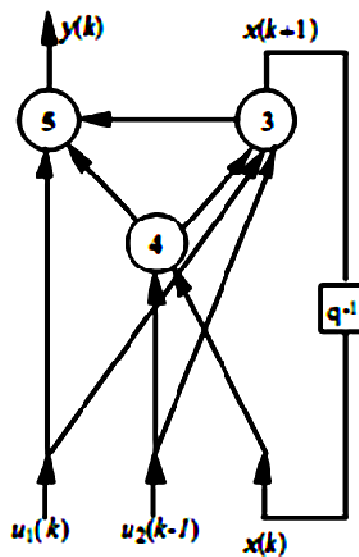


Figure I.11: La forme canonique du réseau de la figure I.10 [2].

I.5.6. Mise en œuvre des réseaux de neurones et domaines d'applications

Pour réaliser l'approximation de la fonction de régression cherchée, à partir d'échantillons généralement bruités, à l'aide d'un réseau de neurones, trois étapes successives sont nécessaires :

- il faut tout d'abord choisir l'architecture du réseau, c'est-à-dire les entrées externes, le nombre de neurones cachés, et l'agencement des neurones entre eux, de telle manière que le réseau soit en mesure de reproduire ce qui est déterministe dans les données ; le nombre de poids ajustables est un des facteurs fondamentaux de la réussite d'une application : si le réseau possède un trop grand nombre de poids, c'est-à-dire si le réseau est trop "souple", il risque de s'ajuster au bruit qui est présent dans les données de l'ensemble d'apprentissage, et, même en l'absence de bruit, il risque de présenter des oscillations non significatives entre les points d'apprentissage, donc de posséder de mauvaises propriétés d'interpolation (ou, dans le jargon des réseaux de neurones, de "généralisation") ; si ce nombre est trop petit, le réseau est trop "rigide" et ne peut reproduire la partie déterministe de la fonction. Le problème de la détermination de l'architecture optimale est resté pendant longtemps un problème ouvert, mais il existe actuellement diverses méthodes, mettant notamment en jeu des tests statistiques, qui permettent de déterminer cette architecture pour une vaste classe de réseaux; la Figure I.12 illustre l'influence du nombre de paramètres sur la qualité de l'approximation;

- il faut calculer les poids du réseau - ou, en d'autres termes, estimer les paramètres de la régression non linéaire - à partir des exemples, en minimisant l'erreur d'approximation sur les points de l'ensemble d'apprentissage, de telle manière que le réseau réalise la tâche désirée. Ce calcul des coefficients synaptiques constitue l'apprentissage supervisé pour le réseau de neurones;
- il faut enfin estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage.

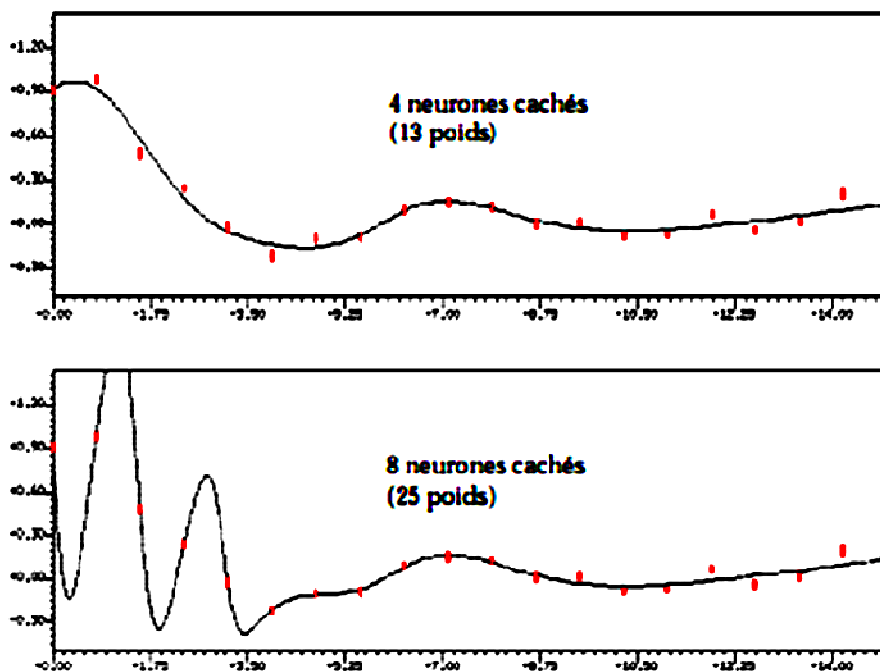


Figure I.12: Le "sur ajustement" : le réseau le plus parcimonieux (4 neurones cachés, soit 13 coefficients) produit un bien meilleur ajustement qu'un réseau trop riche en coefficients (8 neurones cachés, soit 25 coefficients) [2].

Les grands domaines d'application des réseaux de neurones découlent naturellement de leur propriété fondamentale :

La régression non linéaire, ou modélisation de données statiques :

Il existe une immense variété de phénomènes statiques qui peuvent être caractérisés par une relation déterministe entre des causes et des effets, les réseaux de neurones sont de bons candidats pour modéliser de telles relations à partir d'observations expérimentales, sous réserve que celles-ci soient suffisamment nombreuses et représentatives.

La modélisation de processus dynamiques non linéaires:

Modéliser un processus, c'est trouver un ensemble d'équations mathématiques qui décrivent le comportement dynamique du processus, c'est-à-dire l'évolution de ses sorties en fonction de celle de ses entrées ; c'est donc typiquement un problème qui peut être avantageusement résolu par un réseau de neurones, si le phénomène que l'on désire modéliser est non-linéaire. La prédiction de séries chronologiques (prédictions financières, prédiction de consommation, etc.) entre dans ce cadre.

La commande de processus :

Commander un processus, c'est imposer à celui-ci un comportement défini à l'avance en fonction des signaux de commande ; l'ensemble commande + processus peut donc être considéré comme un système qui réalise une fonction (non linéaire) qu'un réseau de neurones peut approcher.

La classification :

Supposons que l'on désire classer des formes en deux catégories, A ou B, en fonction de certaines caractéristiques de ces formes ; on peut définir une fonction ϕ qui vaut +1 pour toutes les formes de la classe A et -1 pour toutes les formes de la classe B. Les réseaux de neurones sont de bons candidats pour réaliser une approximation de cette fonction ϕ , et l'on peut démontrer que cette approximation constitue une estimation de la probabilité d'appartenance de la forme inconnue à la classe A. Les réseaux de neurones fournissent donc une information très riche, qui est loin d'être une simple réponse binaire. Cette propriété remarquable (que les réseaux de neurones partagent avec d'autres classifieurs) n'est malheureusement pas mise à profit dans la plupart des applications.

I.5.7. L'apprentissage des réseaux de neurones formels

Comme nous l'avons indiqué plus haut, l'apprentissage "supervisé", pour les réseaux de neurones formels, consiste à calculer les coefficients de telle manière que les sorties du réseau de neurones soient, pour les exemples utilisés lors de l'apprentissage, aussi proches que possibles des sorties "désirées", qui peuvent être la classe d'appartenance de la forme que l'on veut classer, la valeur de la fonction que l'on veut approcher ou de la sortie du processus que l'on veut modéliser, ou encore la sortie souhaitée du processus à commander. La plupart des algorithmes d'apprentissage des réseaux de neurones formels sont des algorithmes d'optimisation: ils cherchent à minimiser, par des méthodes d'optimisation non linéaire, une fonction de coût qui constitue une mesure de l'écart entre les réponses réelles du réseau et ses réponses désirées. Cette optimisation se fait de manière itérative, en modifiant les poids en fonction du gradient de la fonction de coût : le gradient est estimé par une méthode spécifique aux réseaux de neurones, dite méthode de rétro-propagation, puis il est utilisé par l'algorithme d'optimisation proprement dit. Les poids sont initialisés aléatoirement avant l'apprentissage,

puis modifiés itérativement jusqu'à obtention d'un compromis satisfaisant entre la précision de l'approximation sur l'ensemble d'apprentissage et la précision de l'approximation sur un ensemble de validation disjoint du précédent. Contrairement à des affirmations maintes fois répétées, l'apprentissage des réseaux de neurones n'est pas spécialement lent : il existe des algorithmes d'optimisation non linéaire extrêmement rapides qui permettent de faire des développements industriels sur de simples PC. L'apprentissage des réseaux de neurones bouclés (pour réaliser des modèles dynamiques) est très semblable à celui des réseaux non bouclés.

I.5.8. Modélisation de phénomènes non linéaires statiques

On peut trouver de nombreux exemples d'utilisation, potentielle ou opérationnelle, des réseaux de neurones pour la modélisation de phénomènes statiques. Ainsi, dans le domaine des relations structure-activité (ou QSAR), les réseaux de neurones permettent de prévoir l'activité d'une molécule (caractérisée par un nombre réel) à partir de descripteurs chimiques de cette molécule (qui sont eux-mêmes des nombres réels) ; par exemple, il est possible de prédire la solubilité dans l'eau, le point d'ébullition, le coefficient de partage eau-octanol, ou toute autre propriété caractérisée par un nombre, en fonction de descripteurs tels que la masse moléculaire, le moment dipolaire, la charge portée par les divers atomes, le "volume" de la molécule, etc. (Figure 13). Certains de ces descripteurs sont mesurables, d'autres sont calculés par des méthodes semi-empiriques ou *ab initio*. On peut imaginer de nombreuses extensions de cette approche : prédiction de propriétés pharmacologiques de molécules, formulation de mélanges, prédiction de propriétés mécaniques ou optiques de matériaux, etc. Par exemple, un système de prédiction des propriétés mécaniques des caoutchoucs pour pneumatiques a été développé pour la société Michelin. Dans ces domaines, la combinaison des connaissances physiques ou chimiques de l'expert et la maîtrise des techniques statistiques "neuronales" permet d'obtenir des résultats très supérieurs à ceux qui étaient obtenus auparavant en utilisant les mêmes bases de données.

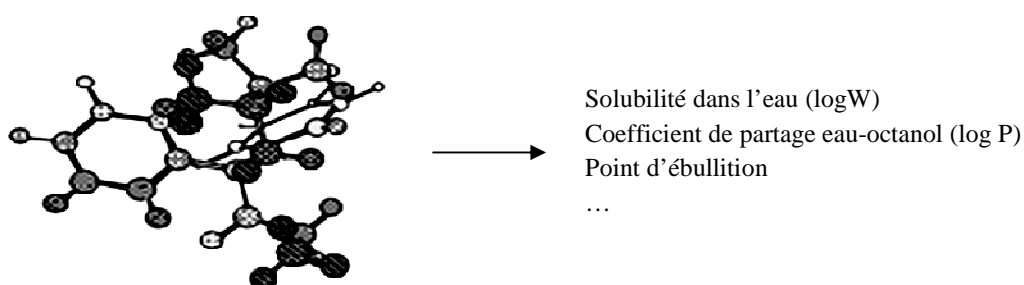


Figure I.13: Relation structure-activité (Quantitative Structure-Activity Relations ou QSAR) [2].

I.5.9. Modélisation de processus non linéaires dynamiques

Comme nous l'avons vu plus haut, la propriété d'approximation universelle parcimonieuse des réseaux de neurones peut avantageusement être mise à profit pour la modélisation dynamique non linéaire de processus très divers. On distingue habituellement

- les modèles de connaissance, dont l'expression mathématique, comprenant un petit nombre de paramètres ajustables, résulte d'une analyse (physique, chimique, économique, etc.) du processus,
- la modèle «boîte noire», qui est établis uniquement à partir des mesures effectuées sur le processus, sans intervention d'autre connaissance.

Les réseaux de neurones sont souvent utilisés comme des modèles "boîtes noires". Néanmoins, nous verrons plus loin qu'ils peuvent aussi mettre en œuvre des connaissances, constituant ainsi un excellent compromis entre les modèles "boîtes noires" et les modèles de connaissances.

I.5.9.1. Les modèles dynamiques "boîtes noires"

Pour réaliser un modèle dynamique, il convient alors d'utiliser des réseaux bouclés, qui, comme indiqué plus haut, sont eux-mêmes des systèmes dynamiques. Comme précédemment, l'apprentissage est l'estimation des paramètres du modèle neuronal utilisé, selon le schéma représenté sur la Figure I.14.

L'objectif de l'apprentissage n'est pas ici d'annuler l'erreur de prédiction, puisque, si tel était le cas, le réseau serait capable de reproduire l'effet des perturbations non mesurables ! Il s'agit plutôt d'obtenir une erreur de prédiction dont la variance est minimale, c'est-à-dire égale à celle du bruit. Si l'on peut obtenir un tel résultat, le réseau reproduit complètement le comportement déterministe du processus, bien que l'apprentissage ait été effectué en présence de perturbations. Des résultats théoriques prouvent que cet objectif est accessible, et de nombreux exemples montrent qu'il est effectivement atteint.

I.5.9.2. Les modèles neuronaux de connaissances

Une des caractéristiques particulièrement avantageuses - et généralement ignorée - des réseaux de neurones est la possibilité d'introduire, dans la conception même du réseau, des connaissances mathématiques résultant de l'analyse de la physico-chimie du processus. De telles connaissances sont très souvent disponibles, et il serait déraisonnable de ne pas en tirer profit. Ainsi, on bénéficie à la fois de l'intelligibilité des modèles de connaissance et des capacités d'adaptation des réseaux de neurones : loin de constituer des "boîtes noires", ces derniers sont alors de véritables modèles neuronaux de connaissances. Citons, à titre d'exemple,

un modèle de colonne à distiller réalisé par NETRAL S.A. pour Elf-Atochem, qui permet de détecter précocement des anomalies de fonctionnement, par examen de la différence entre les prédictions du modèle (qui fonctionne en temps réel sur PC) et les grandeurs effectivement mesurées sur la colonne.

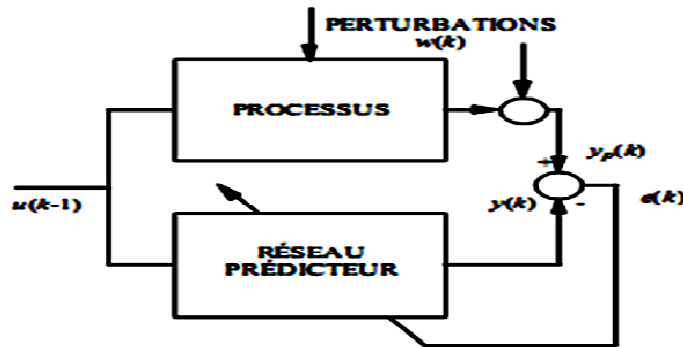


Figure I.14: Identification des paramètres d'un modèle : le réseau de neurones doit, à partir des commandes à l'instant $k-1$, prédire la sortie qu'aurait le processus à l'instant k si les perturbations $w(k)$, non mesurables, n'existaient pas [2].

I.5.10. Commande de processus

L'utilisation des réseaux de neurones pour la commande (adaptative ou non) de processus non linéaires découle naturellement des aptitudes de ces derniers à la modélisation. Il s'agit essentiellement d'une extension non linéaire de la commande optimale avec coût quadratique sur un horizon infini. Considérons en effet une structure de commande typique (la commande avec modèle interne) représentée sur la Figure I.15.

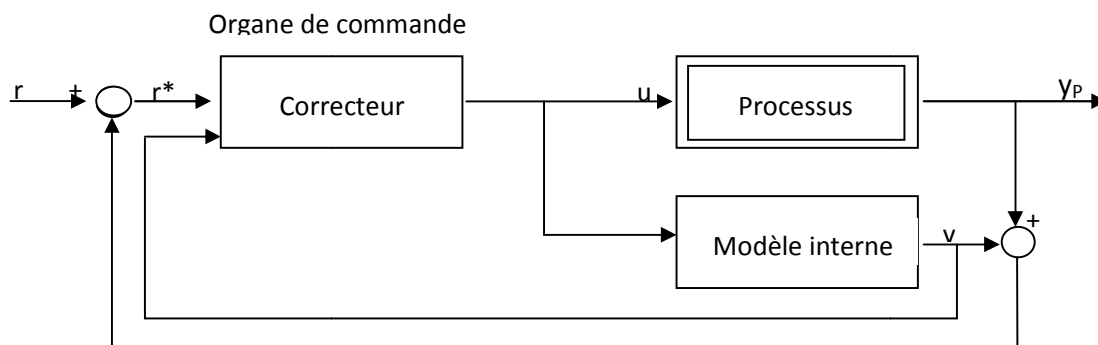


Figure I.15: Exemple d'architecture pour la commande neuronale robuste de processus non linéaires : la commande avec modèle interne [2].

Elle comprend :

- Un modèle neuronal, obtenu comme indiqué au paragraphe précédent
- un correcteur neuronal dont les coefficients sont mis à jour périodiquement si la commande est adaptative ; dans le cas contraire, ses coefficients sont fixes une fois l'apprentissage terminé.

Pour l'apprentissage de systèmes de poursuite, il est nécessaire, de surcroît, d'utiliser un modèle de référence qui traduit le cahier des charges en termes de dynamique de poursuite désirée. La commande de processus non linéaires semble être l'un des domaines les plus prometteurs pour les réseaux de neurones à l'heure actuelle. Ainsi, la société SAGEM, en collaboration avec l'ESPCI, a réalisé la commande d'un véhicule 4x4 Mercedes : le volant, l'accélérateur et le frein sont commandés par des réseaux de neurones, de telle manière que le véhicule reste sur une trajectoire déterminée, avec un profil de vitesse déterminé, en tout-terrain, indépendamment des perturbations possibles (glissement, terrain en dévers, vent, ...) ; le véhicule est équipé pour cela de capteurs proprioceptifs (permettant de connaître l'angle du volant, la pression dans le circuit de freinage, l'angle du papillon d'admission, la vitesse de rotation des roues) et extéroceptifs (permettant de connaître la position du véhicule dans son environnement) . Les comparaisons entre commandes "neuronales" (faisant intervenir des connaissances a priori sur le processus) et commandes non linéaires traditionnelles ont montré que les réseaux de neurones permettent d'obtenir des résultats au moins aussi bons, et souvent meilleurs, mais surtout qu'ils sont de mise en œuvre beaucoup plus simple en raison du caractère générique des algorithmes mis en œuvre : quelle que soit l'architecture du réseau bouclé utilisé, c'est toujours le même algorithme qui est mis en œuvre.

CHAPITRE II

*Application des réseaux de
neurones artificiels au domaine
de la haute tension*

Vu leurs performances, les méthodes d'intelligence artificielle sont de plus en plus utilisées pour traiter des sujets de recherches avancés dans tous les domaines en particulier le domaine de la haute tension. Dans ce chapitre nous présentons les résultats de certains travaux de recherche utilisant les réseaux de neurones artificiels dans le domaine de la haute tension, répartis en plusieurs sections selon le sujet de recherche.

II.1. Les décharges partielles

On doit les deux premières publications faites dans le domaine aux deux chercheurs A.A.Mazroua et al. [3,4]. Elles traitent de la reconnaissance des tailles d'impulsions de décharges partielles produites par différentes sources de décharges en utilisant un réseau compétitif: classificateur du plus proche voisin, un réseau LVQ (Learning Vector Quantization) et un réseau multicouches dont l'algorithme d'apprentissage est BP (Back-Propagation). Les trois réseaux ont donné les mêmes résultats sauf dans le cas des petites cavités. Le LVQ s'est montré plus performant pour reconnaître les décharges partielles produites par une seule cavité et celles produites par une arborescence électrique.

En 1995 K.Engel et D.Peier [5] ont utilisé les réseaux de neurones pour étudier l'influence du développement des décharges partielles sur l'identification du type de défaut. Les données mesurées et leur classification sont déterminées par le changement d'activité des décharges partielles. Ils ont effectué la classification des décharges partielles en utilisant la carte de Kohonen.

T.Kexioug et Z.Deheng [6] ont travaillé sur la reconnaissance des modèles de décharges partielles tridimensionnels. Ils ont étudié l'influence du nombre de neurones des couches d'entrées et cachées afin de choisir la meilleure structure du réseau. La capacité des réseaux de neurones à reconnaître les décharges partielles est testée par une comparaison entre les résultats de simulation et ceux de l'expérience.

S.Happe et H.G.Kranz [7] ont développé en 1997 un système de suppression en temps réel des signaux d'impulsions des décharges partielles en appliquant les réseaux de neurones. Les taux de répétition stochastiques et les pulsations sonores sont très nuisibles et entraînent une mauvaise interprétation des activités des décharges partielles. Ce problème affecte tous les systèmes de mesure des décharges partielles et sa suppression conduit à leur évaluation. Le réseau de neurones de Kohonen classe les signaux inconnus. Cette classification est utilisée pour l'apprentissage du réseau. Pour son adaptation ils utilisent l'algorithme de back-propagation (BP).

II.2. Les huiles de transformateurs

En 1995, T. Nogami [8] a publié dans une revue japonaise un travail sur la pré-détection d'anomalies dans l'huile d'un transformateur en utilisant des réseaux de neurones. Aucun détail n'a été donné concernant le réseau de neurones utilisé. Les résultats obtenus sont comparés à ceux des mesures chromatographiques. Ils donnent les types de défauts pouvant apparaître dans l'huile en fonction des gaz dissous captés et leurs concentrations.

En 1997, Z.Wang et P.Giffin [9] ont publié leur travail sur la combinaison d'un réseau de neurones et un système expert pour le diagnostic de défauts d'un transformateur en utilisant la méthode de l'analyse des gaz dissous.

En 2002, L. Mokhnache [10] a utilisé un réseau BP entraîné par l'algorithme L.M pour le diagnostic des huiles de transformateur utilisée par SONELGAZ en se basant sur les tests physico-chimiques.

En 2003, L. Mokhnache [11] a présenté un travail sur l'application des réseaux de neurones de Kohonen pour prédire le vieillissement thermique du PVC et de la BORAK22. Dans ce travail elle a essayé de rendre la prédiction indépendante de la nature de la caractéristique d'isolation et cela en transformant la base de données en retranchant la valeur moyenne de la valeur réelle et en divisant le résultat par la variance des données et de cette façon on utilise les profils au lieu des valeurs réelles. Ce même réseau a été utilisé pour la prédiction de la tension disruptive des systèmes pointe-barrière-plan en fonction de la longueur de l'intervalle et la largeur de la barrière.

II.3. Les isolateurs de haute tension

En 1995, P.S.Ghosh et al. [12] ont modélisé un arc électrique d'un isolateur pollué en utilisant un réseau de neurones artificiel entraîné par l'algorithme d'apprentissage *Resilient Propagation Algorithm* (RPROP). Dans le RPROP, l'adaptation des poids ou leur modification dépend des signes des dérivées de la fonction erreur E et non pas de leur amplitudes. Elle est basée sur ce qu'on appelle la règle d'apprentissage de Manhattan.

A.Ailam [13] a utilisé les RBF pour prédire les caractéristiques non-linéaires d'un isolateur pollué.

Zafer Aydogmus [14] a entrepris un travail sur l'estimation du champ électrique bidimensionnel le long d'une chaîne d'isolateurs de haute tension à profil antibrouillard en utilisant les réseaux de neurones artificiels. Deux méthodes sont soumises pour la validation du modèle. Premièrement, le champ électrique a été calculé pour différents niveaux de tension et de conductivité de la pollution par la méthode des éléments finis. Certaines données ont été

utilisées pour entraîner le réseau de neurones, les autres ont été utilisées pour le teste. Les coordonnées x et y représentant les coordonnées des nœuds composant le profil de l'isolateur discrétisé, ainsi que la tension appliquée sont les entrées du premier réseau pour estimer le champ électrique sur la surface de l'isolateur pour la première application. Puis les coordonnées x , y et le niveau de pollution de la surface ont été prises comme entrées du deuxième réseau toujours pour l'estimation du champ électrique sur la surface de l'isolateur pour la deuxième application. Ces modèles développés rendent la détermination du champ électrique plus aisée et diminuent le temps de calcul, aussi ces applications montrent une grande précision pour les deux modèles. Les réseaux de neurones présentés peuvent être utilisés pour concevoir et développer des isolateurs pour différentes tensions de service et niveaux de pollution.

Un autre travail fait par A.S.Farag [15] sur l'estimation du temps de claquage des isolateurs pollués par les réseaux de neurones artificiels. Il utilise des données expérimentales faites sur des modèles d'isolateurs pollués sous tension représentant plusieurs configurations de pollution allant de la plus faible à la plus sévère pour l'entraînement de ses applications. En faisant des études et testes plus approfondis, il a pu trouver les paramètres du réseau de neurones qu'il a utilisé afin de donner les meilleurs résultats possible et juger l'effet de la présence de données inadéquates dans l'apprentissage. La nouvelle approche utilisant les réseaux de neurones comme fonction d'estimation est employée pour modéliser précisément la relation $t = f(V, L, R_p)$ sachant que V représente la tension, L la longueur totale de l'isolateur et R_p la résistance par unité de longueur. Une fois l'apprentissage terminé, le réseau de neurones est capable d'estimer le temps de claquage de façon très efficace même si les données inadéquates sont incorporées dans l'apprentissage.

M.T.Gençoğlu et M.Cebeci [16] ont présenté un travail sur le claquage des isolateurs de haute tension pollués et les mécanismes physiques qui gouvernent ce phénomène en utilisant les réseaux de neurones artificiels. Dans cette étude ils ont modélisé la fonction $V_c = f(H, D, L, \sigma, n, d)$. H étant la hauteur, D le diamètre, L la longueur, σ la conductivité surfacique, n le nombre d'éléments de la chaîne, d le nombre de rainures. Pour cela ils ont utilisé le réseau de neurones Feed-Forward avec back-propagation comme algorithme d'apprentissage. Ce modèle a permis la prédiction de la tension de claquage de différents types de chaînes d'isolateurs avec grand succès en donnant des résultats très satisfaisants.

Un travail sur la prédiction du courant de fuite dans les isolateurs non céramiques sous brouillard salin par réseaux de neurones artificiels a été effectué par Ayman H.EL-Hag, Ali Naderian Jahromi et Madjid Sanaye-Pasand, [17]. Une cinquantaine d'isolateurs en caoutchouc de silicone avec trois classes de tension ont été testés dans une chambre de brouillard salin où le courant de fuite est constamment enregistré pendant 100h au minimum. Une limite pour la période de vieillissement prématuré est définie par le taux de changement du courant de fuite plutôt qu'une valeur initiale fixée. Le réseau Gaussien à base radiale est utilisé pour prédire le niveau du courant de fuite à la première étape du vieillissement des isolateurs en caoutchouc de

silicone et est comparé à un réseau classique. Les valeurs initiales du courant de fuite et ses taux de changement sur 10 minutes d'intervalle pendant les 5 premières heures sont utilisées comme entrées du RN. La valeur finale du courant de fuite de la période de vieillissement est considérée comme la sortie du réseau. Il est trouvé que le réseau Gaussien à base radiale, peu importe la méthode d'apprentissage, donne de très bons résultats avec une précision de 3.5-5.3%, si les données d'apprentissage et de teste sont prises du même type d'isolateur.

Ahmed S.Ahmad et al. [18] ont travaillé sur l'estimation de la densité équivalente des dépôts de sel sur les isolateurs de haute tension par les réseaux de neurones artificiels. Dans cette étude, ils ont étudié la contamination des isolateurs sous pollution marine. Une nouvelle approche se basant sur les RNA comme fonction d'estimation a été développée et utilisée pour modéliser la relation entre la densité équivalente des dépôts salins (ESDD) et la température, l'humidité, la pression, la précipitation, et la vitesse du vent. Les résultats obtenus ont été comparés aux valeurs mesurées expérimentalement. Ce travail a permis de trouver une méthode fiable d'estimation de la fonction non linéaire $ESDD=f(T,H,Pr,P,VV)$, ceci va aider à améliorer la maintenance et trouver une solution efficace contre le claquage des isolateurs de haute tension pollués.

S.M.Gubanski et al. [19] ont publié un travail sur le diagnostic des isolateurs en polymère en plein air. Ils ont exposé les principaux problèmes et défauts enregistrés sur les isolateurs en polymère. Une étude rétrospective des travaux effectués pour développer des méthodes fiables et robustes de diagnostic des isolateurs a été également réalisée.

En 2006, V.T.Kontargyri, A.A.Gialketsi, et al. [20] se sont intéressés à l'estimation du claquage des isolateurs par les RNA. Ils ont appliqué les RNA pour estimer la tension critique de claquage des isolateurs pollués. Les entrées du RNA sont le diamètre, la hauteur, et l'ESDD pour donner en sortie la tension de claquage. La base de données expérimentale et les résultats des modèles mathématiques ont été utilisés pour l'apprentissage et le teste des performances du RNA. L'avantage fondamental de cette méthode est l'agilité de trouver le choix optimal des paramètres comme le taux d'apprentissage, le nombre de neurones... et on a remarqué que plus la base de données expérimentales est grande plus l'estimation est précise que lorsqu'on utilise les modèles mathématiques.

En 2010, C.Volat, et al. [21] ont travaillé sur le contrôle du courant de fuite dans les stations d'isolateurs recouvertes de glace en période de fonte par les réseaux de neurones artificiels. Les testes, effectués sous un régime de contamination par la glace sous différentes conditions expérimentales, ont montré que l'établissement permanent des arcs blancs connu comme régime permanent mènent au claquage dans la majorité des cas. En se basant sur ces observations, le développement d'une méthode de contrôle qui a pour objectif la prévention de l'approche de l'arc blanc principal pendant la période de fonte est proposé. La méthode de contrôle utilise différents RNA pour prédire l'apparition de l'arc blanc basée sur l'identification, la classification et l'analyse de la fréquence d'apparition des décharges électriques. Les résultats

montrent que le modèle de contrôle par RNA développé est capable de prédire le début du régime permanent sous différentes conditions expérimentales. D'où il a été trouvé que le délai entre la prédiction du début du régime permanent donnée par le RNA et sa réalisation était égale approximativement à 9 minutes. Ce qui prouve que le modèle présenté peut être utilisé dans le système de contrôle pour les isolateurs durant les périodes de grand froid pour la protection contre les risques de claquage.

II.4. Autres

En 1999, S.Chowdhary [22] s'est intéressé à la détermination de la tension de claquage du SF₆ sous un champ non-uniforme. Le réseau de neurones a été entraîné par le PPLN (Project Pursuit Learning Network). C'est un genre de réseau qui se construit au fur et à mesure de l'apprentissage. Il commence par un neurone caché. En ajoutant un nouveau neurone, l'apprentissage se fait uniquement pour le nouveau neurone. Ce réseau a de bonnes capacités de généralisation.

Les réseaux de neurones ont aussi été utilisés pour la détermination de l'intensité du courant électrique et de la valeur critique de la tension disruptive dans les sous-stations, en tenant compte des facteurs atmosphériques et géométriques. L'objectif principal de cette étude est de trouver une méthode perspective pour le dimensionnement des sous-stations de 138 kV. Un réseau composé de deux réseaux de neurones en cascade est utilisé pour faire ce travail. Le premier calcule la tension disruptive qui devient l'une des entrées du second réseau. Ce dernier calcule l'intensité du champ électrique entre les jeux de barres [23].

L. Mokhnache [24] a présenté un travail traitant de la prédiction de l'effet du vieillissement thermique du PVC sur ses caractéristiques mécaniques en utilisant des réseaux de neurones RBFG (Radial Basis Function Gaussian) entraînés par la ROM (Random Optimisation Method). Ce travail de recherche a été effectué sous différentes températures de vieillissement. La durée de vie du PVC a été prédite aussi en fonction du temps de vieillissement.

En 2001, L. Mokhnache s'est intéressée à l'application des réseaux de neurones pour prédire la tension de claquage des intervalles d'air pointe-barrière-plan avec différentes positions de la barrière et différentes tensions appliquées [25].

CHAPITRE III

Présentation des méthodes utilisées

III.1. Introduction

Le but de notre travail est d'arriver à prédire certaines caractéristiques comme la tension de contournement, le courant de fuite en fonction des paramètres spécifiques du modèle et ce pour le diagnostic des isolateurs de haute tension en utilisant les réseaux de neurones artificiels.

Dans notre travail, nous nous basons sur les testes effectués au laboratoire de Haute Tension de l'ENP sur le modèle plan de l'isolateur (IEC 60 305/1995) dans le cas de la pollution uniforme. Ces essais nous ont permis de décrire l'évolution de la tension de contournement et du courant de fuite sur la surface du modèle plan pour différentes conductivités et différents niveaux de tension appliqués. Les solutions polluantes sont préparées pour plusieurs concentrations en Kaolin (40 g/l, 100 g/l et 200 g/l). Les conductivités considérées pour ces testes sont 1.2 mS/cm 3.5 mS/cm 6 mS/cm 10 mS/cm 20 mS/cm. Ces valeurs sont choisies pour avoir une bonne visibilité concernant la tension de contournement et du courant de fuite pour le cas de la pollution uniforme.

Dans ce chapitre, nous exposons les méthodes et algorithmes utilisés pour le traitement des résultats expérimentaux. Ce travail nous permet de faire la prédiction de la tension de contournement et du courant de fuite en fonction respectivement de la conductivité de la surface polluée et de la tension appliquée pour cela nous utilisons deux types de réseaux de neurones. Premièrement, la Back Propagation (BP) qui entraîne les réseaux multicouches dans un cadre supervisé, les cartes auto-organisées de Kohonen (Self-Organization Maps) qui rentrent dans la famille des réseaux de neurones à apprentissage non-supervisé.

Le fait que nous soyons autodidactes dans le domaine de l'intelligence artificielle et plus particulièrement dans le domaine des réseaux de neurones, le choix de ces deux méthodes c'est fait sur la base de la documentation qui était à notre disposition. Au cours de notre recherche bibliographique, nous avons trouvé que l'architecture Feed Forward Back Propagation était l'architecture la plus commune et la plus simple a utilisé [26]. Pour la deuxième architecture SOM. Nous nous sommes basé sur les différents travaux de recherche dans le domaine de la haute tension liés à la prédiction des caractéristiques des équipements de haute tension.

III.2. Méthode 1: Feed Forward-Back Propagation

La technique d'apprentissage back propagation est utilisée pour entraîner les réseaux de neurones d'architecture de type Feed Forward. C'est une architecture multicouche ce qui veut dire que la couche d'entrée est connectée à une couche cachée qui peut à son tour être connectée à une autre couche cachée ou directement raccordée à la couche de sortie. En général on ne trouve qu'une seule couche cachée dans ce genre d'architecture. Il est très rare de trouver plus de deux couches, mathématiquement cela ne change rien au fonctionnement du réseau [26].

III.2.1. La couche d'entrée:

La couche d'entrée du RNA est la conduite qui permet au monde extérieur de présenter un modèle au RNA. Une fois ce modèle posé à l'entrée du réseau la couche de sortie va produire un autre modèle. La couche d'entrée doit représenter les conditions pour laquelle le réseau est entraîné. Chaque neurone d'entrée doit représenter une variable indépendante qui a une influence sur la sortie du réseau.

III.2.2. La couche de sortie:

Cette couche est celle qui présente le modèle à l'environnement extérieur, tout résultat donné par la couche de sortie peut être tracé jusqu'à la couche d'entrée. Le nombre de neurones de la couche de sortie dépend directement de la fonction du RNA. Par exemple, s'il s'agit de classification en groupes, il est préférable d'avoir un seul neurone pour chaque groupe.

III.2.3. Les couches cachées:

Il y a deux décisions à prendre concernant les couches cachées. La première concerne le nombre de couches cachées à utiliser et la deuxième concerne le nombre de neurones dans chaque couche cachée.

III.2.4. Le nombre de couches cachées:

Les réseaux de neurones à deux couches cachées représentent des fonctions avec n'importe quelle forme. Il n'y a aucune utilité théorique à l'utilisation de réseaux à plus de deux couches cachées. Plusieurs problèmes pratiques n'ont pas besoin de plus d'une couche cachée. Le tableau III.1 montre la différence entre les nombres de couches cachées.

Nombre de couches cachées	Résultat
0	Capable de représenter uniquement les fonctions ou décisions linéairement séparables
1	Peut approximer arbitrairement n'importe quelle fonction qui contient une cartographie continue d'un espace fini à un autre
2	Peut représenter une frontière de décision arbitraire de précision avec les fonctions d'activation rationnelles et peut approximer toute cartographie pour toute précision.

Tableau III.1: Détermination du nombre de couches cachées.

III.2.5. Le nombre de neurones dans les couches cachées:

Le choix du nombre de neurones de la couche cachée est une partie très importante dans le choix de l'architecture du réseau. Même si ces couches n'échangent pas directement avec l'environnement extérieur, elles ont une influence immense sur la sortie finale.

L'utilisation d'un trop petit nombre de neurones dans la couche cachée va provoquer un "underfitting", ce qui signifie que le réseau n'arrive pas à détecter les signaux dans un ensemble de données.

Utiliser un nombre trop important de neurones cachés peut entraîner un problème d'"overfitting" qui se produit quand la base de données n'est pas assez suffisante pour entraîner tous les neurones des couches cachées, ce qui augmente le temps d'apprentissage ainsi que de traitement des données, jusqu'à ce que le réseau ne puisse plus être entraîné de façon adéquate.

Un compromis doit être fait. Nous présentons ci-dessous quelques règles qui peuvent aider au choix du nombre de neurones cachés.

- le nombre de neurones cachés doit être dans la gamme entre la taille des couches d'entrée et de sortie.
- le nombre de neurones cachés doit être 2/3 de la taille de la couche d'entrée plus la taille de la couche de sortie.
- le nombre de neurones cachés doit être inférieur au double de la taille de la couche d'entrée.

Ces trois règles sont des points de départ que l'on peut considérer. D'autres méthodes sont utilisées pour trouver le nombre optimal de neurones dans la couche cachée.

Dans notre travail nous avons utilisé comme fonction d'activation la fonction sigmoïde et comme algorithme d'apprentissage la méthode de Levenberg-Marquardt exposée ci-après

III.2.6. Méthode de Levenberg-Marquardt (L.M)

Cette méthode d'apprentissage donne de très bons résultats en plus de sa rapidité de convergence, ce qui fait d'elle une méthode très populaire dans la recherche. Elle consiste à considérer la surface d'erreur quadratique en se basant sur la fonction d'énergie suivante:

$$E(k + 1) = E(k) + \dot{E}(k) \cdot \Delta W(k) + \frac{1}{2} \Delta W(k)^T \cdot \ddot{E}(k) \cdot \Delta W(k) \quad \text{III.1}$$

En posant pour chaque neurone i d'une couche l le vecteur de poids lui parvenant des neurones de la couche qui le précède par w_i^l , minimiser la fonction d'énergie revient à résoudre l'équation:

$$\frac{\partial E(k+1)}{\partial \Delta W(k)} = 0 \quad \text{III.2}$$

Avec $\Delta W(k)$ est le vecteur global des taux d'adaptation de tous les vecteurs de poids w_i^l .

Soit F_i^l les vecteurs des dérivées de l'erreur à la sortie du réseau par rapport à ces vecteurs poids w_i^l pour $l=1, \dots, L$, et $i=1, \dots, n_l$ ou n_l est le nombre de neurones dans la dernière couche et L le nombre de couches cachées.

$$F_i^l = \sum_{p=1}^K \sum_{i=1}^{n_l} (d_i^p - y_i^{(L)}) \left[-\frac{\partial y_i^p}{\partial w_i^l} \right] \quad \text{III.3}$$

K représente le nombre d'exemples d'entraînement. On définit F le vecteur rassemblant les vecteurs F_i^l qui est le Jacobien de l'erreur par rapport aux vecteurs poids. Afin de trouver l'optimum de III.2, on obtient:

$$H \Delta W = -F \quad \text{III.4}$$

H représente la matrice Hessienne (ou le Hessien).

L'équation III.4 représente un système d'équations linéaires dont la résolution peut être faite par la méthode de Gauss-Newton.

La méthode de Levenberg-Marquardt remplace le calcul du hessien par une approximation numérique donnée par :

$$H = FF^T + \lambda \Omega \quad \text{III.5}$$

λ est un coefficient positif, la matrice de FF^T étant définie semi-positive, le coefficient λ est utilisé afin de mieux conditionner la matrice H . Ainsi, la matrice Ω doit être d'un choix approprié. Le choix d'une matrice diagonale dont les éléments sont égaux aux éléments diagonaux de la matrice FF^T donne de bons résultats. Le choix d'une matrice identité peut être pris en considération. Les valeurs que doit prendre le coefficient λ doivent être convenablement choisies

En effet, la recherche doit être orientée dans la région de l'espace permettant de représenter adéquatement le problème non-linéaire par le modèle quadratique, afin de garantir la convergence de l'algorithme vers un minimum local.

La figure III.1 montre l'algorithme utilisé pour accomplir cette tâche. Il consiste à commencer avec une petite valeur et la faire augmenter jusqu'à ce que l'erreur diminue, la valeur maximale de ce paramètre doit être bornée.

Plusieurs variantes de cet algorithme ont été mises au point afin de le rendre plus efficace. Elles jouent, généralement, sur le choix optimal de λ , permettant de conduire vers un minimum global. Elles peuvent agir aussi sur le contrôle du gradient qui, en diminuant, ne permet plus de poursuivre la prospection dans d'autres directions, à la recherche d'un minimum.

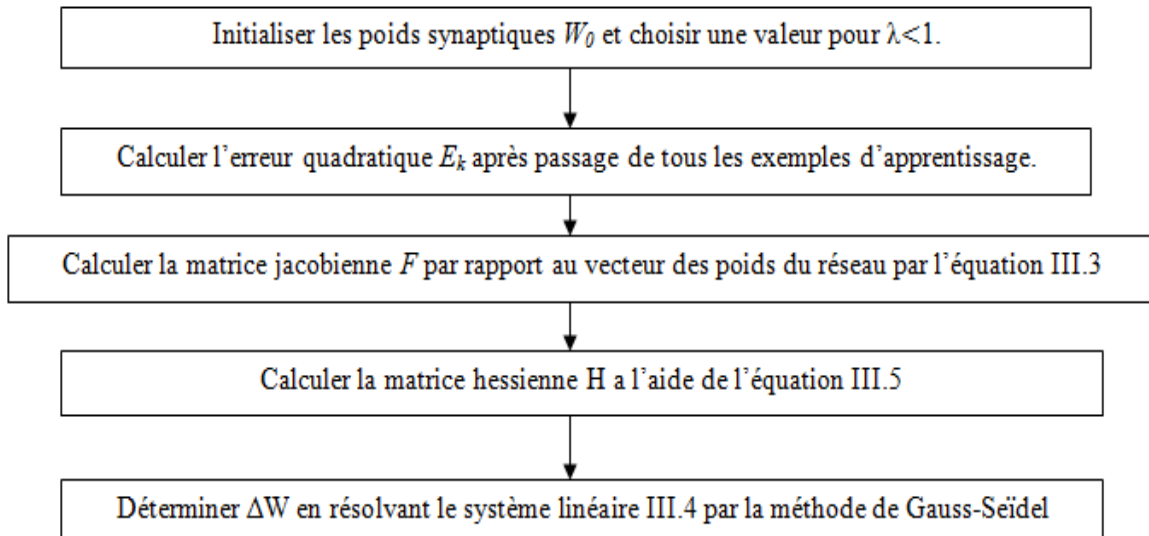


Figure III.1: Algorithme de la méthode de Levenberg-Marquardt

III.3. Méthode 2: Les cartes auto-organisées de Kohonen (Self-Organization Maps)

Les réseaux de neurones non supervisés peuvent classer un ensemble d'exemples de la base d'apprentissage en un nombre spécifique de catégories sans savoir par avance comment les organiser. Ce sont les données elles-mêmes qui servent de réponses car les poids du réseau sont modifiés en fonction de leurs composantes. Dans ce type de réseaux, l'information utile qui participe à l'apprentissage, se trouve donc uniquement dans les données.

Les SOM sont flexibles et faciles à implémenter dans la pratique. Pour rendre la prédiction indépendante de la nature de la caractéristique à prédire, nous transformons la base de données d'entrée par soustraire la valeur moyenne de la propriété considérée de la valeur réelle et diviser le résultat par la variance des données.

Alors, au lieu d'utiliser directement les vraies données, nous employons de cette façon les prétendus 'profils'. On doit noter ici que la prédiction de la caractéristique réelle est séparée du réseau de Kohonen, elle utilise seulement la base de données des profils prédits par les SOM.

Dans cette application, Nous avons utilisé une architecture dont le rayon du voisinage est $D=1$, et la topologie des neurones est hexagonale comme le montrent les figures III.2 et III.3.

Pendant l'apprentissage, le SOM trouve le neurone de sortie qui a la plus petite distance à partir des données d'apprentissage. Le SOM change les poids des neurones pour augmenter leurs similitudes avec les données, et il influe sur les poids des neurones voisins qui ont des relations aléatoires avec les données d'apprentissage.

Les neurones de sortie sont déplacés aux positions qui tracent la distribution des données de l'apprentissage. Après l'entraînement, les poids de chaque neurone modélisent les caractéristiques d'une classe dans les données [1].

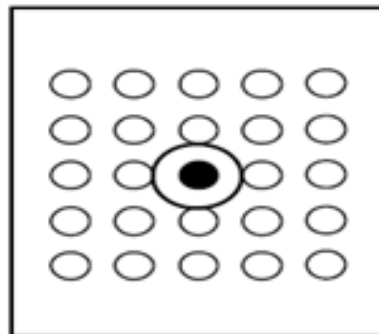


Figure III.2: Rayon de voisinage autour du neurone gagnant ($D=1$).

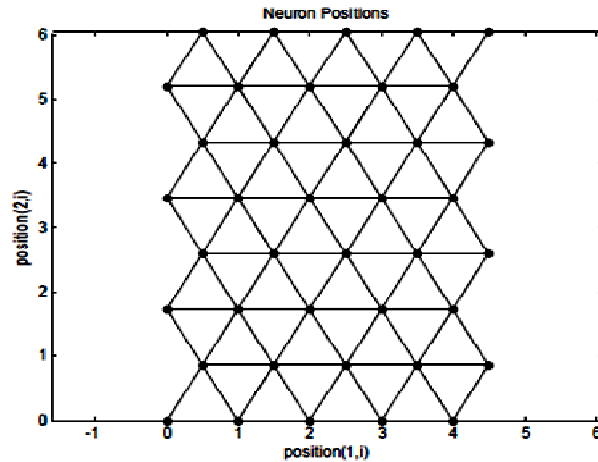


Figure III.3: Distribution et positions des neurones dans la carte de Kohonen en topologie hexagonale

III.3.1. Réseau de neurones SOM et règle d'apprentissage

Le réseau de Kohonen (Self-Organization Maps), souvent appelé la carte de Kohonen, se compose d'une seule couche de neurones fonctionnant en mode compétitif auto-organisé. Chaque neurone est alimenté par les n composantes du vecteur d'entrée, par les coefficients de pondération, formant un vecteur w_k à dimension N .

L'algorithme de l'auto-organisation (self-organization) est constitué par les séquences des opérations suivantes :

- Présenter le vecteur d'entrée e au réseau.
- Trouver la région dans le réseau où un neurone spécifique répond le plus fortement au vecteur e précédemment présenté. L'unité gagnante N_w est celle dont le vecteur de poids est le plus proche dans le sens d'une mesure de distance par rapport au vecteur d'entrée.
- Adapter les poids des neurones choisis de cette région. Répéter ces séquences plusieurs fois (selon les données d'entrée, la taille du réseau et les facteurs d'adaptations) mène le réseau à un état organisé, dans lequel chaque neurone représente une classe séparée des données. Dans l'algorithme généralisé de Kohonen, l'adaptation des poids des neurones trouvés dans le voisinage autour du neurone gagnant N_w , se fait selon la règle suivante:

$$w(i+1) = w(i) + \alpha_i \cdot G(i, e)[e - w_i] \quad \text{III.6}$$

e_i est le vecteur des poids du $i^{\text{ème}}$ neurone trouvé dans le voisinage du neurone gagnant et α_i est le taux d'apprentissage qui diminue avec le temps. Habituellement, c'est la diminution linéaire qui est utilisée, à partir d'une certaine valeur initiale et jusqu'à $\alpha_i = 0$.

Le voisinage du neurone gagnant diminue aussi avec le temps. L'une des méthodes les plus utilisées dans l'ajustement du voisinage est l'application de la fonction gaussienne définie autour du gagnant et décrite par la relation:

$$G(i, e) = \exp \frac{d_t^2(i, w)}{2\lambda^2} \quad \text{III.7}$$

$d_t(i, w)$ est la distance euclidienne entre le gagnant et le $i^{\text{ème}}$ neurone et λ est une constante.

Dans cette application, nous avons employé la règle d'apprentissage de Kohonen, pour l'ajustement des poids en utilisant la règle citée ci-dessus dans l'équation (III.6).

Supposons que le $i^{\text{ème}}$ neurone gagne, la $i^{\text{ème}}$ rangée de la matrice de poids d'entrée est ajustée comme suit:

$$w(i) = w(i-1) + \alpha_i \cdot (e(i) - w(i-1)) \quad \text{III.8}$$

$e(i)$ est le vecteur d'entrée correspondant.

La règle de Kohonen permet aux poids d'un neurone d'apprendre à reconnaître un vecteur d'entrée. C'est pour cette raison qu'il est utile dans la classification.

Le neurone dont le vecteur de poids était le plus proche du vecteur d'entrée est adapté pour être encore plus proche. Le résultat est que le neurone gagnant va encore gagner la compétition la prochaine fois qu'un vecteur semblable est présenté alors que la probabilité est moindre pour lui de gagner quand un vecteur très différent est présenté.

Par la suite, s'il y a assez de neurones, chaque classe de vecteurs d'entrée semblables aura un neurone qui donne une sortie égale à "1" quand un vecteur dans cette classe est présenté, tout en produisant un "0" à toutes autres fois. Ainsi, le réseau compétitif apprend à classer les vecteurs d'entrée qu'il voit. L'algorithme général des SOM pour la classification est donné à la figure.

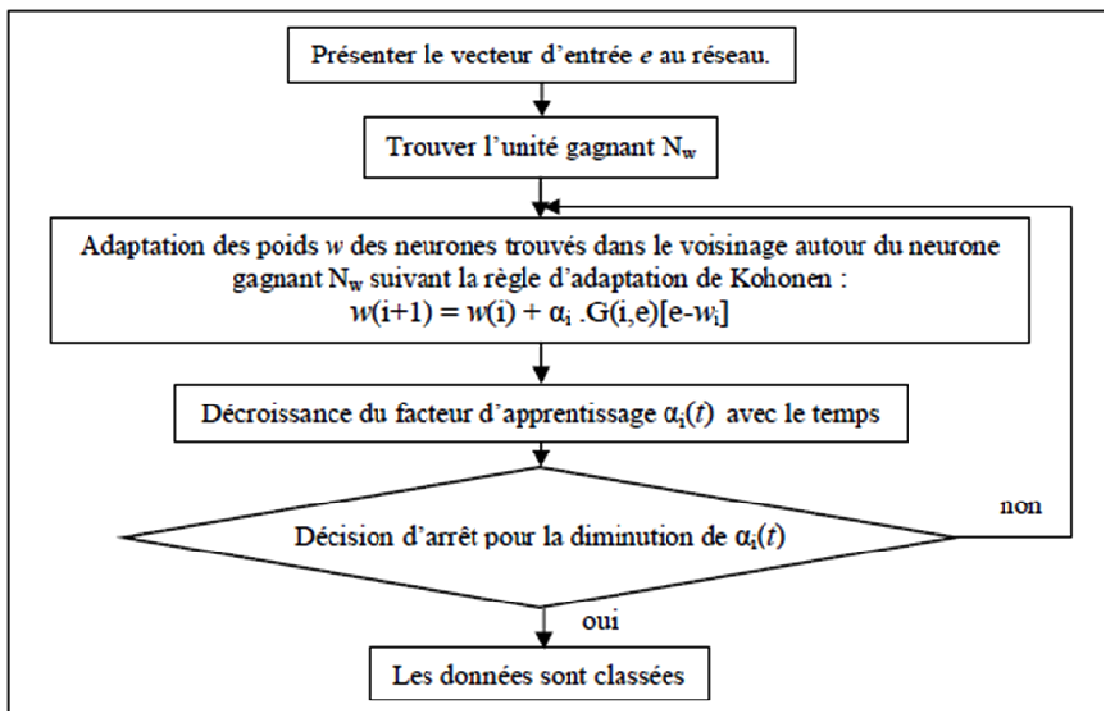


Figure III.4: Algorithme du classement des données utilisant la carte de Kohonen. [1]

III.3.2. Phase d'apprentissage du réseau de prédiction de Kohonen

Dans cette méthode, pour prédire une caractéristique, nous utilisons dans l'apprentissage deux autres données présentées avec celle à prédire.

Pour rendre la prédiction indépendante de la nature de la caractéristique à prédire, nous transformons la base de données d'entrée en faisant soustraire la valeur moyenne $P_m(t)$ de la propriété considérée de la valeur réelle $P(t)$ et en divisant le résultat par la variance $v(t)$ des données. Au lieu d'utiliser directement les vraies données, nous employons la méthode des prétendus "profils" définis par:

$$pr(t) = \frac{p(t) - p_m(t)}{v(t)} \quad \text{III.9}$$

Pour toutes les caractéristiques d'isolation participant à l'apprentissage, l'ensemble des profils forme les données d'apprentissage du réseau. Une fois que le réseau est entraîné, chaque neurone représenté les données les plus proches de son vecteur de poids dans le sens métrique choisi.

En mode d'apprentissage, quand nous voulons faire une prédiction pour une caractéristique, nous fournissons ses données expérimentales réelles et le SOM trouve les neurones gagnants correspondants. Le profil prédit $pr_d(t)$, est alors la moyenne pondérée de tous les neurones gagnants trouvés :

$$pr_d(t) = \frac{\sum_{i=1}^{n_g} k_i w_i}{\sum_{i=1}^{n_g} k_i} \quad \text{III.10}$$

n_g est le nombre de neurones gagnants.

k_i est le nombre d'apparitions du $i^{\text{ème}}$ neurone parmi les gagnants pour la caractéristique choisie et n_g est le nombre de neurones gagnants qui est égal au nombre de vecteurs d'entrée utilisées. w_i est le facteur de pondération attribué à chaque neurone gagnant i correspondant à l'une des trois caractéristique utilisées.

La prédiction de la caractéristique réelle est alors calculée a partir de l'équation (III.9) par:

$$P_d(t) = v_d(t).pr_d(t) + P_{md}(t) \quad \text{(III.11)}$$

Sur la base du profil prédit, la prédiction de la caractéristique réelle est calculée à partir de l'équation (III.11). Le point le plus important après la détermination du profil prédite $pr_d(t)$, est de prédire la valeur moyenne $P_{md}(t)$ et la variance $v_d(t)$ au temps (t). Pour cette tâche nous avons appliqué un modèle linéaire de prédiction, tenant compte des valeurs appropriées ($P_m(t)$ et $v(t)$) des autres caractéristiques. La prédiction linéaire de la valeur moyenne de la caractéristique $P_m(t)$ est donnée par *S.Osowski* :

$$P_m(t) = a i + b \quad \text{(III.12)}$$

i est une constante.

Nous dénotons par w_k les valeurs de pondération de chaque caractéristique d'apprentissage. a et b sont des solutions du problème dont la fonction coût est définie par :

$$E = \sum_{j=1}^n [(a_j + b) - P_j] w_j \quad \text{III.13}$$

$$b = \frac{\sum_{k=1}^n k w_k \frac{\sum_{k=1}^n k P_k w_k}{k^2 w_k} - \sum_{k=1}^n P_k w_k}{\sum_{k=1}^n k w_k \frac{\sum_{k=1}^n k w_k}{k^2 w_k} - \sum_{k=1}^n w_k} \quad \text{III.14}$$

$$a = \frac{\sum_{k=1}^n k P_k w_k - b \sum_{k=1}^n k w_k}{\sum_{k=1}^n k^2 w_k} \quad \text{III.15}$$

Dans la méthode de calcul de la moyenne pondérée donnée par (III.10), le neurone gagnant de la caractéristique à prédire est pondéré par la valeur la plus élevée proche de 1 et les autres caractéristiques sont pondérées avec des valeurs décroissantes. Ce mode de pondération est employé quand l'allure de la caractéristique à prédire est croissante. L'opposé est appliqué pour les caractéristiques décroissantes.

Différentes simulations, pour la méthode de pondération adoptée ci-dessus, ont été faites pour trouver les meilleures valeurs pour les coefficients de pondérations w_k . Comme résultat de ces simulations et pour une courbe croissante, nous avons posé $w_1=0.9$ pour les données de la caractéristique à prédire, $w_2=0.6$ et $w_3=0.1$ pour les autres caractéristiques d'apprentissage. Pour une courbe descendante : $w_1=0.1$, $w_2=0.5$, $w_3=0.9$.

Si nous voulons prédire pour une période plus longue que celle de l'expérience, nous entraînons le réseau ayant n valeurs dans le vecteur d'entrée. Après la première prédiction, le vecteur de sortie contiendra **n+1** valeurs. Avec ce vecteur contenant **n+1** valeurs, l'apprentissage sera lancé encore pour prédire une valeur **n+2** et ainsi de suite. Nous devons noter ici que la prédiction n'est pas totalement faite par le réseau de Kohonen SOM, mais elle l'utilise dans la phase de prédiction des profils.

CHAPITRE IV

Présentation des résultats

IV.1. Introduction

Après avoir présenté les méthodes de la Back-propagation BP et des cartes auto-organisées de Kohonen (Self-Organization Maps), nous présentons dans les sections qui viennent l'application des réseaux de neurones à la prédiction de l'évolution de la tension de contournement et du courant de fuite à la surface du modèle plan pour différentes conductivités et différents niveaux de tension appliqués.

Pour les deux méthodes, si nous voulons prédire pour un intervalle plus long que celui de l'expérience, nous entraînons le réseau ayant n valeurs dans le vecteur d'entrée. Après la première prédiction, le vecteur de sortie contiendra $n+1$ valeurs. Avec ce vecteur contenant $n+1$ valeurs, l'apprentissage sera lancé encore pour prédire une valeur $n+2$ et ainsi de suite.

L'outil mathématique qui nous a permis de faire une comparaison entre les architectures est la fonction MSE (Mean squared error performance function). C'est une fonction qui mesure la différence entre les courbes expérimentales et prédites selon l'erreur des moindres carrés.

IV.2. Méthode1 : La Back-propagation

Plusieurs variantes de l'algorithme de la back-propagation sont utilisées. Pour chaque algorithme, plusieurs architectures sont proposées. Nous avons remarqué l'influence de plusieurs paramètres :

- Le nombre de couches cachées.
- Le nombre de neurones dans la couche.
- L'ajout et l'inhibition des biais pour chaque couche.
- les fonctions d'activations pour chaque couche.
- La variation des paramètres d'apprentissage.

Pour l'algorithme de la BP, trois méthodes existent : la BP avec momentum, la BP avec momentum et à adaptation de taux d'apprentissage (MADPTA) et la méthode de Levenberg-Marquardt (L.M).

Nous avons choisi pour la prédiction de l'évolution de la tension de contournement et du courant de fuite d'utiliser la méthode de Levenberg-Marquardt (L.M). Pour cette méthode le paramètre dominant est le nombre de couches cachées [1], c'est pour cette raison que nous avons présenté deux architectures, la première à une couche cachée et la deuxième à deux couches cachées.

IV.2.1. Prédiction de la tension de contournement :

Pour la prédiction de la tension de contournement nous avons utilisé plusieurs architectures pour les différents cas de pollution (40g/l, 100g/l et 200g/l de Kaolin). Ces architectures diffèrent essentiellement dans le nombre de couches cachées et le nombre de neurones dans chaque couche.

Nous avons utilisé deux architectures contenant respectivement 1 et 2 couches cachées. Nous avons choisi de ne pas concevoir d'architectures dont le nombre de couches cachées dépasse 2 car il a été démontré dans des travaux précédents qu'il n'y a pas d'intérêt théorique à avoir des architectures comportant plus de 2 couches cachées [26].

Nous choisissons 3 variantes différentes pour les différents cas de concentration de Kaolin identiques, où le nombre de neurones utilisé dans chaque couche cachée est 3, 5 et 10

Le vecteur de sortie qui caractérise les valeurs prédites a été calculé sur 2 étapes : la première étape consiste à trouver la première valeur prédite en dehors de l'intervalle (dans notre cas c'est pour une conductivité de 25mS/cm) et la deuxième étape consiste à modifier le vecteur d'entrée tout en ajoutant une nouvelle valeur (qui a été calculée dans l'étape une). Le nouveau vecteur de sortie aura donc deux valeurs prédites (qui correspondent aux conductivités 25mS/cm et 30mS /cm).

a.) Pour une pollution de 40g/l de Kaolin :

a.1) Réseau à une couche cachée :

Réseau à une couche cachée de 3 neurones :

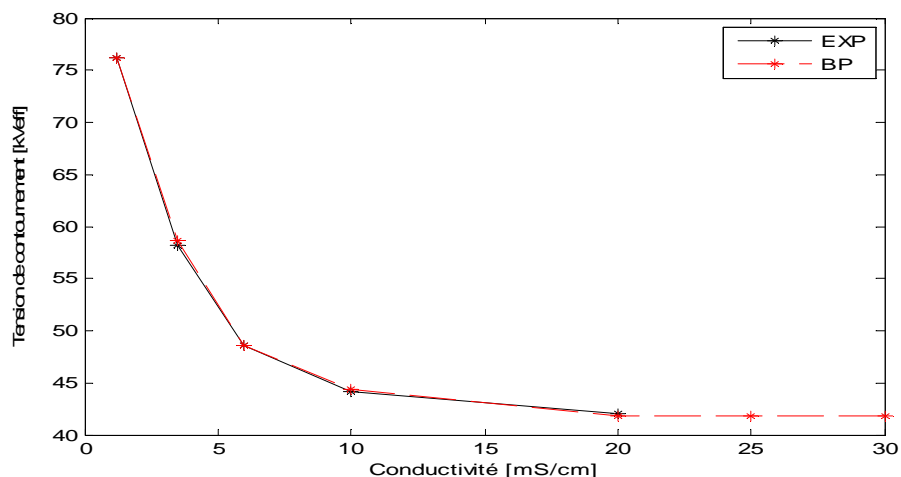


Figure IV.1: Prédiction de la tension de contournement pour une pollution de 40g/l de Kaolin utilisant un réseau FFBP à une couche cachée de 3 neurones. mse=0.211 V².

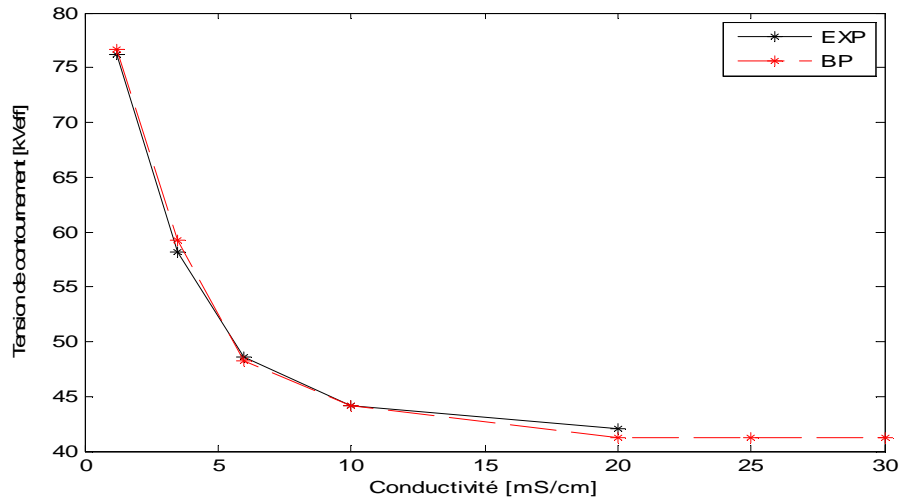
Réseau à une couche cachée de 5 neurones

Figure IV.2: Prédiction de la tension de contournement pour une pollution de 40g/l de Kaolin utilisant un réseau BP à une couche cachée de 5 neurones. $mse=0.517 V^2$.

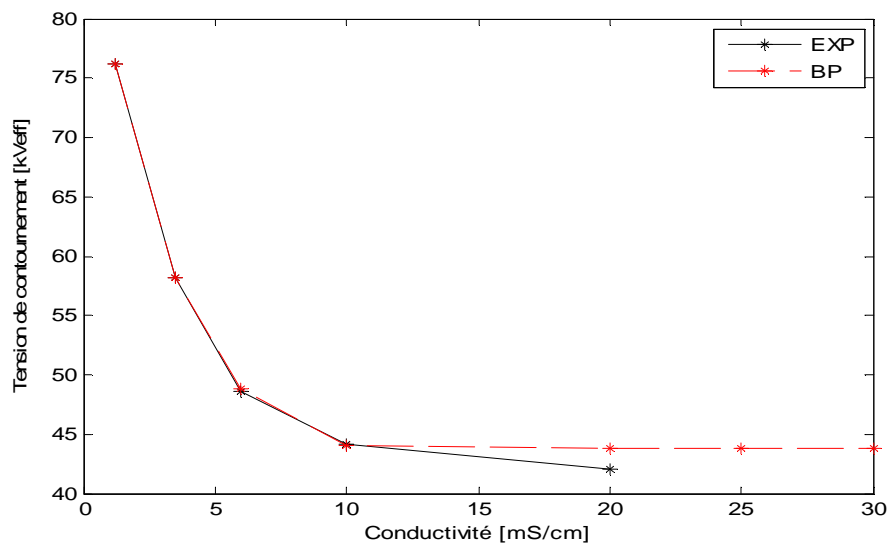
Réseau à une couche cachée de 10 neurones :

Figure IV.3: Prédiction de la tension de contournement pour une pollution de 40g/l de Kaolin utilisant un réseau BP à une couche cachée de 10 neurones. $mse=0.653 V^2$.

a.2) Réseau à 2 couches cachées:

Réseau à 2 couches cachées ,3 neurones dans chaque couches :

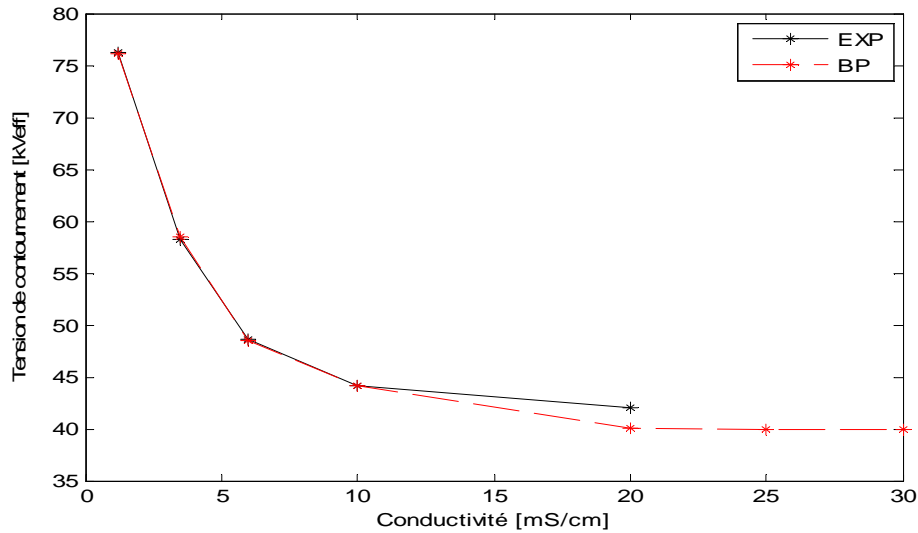


Figure IV.4: Prédiction de la tension de contournement pour une pollution de 40g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,3 neurones dans chaque couche. mse=0.785 V².

Réseau à 2 couches cachées ,5 neurones dans chaque couches :

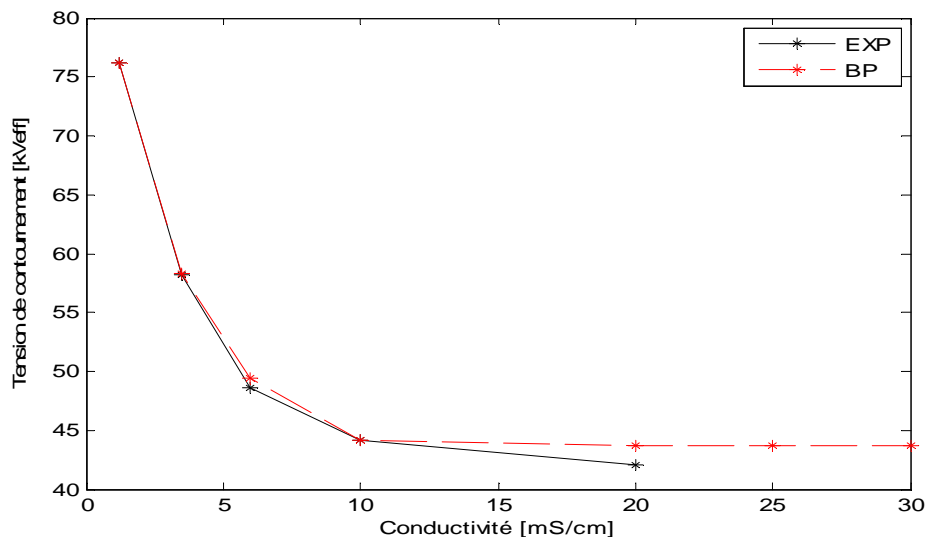


Figure IV.5: Prédiction de la tension de contournement pour une pollution de 40g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,5 neurones dans chaque couche. mse=0.696 V².

Réseau à 2 couches cachées ,10 neurones dans chaque couches :

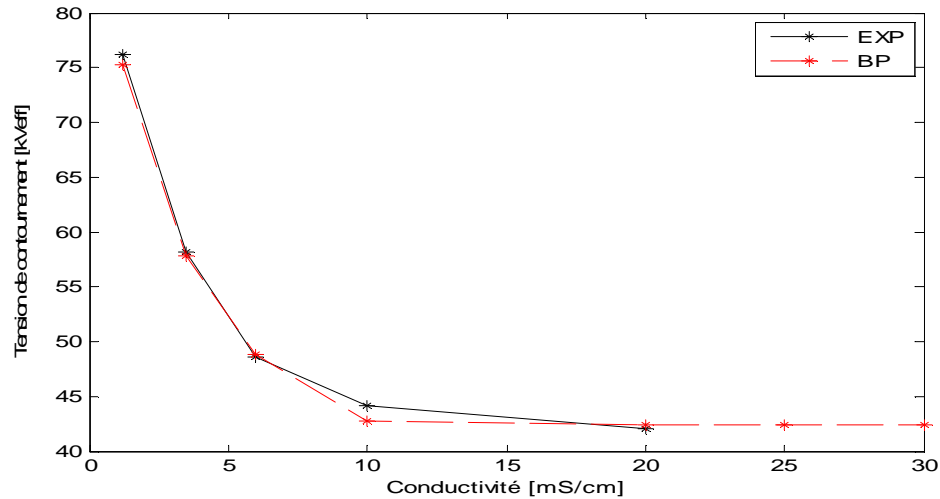


Figure IV.6: Prédiction de la tension de contournement pour une pollution de 40g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,10 neurones dans chaque couche. mse=0.699 V².

b) Pour une pollution de 100g/l de Kaolin :

b.1) Réseau à une couche cachée :

Réseau à une couche cachée de 3 neurones :

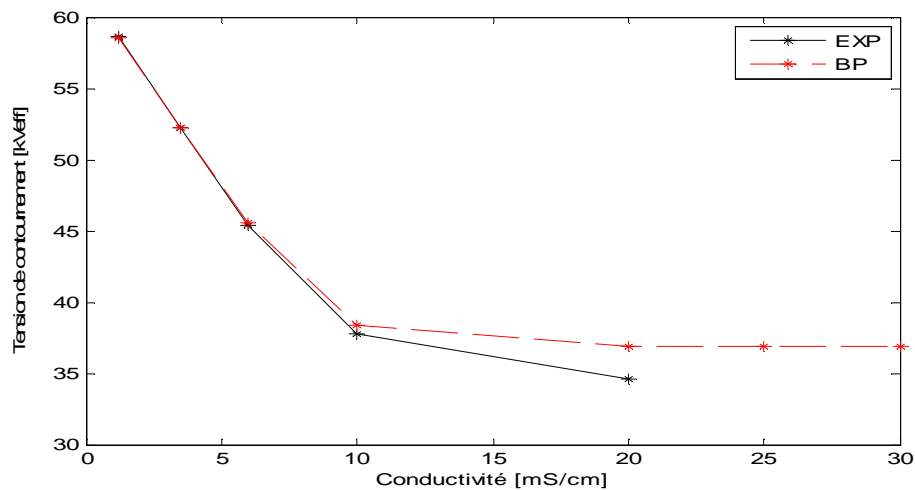


Figure IV.7: Prédiction de la tension de contournement pour une pollution de 100g/l de Kaolin utilisant un réseau BP à une couche cachée de 3 neurones. mse=1.135 V².

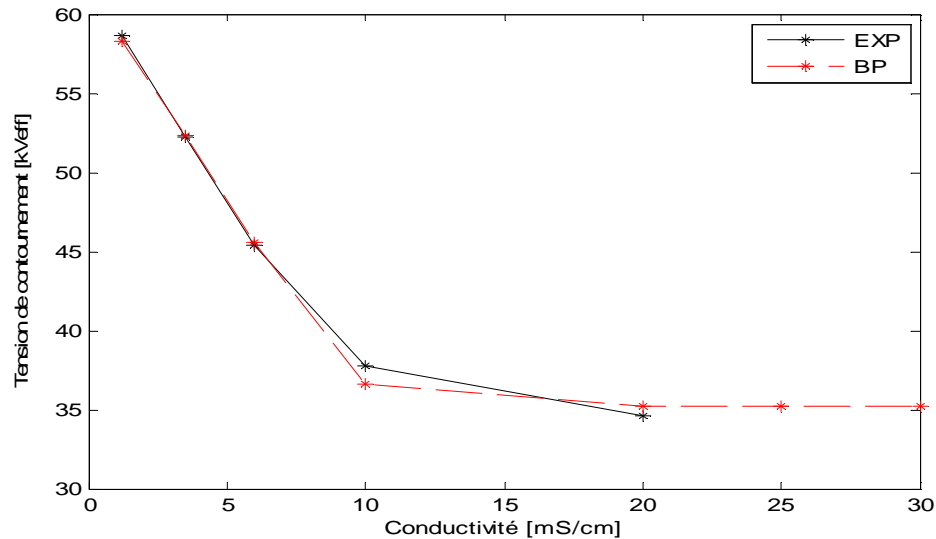
Réseau à une couche cachée de 5 neurones :

Figure IV.8: Prédiction de la tension de contournement pour une pollution de 100g/l de Kaolin utilisant un réseau BP à une couche cachée de 5 neurones. $mse=0.392 V^2$.

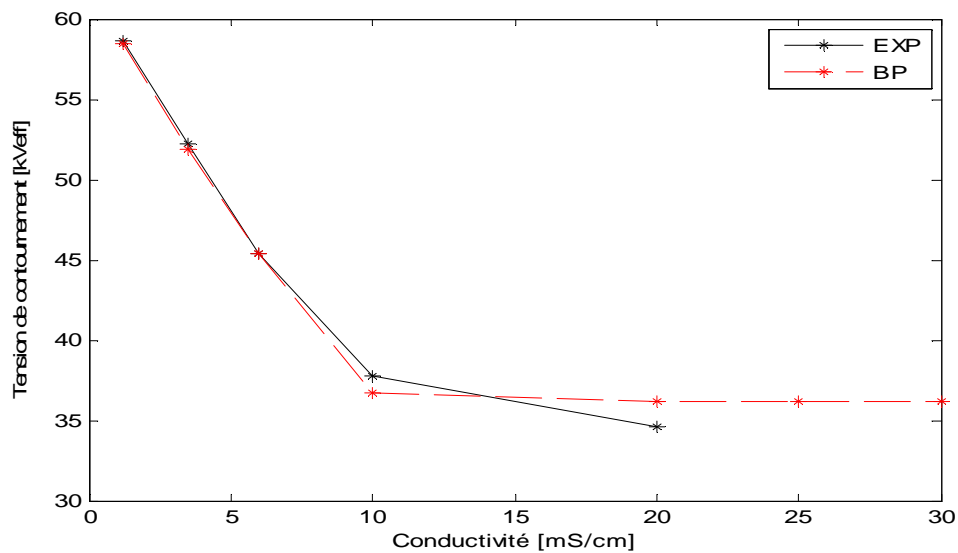
Réseau à une couche cachée de 10 neurones :

Figure IV.9: Prédiction de la tension de contournement pour une pollution de 100g/l de Kaolin utilisant un réseau BP à une couche cachée de 10 neurones. $mse=0.766 V^2$.

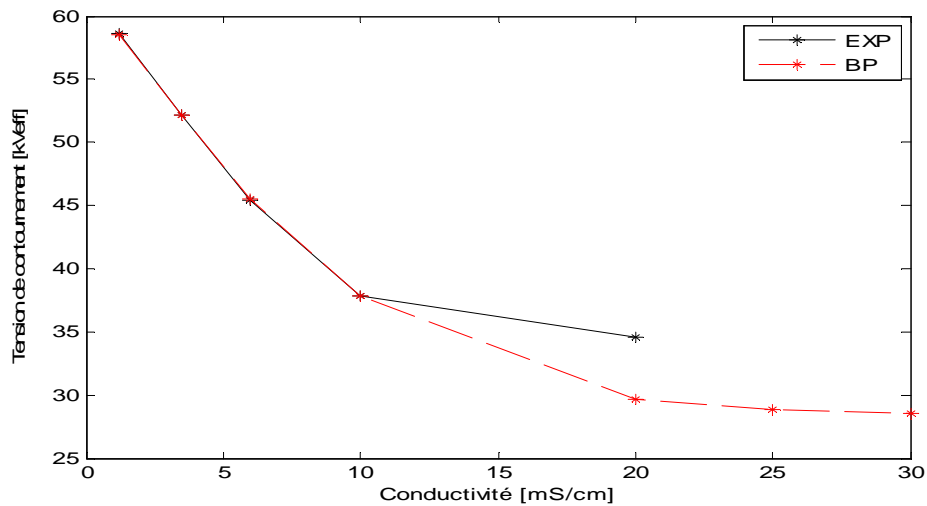
b.2) Réseau à 2 couches cachées:**Réseau à 2 couches cachées ,3 neurones dans chaque couches :**

Figure IV.10: Prédiction de la tension de contournement pour une pollution de 100g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,3 neurones dans chaque couche. mse=4.889 V².

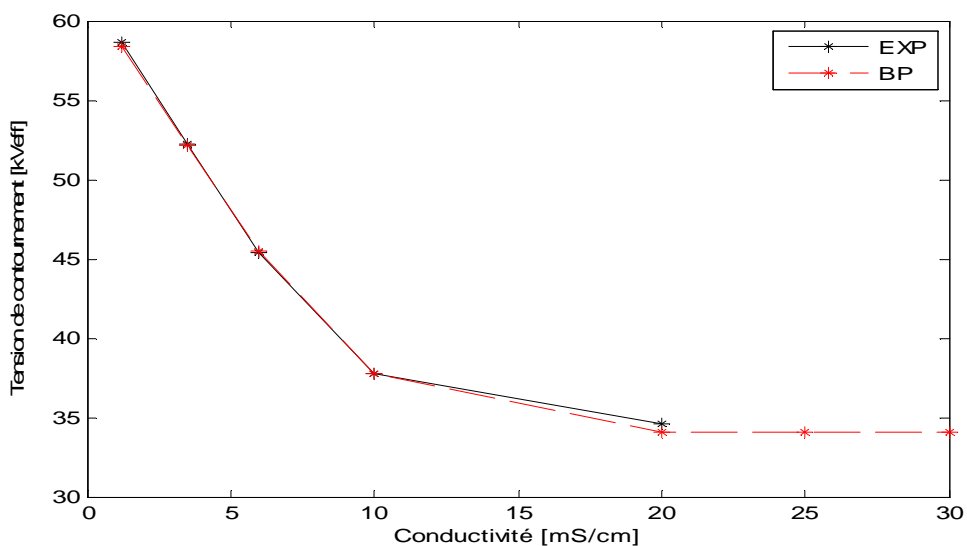
Réseau à 2 couches cachées ,5 neurones dans chaque couches :

Figure IV.11: Prédiction de la tension de contournement pour une pollution de 100g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,5 neurones dans chaque couche. mse=0.074 V².

Réseau à 2 couches cachées ,10 neurones dans chaque couches :

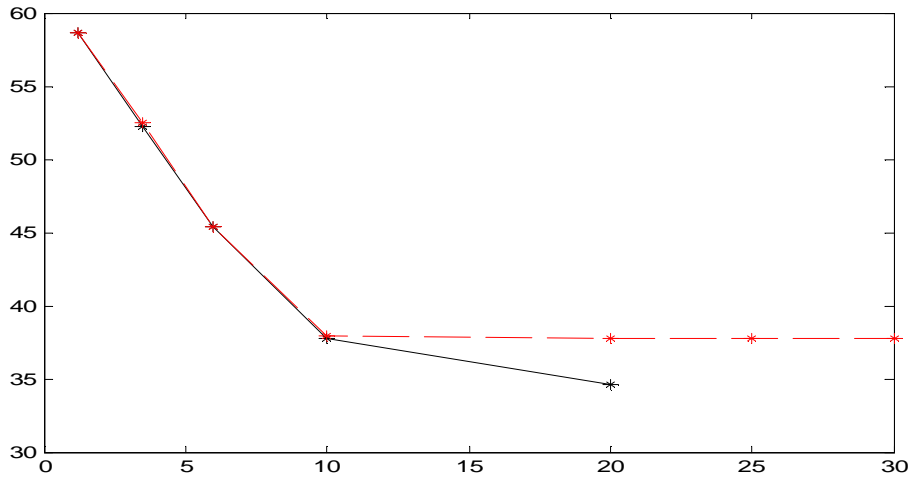


Figure IV.12: Prédiction de la tension de contournement pour une pollution de 100g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,10 neurones dans chaque couche. $mse=4.456 V^2$.

c) Pollution de 200g/l de Kaolin :

c.1) Réseau à 1 couche :

Réseau à 1 couche de 3 neurones :

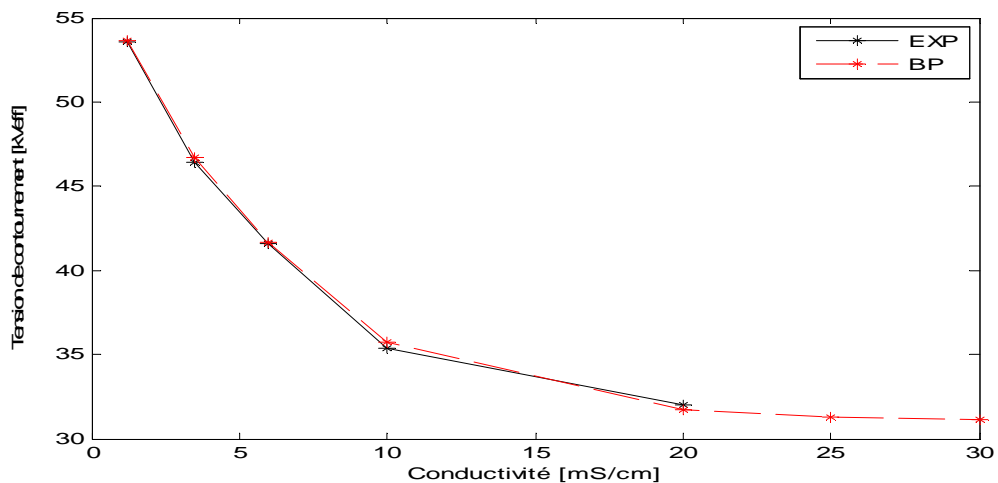


Figure IV.13: Prédiction de la tension de contournement pour une pollution de 200g/l de Kaolin utilisant un réseau BP à une couche cachée de 3 neurones. $mse=0.061 V^2$.

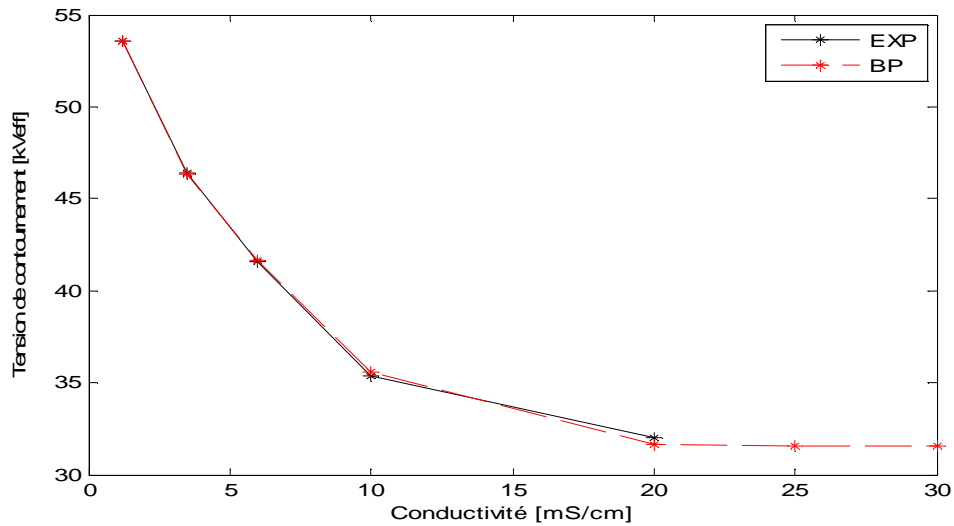
Réseau à 1 couche de 5 neurones :

Figure IV.14: Prédiction de la tension de contournement pour une pollution de 200g/l de Kaolin utilisant un réseau BP à une couche cachée de 5 neurones. $mse=0.036 V^2$.

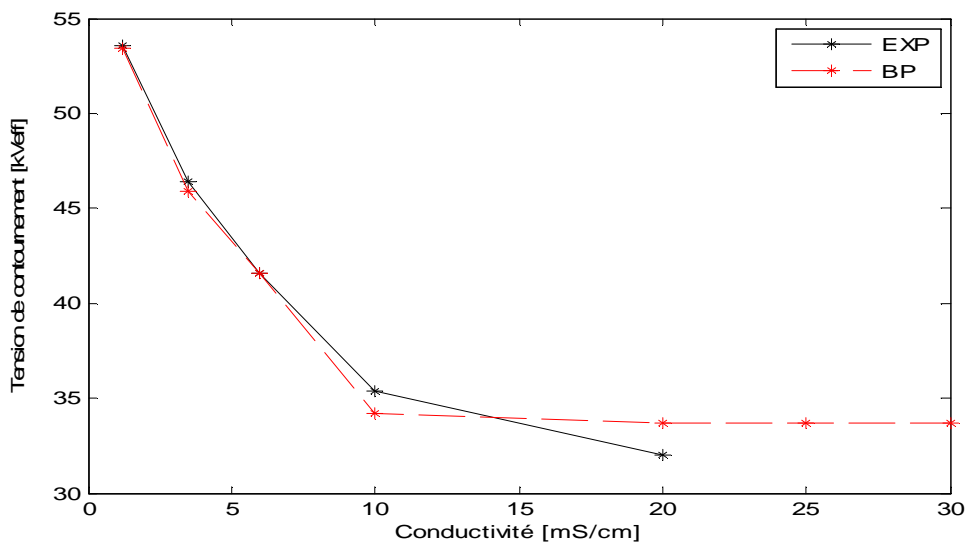
Réseau à 1 couche de 10 neurones :

Figure IV.15: Prédiction de la tension de contournement pour une pollution de 200g/l de Kaolin utilisant un réseau BP à une couche cachée de 10 neurones. $mse=0.940 V^2$.

c.2) Réseau à 2 couches :

Réseau à 2 couches de 3 neurones :

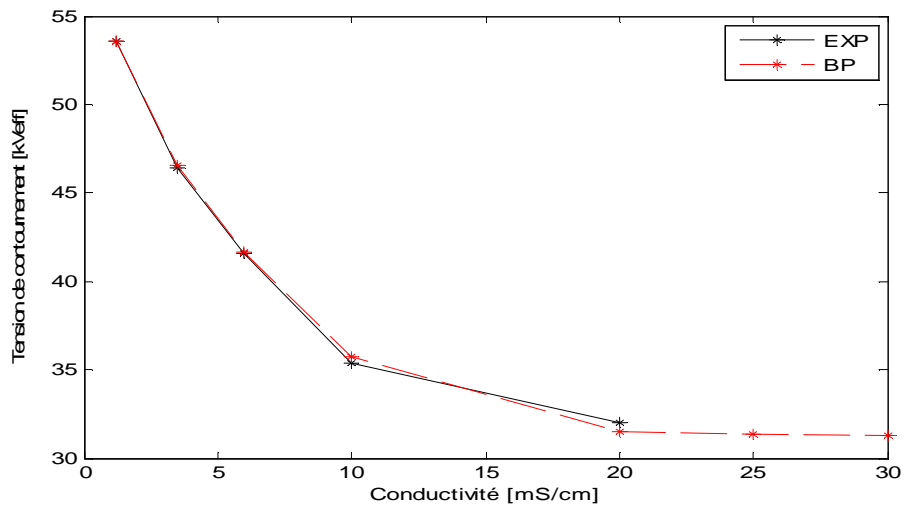


Figure IV.16: Prédiction de la tension de contournement pour une pollution de 200g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,3 neurones dans chaque couche. mse=0.082 V².

Réseau à 2 couches de 5 neurones :

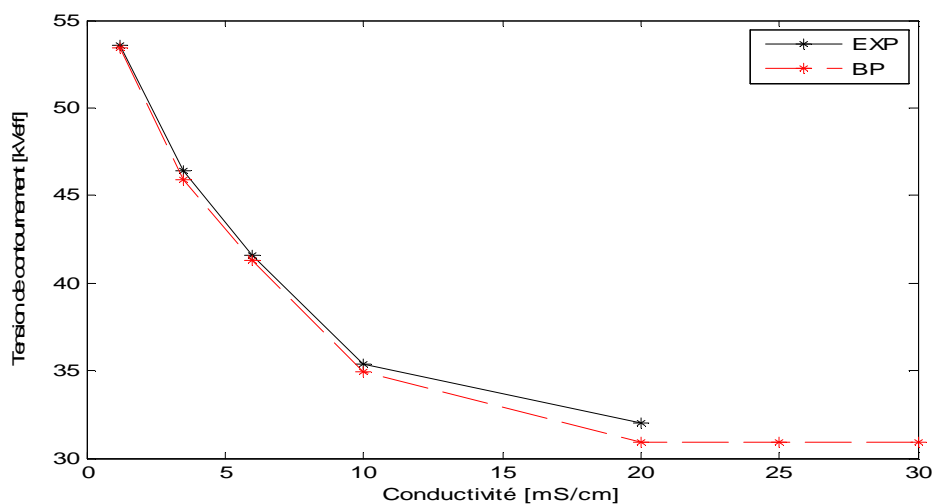


Figure IV.17: Prédiction de la tension de contournement pour une pollution de 200g/l de Kaolin utilisant un réseau BP à 2 couches cachées ,5 neurones dans chaque couche. mse=0.361 V².

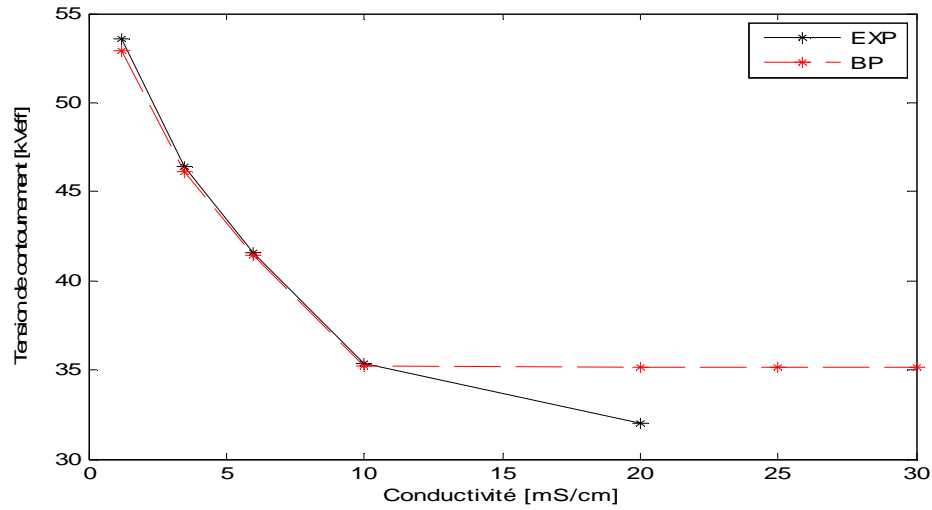
Réseau à 2 couches de 10 neurones :

Figure IV.18: Prédiction de la tension de contournement pour une pollution de 200g/l de Kaolin utilisant un réseau BP à 2 couches cachées, 10 neurones dans chaque couche. $mse=1.684V^2$.

IV.2.2. Prédiction du courant de fuite :

Comme pour la tension de contournement, nous avons essayé de donner l'évolution du courant de fuite en dehors de l'intervalle de l'expérience et nous avons pu prédire la valeur du courant pour des tensions appliquées de 25kVeff et 35kVeff.

L'architecture utilisée est à 1 couche cachée de 3 neurones.

a) Pour une pollution de 40g/l de Kaolin :

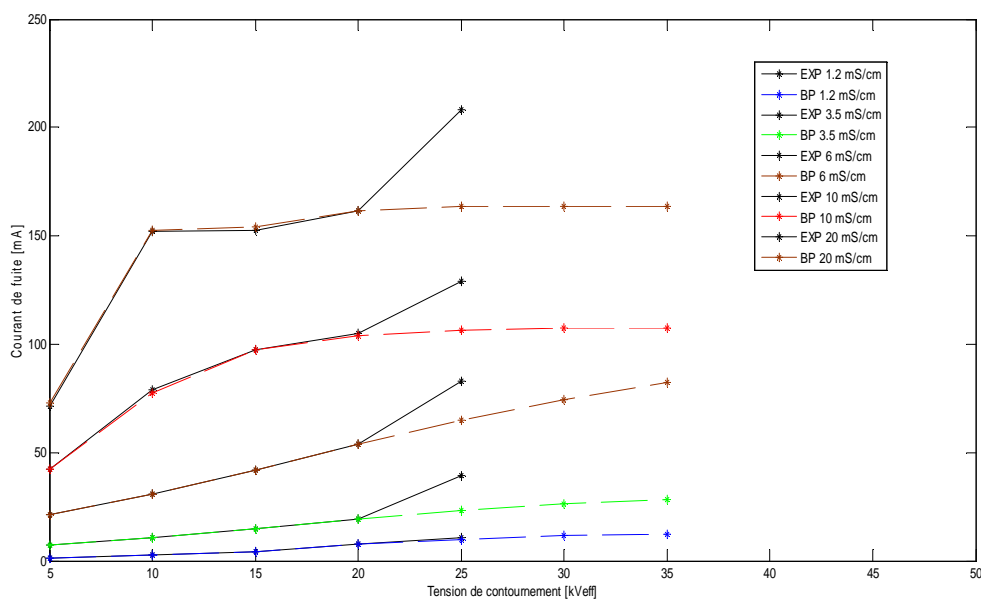


Figure IV.19 Prédiction du courant de fuite pour une pollution de 40g/l de Kaolin utilisant un réseau BP à une couche cachée de 10 neurones.

b) Pour une pollution de 100g/l de Kaolin :

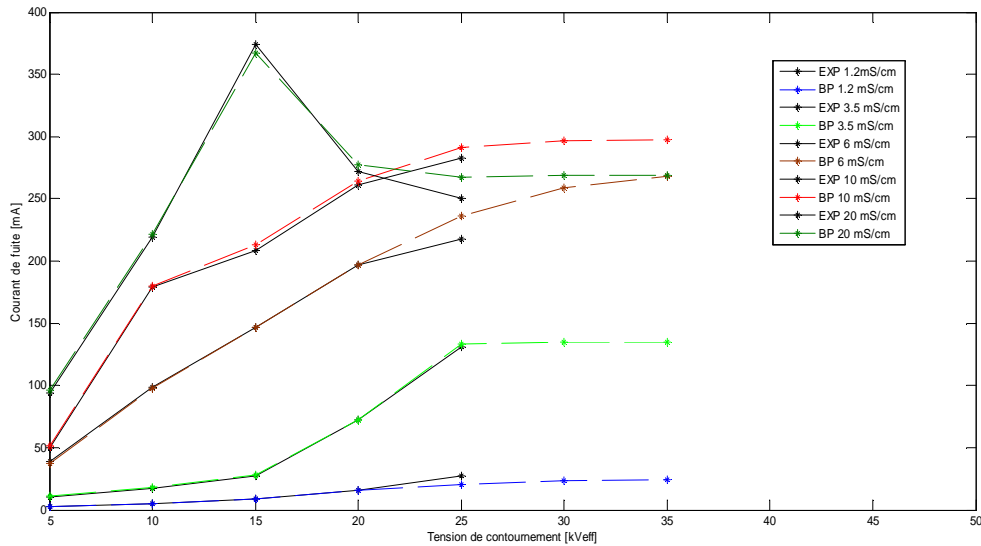


Figure IV.20 Prédiction du courant de fuite pour une pollution de 100g/l de Kaolin utilisant un réseau BP à une couche cachée de 10 neurones.

c) Pollution pour 200g/l

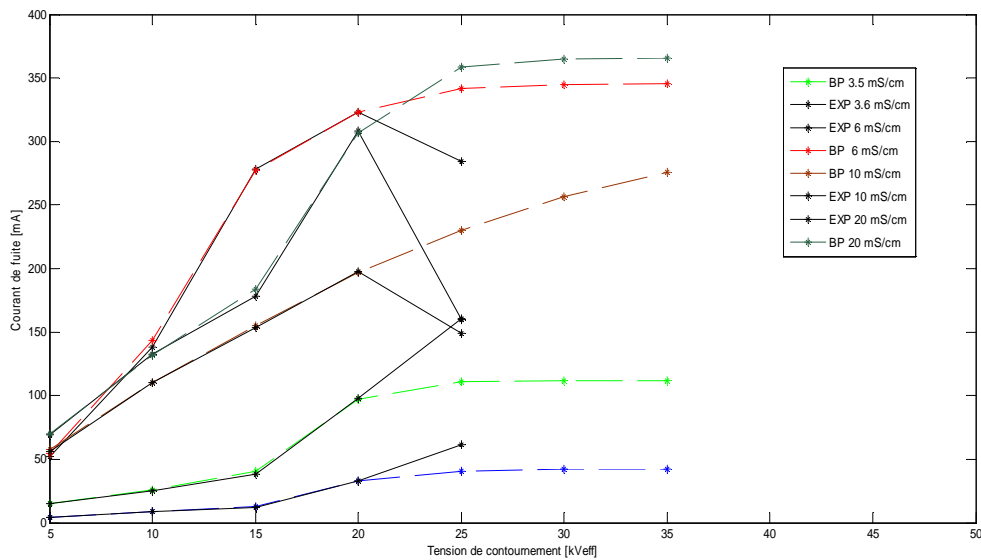


Figure IV.21 Prédiction du courant de fuite pour une pollution de 200g/l de Kaolin utilisant un réseau BP à une couche cachée de 10 neurones.

Le calcul de la performance de l'erreur des moindres carrées pour la BP est donné par le tableau suivant :

Cond Poll	1.2	3.5	6	10	20
40	0.101	50.506	63.257	100.210	405.705
100	9.716	1.372	68.954	21.866	58.111
200	86.115	495.251	500.517	1317	7937

Cond : la conductivité en [mS/cm] ;

Poll : la pollution en [g/l] de Kaolin

Tableau.IV.1. la mse [A^2] calculée pour la prédiction du courant de fuite par la BP

IV.3. Méthode 2 : Réseau de neurones SOM :

Pour cette méthode, nous allons utiliser la même base de données utilisée précédemment pour la BP.

Les données expérimentales de la tension de contournement et du courant de fuite ont été d'abord converties en 'profils', qui sont employés après, dans la phase d'apprentissage. Dans ce cas nous considérons la même architecture que celle considérée par L.Mokhnache[1] pour pouvoir comparer nos résultats au siens et d'arriver éventuellement à un résultat. Le réseau de Kohonen utilisé donc contient 40 neurones distribués sur des hexagones et le rayon de voisinage est donné par $D=1$. Le nombre d'itération est 150.

Après l'apprentissage, les neurones sont groupés dans des classes séparées correspondant au type de courbes présentées: ascendante (pour le courant de fuite) et descendante (pour la tension de contournement). Cette phase est nécessaire pour prédire le profil. Nous notons que toutes les courbes expérimentales utilisées sont non linéaires.

IV.3.1. Prédiction de la tension de contournement :

Nous présentons ci-dessous la prédiction de la tension de contournement pour des valeurs de conductivité dépassant 20mS/cm.

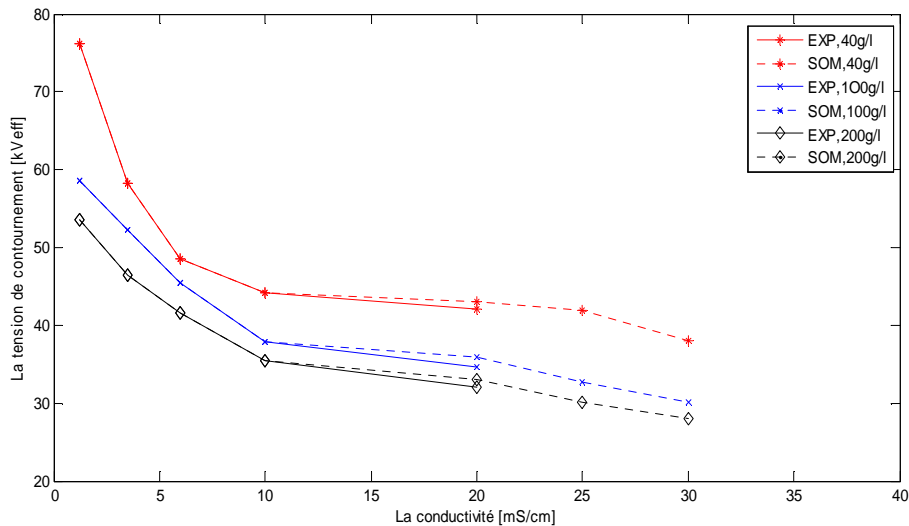


Figure IV.22 Prédiction de la tension de contournement utilisant une carte de Kohonen de 40 neurones. (Architecture hexagonale, $D=1$ et le nombre d'itération est 150).

Le calcul de la performance de l'erreur des moindres carrés a donné les valeurs suivantes : 0.20, 0.39 et 0.192 pour les valeurs de pollution 40 g/l, 100 g/l et 200 g/l respectivement.

IV.3.2. Prédiction du courant de fuite :

Nous présentons dans la section qui suit la prédiction du courant de fuite pour des valeurs de tensions appliquées allant de 5kVeff jusqu'à 35kVeff.

a) Pour une pollution de 40g/l :

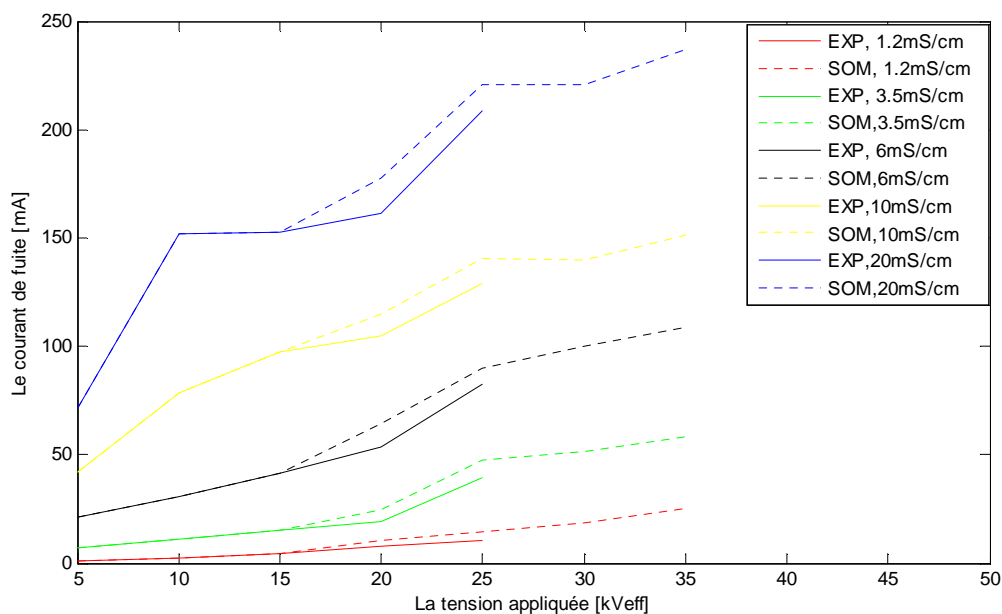


Figure IV.23 Prédiction du courant de fuite pour une pollution de 40g/l de Kaolin utilisant une carte de Kohonen de 40 neurones. (Architecture hexagonale, D=1et le nombre d'itération est 150).

b) Pour une pollution de 100g/l :

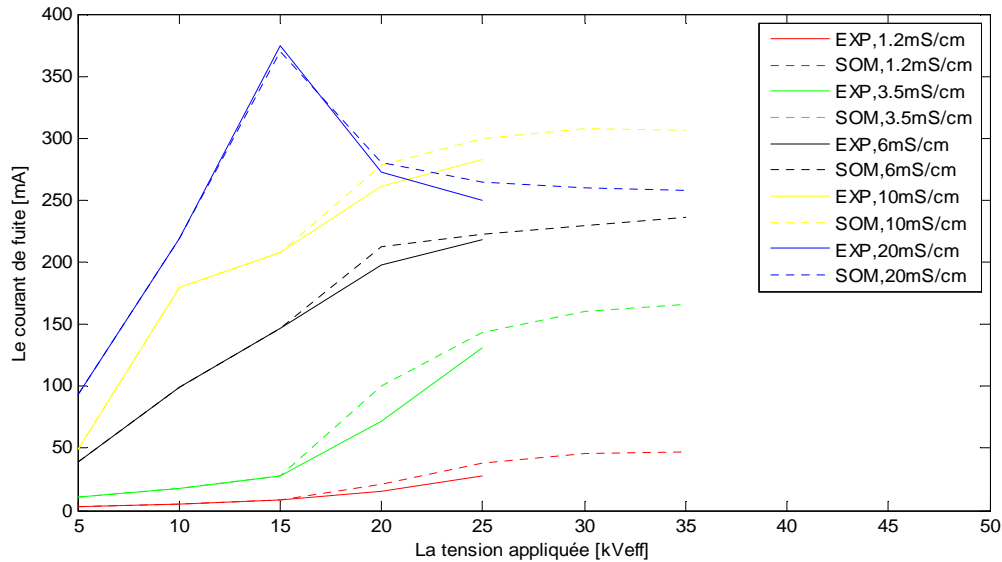


Figure IV.24 Prédiction du courant de fuite pour une pollution de 100g/l de Kaolin utilisant une carte de Kohonen de 40 neurones. (Architecture hexagonale, D=1 et le nombre d'itération est 150).

c) Pour une pollution de 200g/l :

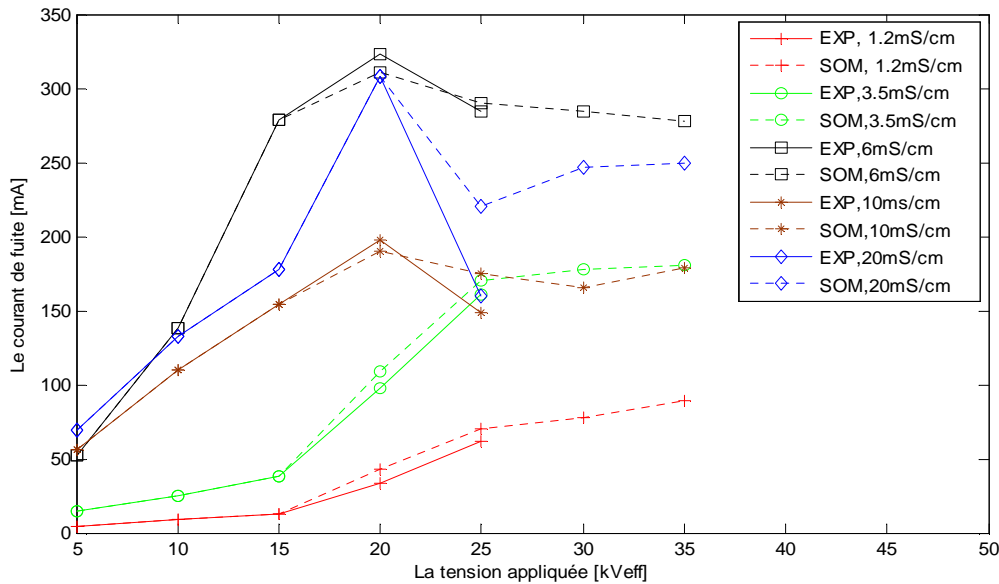


Figure IV.25 Prédiction du courant de fuite pour une pollution de 200g/l de Kaolin utilisant une carte de Kohonen de 40 neurones. (Architecture hexagonale, D=1 et le nombre d'itération est 150).

Le calcul de la performance de l'erreur des moindres carrées pour les SOM est donné par le tableau suivant :

Cond Poll	1.2	3.5	6	10	20
40	4.77	19.68	33.70	47.23	82.63
100	24.72	180.64	54.63	115.33	61.00
200	36.00	43.12	39.61	147.70	729.63

Cond : la conductivité en [mS/cm] ; Poll : la pollution en [g/l] de Kaolin

Tableau IV.2 la mse calculée pour la prediction du courant de fuite pour par les SOM

IV.4. Observations :

Après traitement des résultats expérimentaux par les RNA nous notons les observations suivantes :

- Pour la tension de contournement :

Méthode 1 : La FFBP

Pour le cas de la solution polluante de 40 g/l de Kaolin, l'architecture à 3 neurones est la plus performante pour le réseau à une couche cachée, tandis que l'architecture à 5 neurones s'est avérée être la meilleure pour le réseau à deux couches cachées.

Pour le cas de la solution polluante de 100 g/l de Kaolin, l'architecture à 5 neurones est la plus performante dans les réseaux à une et deux couches cachées.

Pour le cas de la solution polluante de 200 g/l de Kaolin, l'architecture à 5 neurones est la plus performante pour le réseau à une couche cachée, tandis que l'architecture à 3 neurones s'est avérée être la meilleure pour le réseau à deux couches cachées.

Méthode 2 : Les SOM

Nous remarquons que les valeurs de la mse pour la méthode des SOM sont inférieures à celles de la méthode BP.

-Pour le courant de fuite :

Nous remarquons d'après les tableaux IV.1 et IV.2 que pour des valeurs de conductivité élevées (plus de 6 mS/cm), la méthode des SOM sera plus performante que la BP.

IV.5. Conclusion:

- L'évolution des caractéristiques prédites (tension de contournement et courant de fuite) dépend essentiellement de l'architecture du réseau utilisé.
- La taille de la base de données influe considérablement sur la qualité de l'apprentissage des réseaux de neurones.
- la base de données étant réduite, les réseaux à une couche se sont montrés plus performants que les réseaux à deux couches pour la Feed Forward Back-Propagation.
- Les paramètres d'apprentissage (nombre de neurones, leurs distribution...) influe sur le déroulement de l'apprentissage et sa qualité. Leur optimisation n'est facile à faire.
- Pour une même architecture, l'erreur change selon la base de données offerte à l'architecture. Cette base de données, représentée par sa courbe expérimentale, est un paramètre qui contrôle l'apprentissage et la courbe qui contient un point d'inflexion reste difficile à être soumise à l'apprentissage.
- La méthode des cartes auto-organisées de Kohonen a offert une bonne qualité de prédiction comparant à la méthode de BP.
- les réseaux de neurones ne donnent pas de bons résultats pour les fonctions non monotones.

Conclusion Générale

Conclusion générale

Dans ce travail nous avons appliqué l'une des méthodes d'intelligence artificielle les plus utilisées, à savoir les RNA pour la prédiction du comportement des isolateurs de haute tension soumis à une pollution uniforme. Les conclusions suivantes mettent en avant les résultats obtenus :

- La prédiction de la tension de contournement ainsi que du courant de fuite des isolateurs pollués par des outils informatiques aide considérablement la recherche, vu les limitations du matériel expérimental et des instruments de mesure dans les laboratoires. Sans oublier le gain de temps et d'argent colossal que ces méthodes apportent.
- L'architecture du RNA (nombre de neurones, la topologie,...) est un facteur important décidant de la qualité de l'apprentissage plus que les paramètres d'apprentissage (taux d'apprentissage,...).
- L'autre élément très important dans l'utilisation des RNA est la base de données. La fiabilité du traitement des résultats dépend d'elle directement. Cette base de données doit être assez consistante (en taille et en précision) pour que le réseau sache prédire les valeurs futures avec une erreur minimale.
- Pour la méthode Back-Propagation, plus le nombre de neurones des couches cachées est important moins les résultats sont précis. Dans ce cas on se retrouve devant le phénomène d'overfitting ou les neurones cachés ne sont pas tous entraînés
- La méthode des cartes auto-organisées de Kohonen a offert une bonne qualité de prédiction comparant à la méthode de BP.
- Contrairement à la back-propagation, la méthode des cartes auto-organisées de Kohonen que nous avons utilisé avec des ' profils' a montré une meilleure performance dans le classement des entrées avant de les utiliser dans la prédiction.
- Les SOM sont dans la pratique flexibles et faciles à implémenter. L'exactitude de la prédiction dépend de la qualité des données d'entrées.
- Les réseaux de neurones artificiels donnent de meilleurs résultats pour les fonctions monotones.

Bibliographie

-
- [1]. L. Mokhnache, “ Application des réseaux de neurones dans le diagnostic et la prédiction des isolations de haute tension ”, Thèse de Doctorat, ENP, Décembre 2003.
 - [2]. G. Dreyfus, ‘LES RÉSEAUX DE NEURONES’, cours de l’École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (ESPCI), Laboratoire d’Électronique, septembre 1998.
 - [3]. A.A.Mazroua, R.Bartnikas, M.M.A.Salama, ‘Discrimination between PD pulse shapes using different neural networks paradigms’, IEEE Trans. on Dielectrics and Electrical Insulation, Vol. 1, N°6, pp: 1119-1131, Dec. 1994.
 - [4]. A.A.Mazroua, R.Bartnikas, M.M.A.Salama, ‘Neural networks systems using the multiplayer perceptron technique for the recognition of PD pulse shapes due to cavities and electrical trees’ IEEE Trans. on Power Delivery, Vol.10, N°1, pp:92-96, Jan. 1995.
 - [5]. K.Engel, D.Peier, ‘Influence of PD fault development on fault type recognition using an artificial neural network’, 9th ISH, Graz, Austria, paper N°5861, Aug. 1995.
 - [6]. T.Kexiong, Z.Deheng, W.Zenyuan, ‘Recognition of partial discharges based on artificial neural network’, 10th ISH, Montréal, Canada, Aug. 1997.
 - [7]. S.Happe, H.G.Kranz, ‘Real-Time partial discharge disturbing pulse suppression with a neural network hardware’, 10th ISH, Montréal, Canada, Aug. 25-29, 1997.
 - [8]. T.Nogami & al. , ‘Gas discrimination method for transformer fault by neural network’, Electrical engineering in Japan, Vol.115, N°1, pp:93-102, 1995.
 - [9]. T.Nogami & al. , ‘Gas discrimination method for transformer fault by neural network’, Electrical engineering in Japan, Vol.115, N°1, pp:93-102, 1995.
 - [10]. L.Mokhnach, A.Boubakeur, N.Nait Said, ‘Application of neural networks in the diagnosis of transformer oil used by SONELGAZ’, Numéro special AJOT-Série B, 4^{eme} CNHT, pp: 51-54, Ghardaia, Algérie, Jan. 2002.
 - [11]. L.Mokhnach, A.Boubakeur, ‘Self-organization neural networks in prediction of insulation thermal ageing’, ISH2003, IEEE-DEIS proceedings, Delft-Netherlands, Aug. 2003.
 - [12]. P.S.Ghosh, S.Chakravorti, N.Chatterjee, ‘ANN with RPROP Algorithm applied to modelling of flashover voltage characteristics of polluted insulators’, 9th ISH, Graz, Austria, Aug. 1995.
 - [13]. E.Ailam, M.Teguar, A.Mekhaldi, ‘Utilisation des réseaux de neurones dans le domaine de la pollution des isolateurs’, Numéro spécial AJOT, Série B, 4eme CNHT, pp: 136-139, Ghardaia, Algérie, Jan.2002.
 - [14]. Zafer Aydogmus, ‘A neural network-based estimation of electric fields along high voltage insulators’, Department of Electrical Education, Firat University Technical Education Faculty, Elazığ, Turkey.
 - [15]. A.S. Farag, ‘Estimation of Polluted Insulators Flashover Time Using Artificial Neural Networks’, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, IEEE, 2010.

- [16]. M.T. Gencoğ lu, M. Cebeci, 'Investigation of pollution flashover on high voltage insulators using artificial neural network', Department of Electrical and Electronics Engineering, Faculty of Engineering, Firat University, Elazig, Turkey, 2008.
- [17]. Ayman H. El-Hag, Ali Naderian Jahromi, Majid Sanaye-Pasand, 'Prediction of leakage current of non-ceramic insulators in early aging period', Electrical Engineering Department, American University of Sharjah, Sharjah, United Arab Emirates, Kinectrics Inc., Transmission & Distribution Technologies, Toronto, Canada, Electrical and Computer Engineering Department, University of Tehran, Iran, 2008.
- [18]. Ahmad S. Ahmada, P.S. Ghosh, S. Shahnawaz Ahmed, Syed Abdul Kader Aljunid, 'Assessment of ESDD on high-voltage insulators using artificial neural network', Dhaka 1000, Bangladesh, 2004.
- [19]. S.M. Gubanski, A. Dernfalk, J. Andersson and H. Hillborg, 'Diagnostic Methods for Outdoor Polymeric Insulators', IEEE Transactions on Dielectrics and Electrical Insulation Vol. 14, No. 5; Oct. 2007.
- [20]. V.T. Kontargyri, A.A. Gialketsi, G.J. Tsekouras, I.F. Gonos, I.A. Stathopoulos, 'Design of an artificial neural network for the estimation of the flashover voltage on insulators', Electric Power Systems Research 77, 2007.
- [21]. C. Volat, F. Meghnefi, M. Farzaneh and H. Ezzaidi, 'Monitoring Leakage Current of Ice-covered Station Post Insulators using Artificial Neural Networks', IEEE Transactions on Dielectrics and Electrical Insulation Vol. 17, No. 2; April 2010.
- [22]. S.Chowdhury, M.S.Naidu, 'Application of artificial neural network in SF6 breakdown studies in non-uniform field gaps', 11th ISH, London, UK, paper 5.204.P5, Aug. 1999.
- [23]. I.Nunes Da Silvan, A.Nunes Da Souza, 'Analysis of high-voltage substations design using artificial neural networks', 11th ISH, London, UK, Aug. 1999.
- [24]. A.Boubakeur, L.Mokhnache, A.L.Nemmour, A.Khelifi, 'Neural network application in the field of HV cable insulation thermal ageing' 3rd Int. Conf. on Quality, Reliability & Maintenance, Professional Publishers, IMecE, pp: 122-126, Oxford, UK, Apr. 2000.
- [25]. L.Mokhnach, A.Boubakeur, "Prediction of the Breakdown Voltage in a Point-Barrier-Plane Air Gap using Neural Networks", CEIDP2001, IEEE- Annual report Conf. on Electrical Insulation and Dielectric Phenomena, pp: 369-372, Ontario, Canada, Oct. 2001.
- [26]. Jeff Heaton, 'Introduction to Neural Networks with Java, 1st Edition', Heaton research Inc., 2005.