

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
 MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
 RECHERCHE SCIENTIFIQUE
 ECOLE NATIONALE POLYTECHNIQUE D'ALGER

Département de Génie Electrique
 Filière Electrotechnique

Projet de Fin d'Etude
 pour l'obtention du diplôme
 D'Ingénieur d'état en Electrotechnique

المدرسة الوطنية المتعددة التقنيات
 BIBLIOTHEQUE المكتبة
 Ecole Nationale Polytechnique

**UTILISATION DES RESEAUX DE
 NEURONES ARTIFICIELS DANS LE
 VIEILLISSEMENT THERMIQUE DE
 L'HUILE DE TRANSFORMATEUR**



المدرسة الوطنية المتعددة التقنيات
 Ecole Nationale Polytechnique

Etudié par :
par :

M^r : Ould Noureddine Mrd El Boukhary
 M^r : Bedja A/Rahim

Proposé et dirigé

M^r : A.Boubakeur



E.N.P. 10, Avenue Hassen-Badi, EL-HARRACH, ALGER

Promotion : Juin 2000

Remerciements

*Nous remercions très chaleureusement Monsieur
A. BOUBAKEUR d'avoir bien voulu nous proposer ce sujet et
de nous avoir guider le long de ce travail.*

*Nous remercions très vivement tous ceux qu'ils nous ont fait
l'honneur de participer à ce jury.*

*Nous remercions Mlle HAMIMI et Mlle MOKHNACHE qui
nous ont guidées durant nos recherches bibliographiques.*

*Nous tenons également à remercier tous ceux qui ont
contribué à la réalisation de ce modeste travail.*

SOMMAIRE

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

Introduction générale	1
Chapitre I : Synthèse bibliographique sur l'huile de transformateur	
I.1. Conduction et claquage dans les diélectriques liquides	2
I.1.1. Introduction	2
I.1.2. Mécanisme électronique	2
I.1.2.1. Dissociation renforcée par le champ	2
I.1.2.2. Production de charges aux interfaces	2
I.1.2.3. Conséquence du transfert de charges	2
I.1.3. Mécanisme de charge avec phase gazeuse	3
I.1.3.1. Mécanisme thermique	3
I.1.3.2. Mécanisme de cavitation	3
I.1.4. Mécanisme de charge par pont	3
I.1.5. Evolution des recharges dans les diélectriques liquides	4
I.2. Les huiles de transformateurs et leurs caractéristiques	4
I.2.1. Définition d'un transformateur	4
I.2.2. Principaux éléments d'un transformateur	5
I.2.3. Rôle des diélectriques liquides dans les transformateurs	5
I.2.4. Les huiles minérales isolantes	6
I.2.4.1. Les huiles de base	6
I.2.4.2. Choix d'un diélectrique liquide	6
I.2.4.3. Caractéristiques d'une huile minérale isolante	7
I.2.4.3.1. Propriétés électriques	7
I.2.4.3.2. Propriétés physiques	7
I.2.4.3.3. Propriétés chimiques	8
I.3. Facteurs influants sur la rigidité diélectrique des liquides	9
I.4. Influence de la température sur les propriétés de l'huile de transformateur	11
I.5. Vieillessement des huiles de transformateurs	14
I.5.1. Vieillessement thermique de l'huile de transformateur	14
I.5.2. Lois du vieillessement des isolants	15
I.6. Influence du vieillessement thermique sur les propriétés de l'huile de transformateur	15
I.7. Remplacement et régénération de l'huile isolante	18
I.8. Conclusion	19
Chapitre II : Les réseaux de neurones artificiels	
II.1. Introduction	20
II.2. Le neurone biologique	21
II.2.1. Structure du neurone	21
II.3. Le neurone formel	22
II.3.1. Modèle mathématique général du neurone	23

II.4. Architecture des réseaux de neurones	25
II.4.1. Réseaux statiques	25
II.4.2. Réseaux dynamiques	25
II.5. Apprentissage des réseaux de neurones	25
II.5.1. Apprentissage supervisé	25
II.5.2. Apprentissage non supervisé	26
II.6. Réseaux de neurones à apprentissage non supervisé	26
II.6.1. Mémoires associatives	26
II.6.1.1. Réseau de Hopfield	27
II.6.1.2. Mémoires associatives bidirectionnelles	27
II.6.1.3 Mémoires associatives linéaires optimisées	28
II.6.2. Réseaux compétitifs	29
II.6.2.1. Couche de Grossberg	29
II.6.2.2. Adaptive Resonance Théory (ART)	29
II.7. Réseaux de neurones à apprentissage supervisé	30
II.7.1. Réseaux de neurones basés sur la décision	30
II.7.2. Adaline , Madaline	32
II.7.3. Réseaux multicouches statiques	33
II.7.3.1. Réseaux à Fonction de Base Linéaire LBF	34
II.7.3.1.1. Algorithme du rétropropagation du gradient	34
II.7.3.1.2. Méthode d'Optimisation Aléatoire ROM	38
II.7.3.1.3. Réseaux LBF et approximation de fonctions	39
II.7.3.1.4. Dimension du réseau LBF	39
II.7.3.2. Réseaux à Fonction de Base Radiale RBF	39
II.7.3.2.1. Architecture et fonctionnement des réseaux RBF	40
II.7.3.2.2. Apprentissage des réseaux RBF généralisés	41
II.7.3.2.3. Réseaux RBF et approximation de fonction	42
II.8. Conclusion	42

Chapitre III : Architecture du réseau élaboré

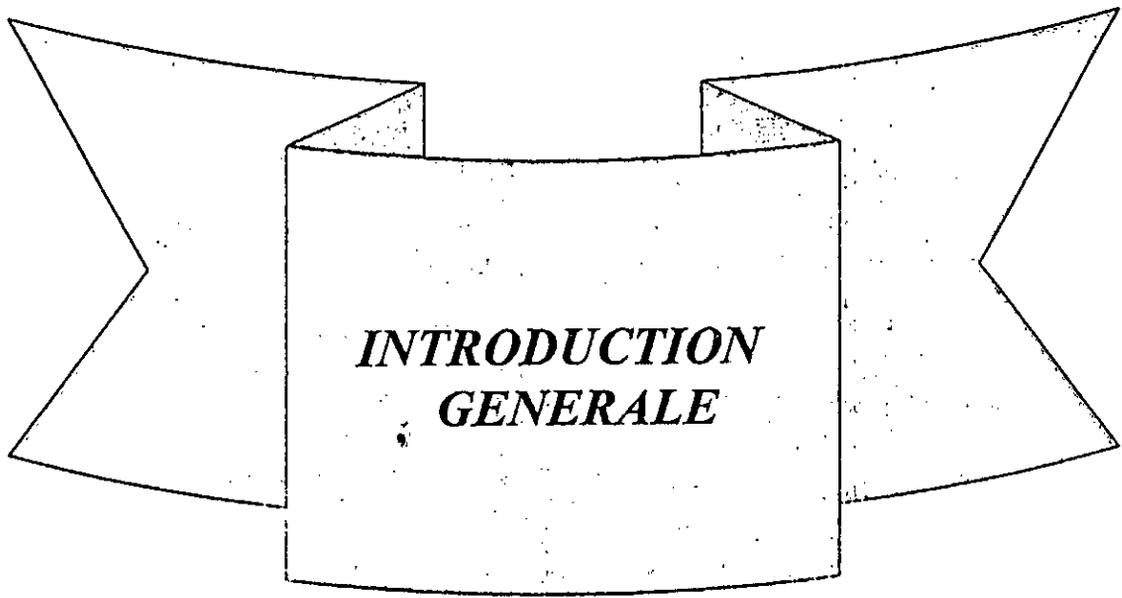
III.1. Introduction	43
III.2. Architecture du réseau élaborée	43
III.3. Choix de l'algorithme d'apprentissage	46
III.4. Application du réseau élaboré dans le vieillissement thermique	47
III.5. Application	57
III.6. Conclusion	56

Chapitre IV : Résultats et interprétations

IV.1. Introduction	56
IV.2. Résultats liés au réseau RBF1	57
IV.2.1. Vieillissement à T = 120°C	57
IV.2.2. Vieillissement à T = 80°C	60
IV.3. Résultats liés au Réseau RBF 2	64
IV.3.1. Vieillissement à T = 120°C	64
IV.3.2. Vieillissement à T = 80°C	67
IV.4. Comparaison entre les deux programmes	70
IV.5. Interprétations	72
IV.5. Conclusion	72

Conclusion générale	73
Annexes	74
Bibliographie	83

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique



INTRODUCTION
GENERALE

INTRODUCTION GENERALE

L'industrie électrique est une grande consommatrice d'huiles isolantes, parmi lesquelles, les huiles minérales occupent une place très prépondérante. Ces dernières sont diversement utilisées, dans l'élaboration de l'isolation du matériel électrique, plus particulièrement celle des transformateurs, des condensateurs, des câbles et des disjoncteurs. Vu les tensions nominales dans ce type de matériel électrique qui peuvent atteindre des centaines de kilo volts, les huiles isolantes sont utilisées à des fins de refroidissement et d'isolation des appareillages électriques en service.

Soumis à différentes contraintes durant l'exploitation, les huiles isolantes se détériorent progressivement sous l'action combinée ou séparée de ces contraintes. Cette dégradation au cours du temps est connue sous le nom de vieillissement.

Les essais de vieillissement ont pour objectif de définir les variations des propriétés de l'huile sous l'effet des contraintes appliquées durant une longue période, de préférence jusqu'à la dégradation définissant la limite de durée de vie. L'objectif est de définir une meilleure maintenance de l'huile de transformateur. Le diagnostic de l'huile en exploitation se fait surtout par « analyse des gaz dissous ».

Au laboratoire de haute tension de l'ENP, une nouvelle approche de l'étude du vieillissement a été introduite. Elle consiste en l'utilisation « des réseaux de neurones artificiels ». Elle nous permettrait de prédire la durée de vie de l'huile à partir des essais de vieillissement de durée relativement courte. Notre étude fait suite à une première étude concernant le vieillissement thermique du PVC utilisé dans les câbles de moyenne tension[1][2].

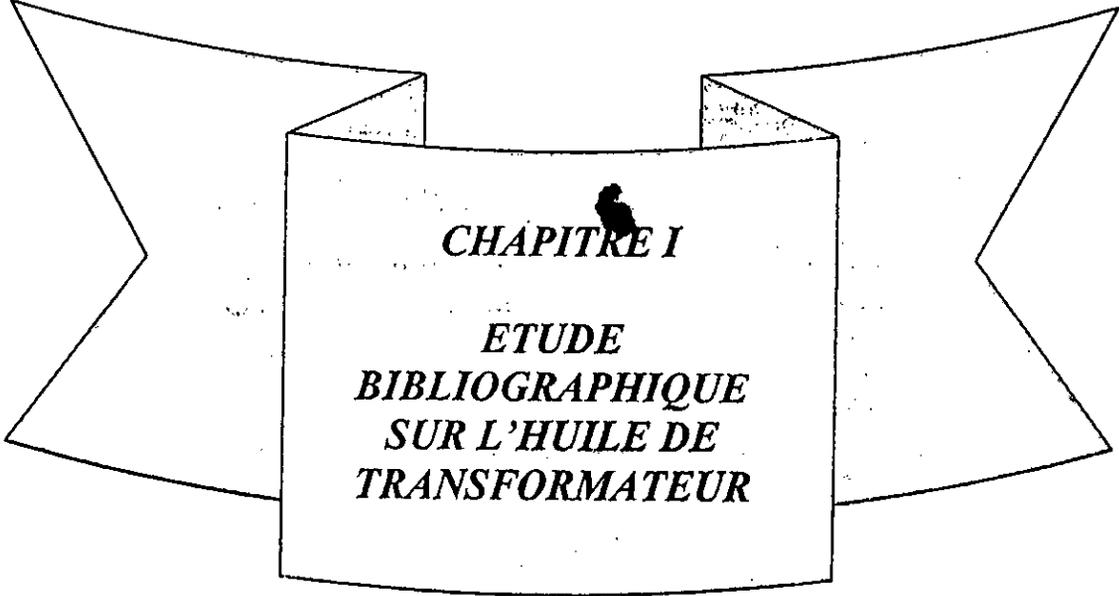
Dans notre étude, nous nous intéressons à l'application des réseaux de neurones artificiels au vieillissement thermique de l'huile de transformateur utilisée par la SONELGAZ et dénommée BORAK 22. Nous avons utilisé les résultats expérimentaux obtenus par d'autres chercheurs du laboratoire[3].

Dans le chapitre I, nous présentons une synthèse bibliographique sur les huiles de transformateurs.

Le chapitre II, présente une synthèse de nos recherches bibliographiques sur les réseaux de neurones artificiels.

Dans le chapitre III, nous détaillons l'architecture du réseau que nous avons élaboré, et dans le chapitre IV, nous présentons tous les résultats obtenus et leurs interprétations.

Finalement, nous terminons notre travail par une conclusion générale.



CHAPITRE I

*ETUDE
BIBLIOGRAPHIQUE
SUR L'HUILE DE
TRANSFORMATEUR*

I.1. CONDUCTION ET CLAQUAGE DANS LES DIÉLECTRIQUES LIQUIDES :

I.1.1. INTRODUCTION :

Les études concernant le phénomène de conduction dans les diélectriques liquides font toujours l'objet de recherche, et cela depuis plus de trente ans, en particulier pour expliquer les mécanismes de création de porteurs de charges et leur nature, ainsi que les processus fondamentaux conduisant au claquage[3].

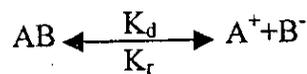
Dans cette partie, on va présenter les différents mécanismes de conduction et de claquage dans les diélectriques liquides purs, c'est à dire qui ne possèdent pas d'ions à l'état libre, et dans les diélectriques liquides contenant des impuretés de différentes nature.

I.1.2. MECANISME ELECTRONIQUE :

La conduction dans les diélectriques liquides, même sous champ intense ($E \geq 1 \text{ MV/cm}$) est presque toujours assurée par des ions de mobilité faible (inférieur à $10^{-3}/\mu\text{s}$), qui préexistent dans les milieux ou qui se créent aux interfaces[3].

I.1.2.1. DISSOCIATION RENFORCEE PAR LE CHAMP :

Il existe une réaction d'équilibre entre les molécules neutres AB d'un liquide et les ions A^+ et B^- , qui s'écrit :



Avec :

K_d : constante de vitesse de dissociation.

K_r : constante de vitesse de recombinaison.

L'application d'un champ électrique au liquide, abaisse l'énergie de liaison de molécules neutres AB. En conséquence de quoi, le nombre des ions dissociés devient supérieur au nombre des ions recombinés [4].

I.1.2.2. PRODUCTION DE CHARGES AUX INTERFACES :

La création d'ions aux électrodes a pour origine des réactions électrochimiques résultant de la mise en contact de l'électrode et du diélectrique. Des échanges électroniques ont lieu lorsque l'ion pénètre dans une zone de 10^{-7} cm d'épaisseur, où règne un champ très intense ($\sim 10^7 \text{ V/cm}$). Cette région est formée par des ions non déchargés fixés à la surface du métal en permanence[4].

I.1.2.3. CONSEQUENCE DU TRANSFERT DE CHARGE :

Le transfert de charge s'accompagne d'une augmentation ou d'une réduction de la résistivité du liquide selon que l'élément qui réagit est un ion ou une espèce neutre[3].

I.1.3. MECANISME DE CLAQUAGE AVEC PHASE GAZEUSE :

Dans les liquides ultra- purs, et en géométrie pointe-plan ,sous tension continue, il a été observé, qu'à partir d'un certain seuil de tension, des bulles gazeuses se forment au voisinage de la pointe, puis sont violemment chassées vers le plan. Cette phase gazeuse résulterait soit de la vaporisation et de la nucléation des bulles, soit de la cavitation[3][5].

I.1.3.1. MECANISME THERMIQUE :

D'après Watson et Sharbaugh , une bulle de vapeur est générée dans le liquide par une injection de courant à partir d'aspérités sur la cathode. Ce courant serait limité par la charge d'espace et a la forme suivante :

$$I=A.V^n \quad (I.1)$$

Avec:

A : est une constante.

V : est la tension appliquée, $1.5 \leq n \leq 2$

Un modèle plus élaboré et tenant compte du régime transitoire a été proposé par KAO. Si l'intensité de courant est suffisante pour ramener la température du liquide à son point d'ébullition , il y a claquage.

I.1.3.2. MECANISME DE CAVITATION :

Le phénomène de cavitation dans un liquide se manifeste lorsque la pression de ce liquide à température constante tombe au dessous de sa vapeur saturante.

D'après Krasucki , les bulles se forment en des points où le champ est élevé (particules solides de très petites dimensions dans le liquide ou attachées aux électrodes).pour qu'il y ait formation d'une cavité, le critère de pression nulle est :

$$P_{em}=P_h + P_s \quad (I.2)$$

P_{em} : Pression électromécanique.

P_h : Pression hydrostatique.

P_s : Pression due à la tension superficielle du liquide.

Un critère analogue a été proposé par Thomas, en supposant que la pression d'électrostriction s'oppose au décollement du liquide.

I.1.4. MECANISME DE CLAQUAGE PAR PONT :

Les impuretés se trouvent dans les isolants liquides sous différentes natures (bulles de gaz, gouttelettes d'eau, particules solides isolantes ou conductrices). La présence de ces impuretés renforce le champ localement, la déformation du champ dépend de plusieurs paramètres comme :

- Forme et dimension des impuretés.
- Intervalle et forme des électrodes.

- Permittivité et conductivité des impuretés.
- Concentration des impuretés entre les électrodes.
- Valeur des charges libres existant à la surface des particules.

L'élévation du champ peut conduire à des claquages partiels dans le cas des impuretés solides conductrices et des impuretés liquides.

Dans le cas des impuretés solides isolantes, celles-ci se polarisent dans le champ électrique et tendent à se déplacer vers les zones où le champ est le plus intense. Pour qu'un pont de particules se forme entre les électrodes, il faut que le temps d'application soit relativement lent. Quand ce pont relie les électrodes, il s'établit un fort courant à travers ce pont qui est plus conducteur que le liquide. On aura ainsi, un échauffement local important qui peut engendrer le claquage.

Dans le cas de l'eau condensée, Les gouttes d'eau sont initialement sphériques et tendent à se déformer sous l'influence du champ électrique. Elles prennent la forme d'ellipsoïdes qui s'allongent dans la direction du champ. Lorsque le volume de la goutte est suffisant, il peut ainsi se former un canal d'eau reliant les électrodes. Ce dernier engendre un fort courant qui provoquerait la vaporisation de l'eau et par suite le claquage dans le canal « vapeur-eau » en cours de formation.

I.1.5. EVOLUTION DES DECHARGES DANS LES DIELECTRIQUES LIQUIDES :

De nombreux travaux ont porté sur la propagation de streamers, dans des systèmes d'électrodes à champ divergent. Il est généralement admis que le claquage est précédé d'une phase de préclaquage pouvant être elle même séparée en deux :

- Une phase de génération caractérisée par un temps t_g dit de génération , Pendant lequel apparaît au voisinage d'une électrode une perturbation révélant La forme d'arborescence.
- Une phase de propagation durant laquelle se développe la perturbation précédemment créée. Cette phase est caractérisée par un temps t_p .

Les streamers sont caractérisés par leur indice de réfraction qui est différent de celui du liquide, ils produisent des courants, et éléments de la lumière.

I.2. LES HUILES DE TRANSFORMATEURS ET LEURS CARACTERISTIQUES :

I.2.1. DEFINITION D'UN TRANSFORMATEUR :

Le transformateur est un appareil électromagnétique statique destiné à transformer un courant alternatif primaire en un autre courant secondaire de même fréquence ayant dans le cas général d'autres caractéristiques, en particulier, une autre tension et un autre courant [6].

I.2.2. PRINCIPAUX ELEMENTS D'UN TRANSFORMATEUR :

Un transformateur se compose des parties principales suivantes :

a. Noyau du transformateur :

On appelle noyau d'un transformateur l'ensemble formant le circuit magnétique avec tout les éléments faisant partie de cet ensemble.

b. Enroulements d'un transformateur :

Les enroulements d'un transformateur doivent satisfaire à plusieurs exigences dont les plus importantes sont :

1. Le régime thermique de l'enroulement doit correspondre aux exigences de la norme, car une température trop élevée réduit la durée de vie du transformateur.
2. L'enroulement doit résister aux contraintes naissant lors d'un court-circuit brusque.
3. L'enroulement doit présenter la rigidité diélectrique nécessaire pour supporter les surtensions.

c. Cuve d'un transformateur dans l'huile :

Les transformateurs dans l'huile sont récemment les plus utilisés, dans lesquels la partie active, c'est à dire le noyau avec les enroulement, sont placés dans une cuve remplie d'huile. En s'échauffant, l'huile commence à circuler à l'intérieure de la cuve et assure ainsi le refroidissement naturel du transformateur.

I.2.3. ROLE DES DIELECTRIQUES LIQUIDES DANS LES TRANSFORMATEURS :

Lors du fonctionnement d'un transformateur, il y a des pertes d'énergie, dont la plus grande partie est localisée dans les enroulements et le circuit magnétique du transformateur[3].

L'échauffement de ces éléments nécessite la mise en œuvre de moyens de refroidissement . c'est la première fonction assignée aux diélectriques liquides : le transfert de chaleur (refroidissement).

La chaleur évacuée est véhiculée par circulation naturelle ou forcée du diélectrique liquide au moyen d'un équipement tel qu'une pompe ou un ventilateur, vers les dispositifs de refroidissement. Le transfert de chaleur de la partie active à l'air a lieu en deux étapes :

- De la partie active à l'huile (à l'intérieur de la cuve).
- De l'huile à l'air.

L'isolement est l'autre fonction essentielle des diélectriques. D'une manière générale, tout transformateur exige : un isolement entre spires, un isolement entre bobines.

I.2.4. LES HUILES MINERALES ISOLANTES :

Les huiles minérales sont issues du raffinage du pétrole brut, elles représentent la fraction lourde de ce dernier, et sont constituées donc de mélanges très complexes de composés appelés « hydrocarbures ».

1.2.4.1. LES HUILES DE BASE :

Les huiles de base sont issues du pétrole brut par distillation et à partir desquelles sont fabriquées les huiles minérales isolantes. Ces huiles de base doivent répondre à trois tendances chimiques de caractères spécifiques et particulières[7][8].

a- La tendance paraffinique :

Elle est représentée par l'ensemble des hydrocarbures saturés à chaîne droite ramifiée ou non, mais non cyclique. Les caractéristiques de cette famille d'huiles sont :

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

b- La tendance naphénique :

Les types de naphènes ayant des hydrocarbures cycliques à liaison saturée . les caractéristiques physiques et chimiques de ces hydrocarbures se traduisent en particulier par :

- Une densité relativement élevée pour une viscosité donnée.
- Des variations assez rapides de viscosité en fonction de la température.
- Un pouvoir solvant relativement élevé.
- Une plus grande volatilité que les huiles paraffiniques.

c- La tendance aromatique :

Ces huiles présentent une densité élevée, un indice de viscosité faible, une très faible résistance à l'oxydation et une facilité d'émulsion avec l'eau.

1.2.4.2. CHOIX D'UN DIELECTRIQUE LIQUIDE :

La production de l'huile isolante est orientée par les tendances des huiles de base ; le choix d'une ou plusieurs huiles de base est à déterminer en fonction :

1. Des caractéristiques physiques et chimiques que l'on désire conférer à l'huile finie (isolante) comme la viscosité, la densité, la stabilité ,....
2. De la nature chimique recherchée (paraffinique, naphénique ou mixte).

1.2.4.3. CARACTERISTIQUES D'UNE HUILE MINERALE ISOLANTE :

Le choix d'une huile isolante est liée à certaines propriétés telles que :
- Les propriétés physiques- Les propriétés chimiques - Les propriétés électriques.

1.2.4.3.1. PROPRIÉTÉS ÉLECTRIQUES :

1- PERMITTIVITÉ :

La permittivité définit la possibilité sous l'action du champ électrique, de libérer des charges. Elle résulte de diverses contributions de déplacements de charges [9].

La permittivité relative « ϵ_r » est définie comme le rapport « ϵ / ϵ_0 » de la permittivité absolue « ϵ » du liquide à celle du vide ($\epsilon_0 = (1/36\pi)10^{-9}$ [F/m]). La permittivité relative des huiles varie peu avec la fréquence.

2- CONDUCTIVITE ET FACTEUR DE PERTES DIELECTRIQUES :

Tout liquide isolant, soumis à une différence de potentiel, donne lieu à un courant de fuite caractérisant la conductivité électrique « σ » du fluide.

En pratique, pour mieux caractériser un diélectrique du point de vue pertes, il est usuel d'utiliser la tangente de l'angle de pertes $\text{tg}(\delta)$. C'est une fonction qui augmente avec la température et dépend de la tension appliquée [10].

3- RIGIDITE DIELECTRIQUE :

La rigidité diélectrique d'une huile est la valeur maximale du champ électrique que l'on peut lui appliquer sans décharge.

Dans l'industrie, elle est maximisée conventionnellement par la tension efficace nécessaire pour provoquer, dans les conditions normalisées, le claquage d'un volume de liquide compris entre deux électrodes dont la forme, distance et nature sont spécifiées [11].

La rigidité diélectrique n'est pas un critère de qualité de fabrication des huiles isolantes, mais un essai conventionnel destiné à déceler leur état plus ou moins grand de pollution physique par l'eau et d'autres matières en suspension et à conclure ou non à l'opportunité de leur faire un traitement avant introduction dans les appareils auxquels, elles sont destinées [12].

1.2.4.3.2. PROPRIETES PHYSIQUES :

1. VISCOSITE :

La viscosité d'un fluide traduit les forces qu'opposent les molécules de ce fluide à une force tendant à les déplacer. C'est donc, la résistance au mouvement que manifestent tous les fluides [13].

La viscosité diminue avec la température et augmente avec la pression. L'expérience montre que, dans le cas d'une huile minérale, la viscosité double quand la pression passe de la pression atmosphérique à 350 bar.

L'indice de viscosité exprime la variation de la viscosité en fonction de la température. Les huiles de transformateur ont un indice de viscosité très bas ; ce qui permet une circulation plus facile, donc, un refroidissement plus efficace.

2. POINT D'ECOULEMENT :

Le point d'écoulement est la température à laquelle les liquides se figent, il nous permet d'évaluer les performances d'une huile à basse température. En particulier, dans les transformateurs avec circulation forcée. Ainsi, la température à laquelle la viscosité permet le fonctionnement des pompes est de l'ordre de -25°C à 35°C pour des points d'écoulement de -45°C [3][10].

3. CONDUCTIVITE THERMIQUE :

Mesurée en Watt par mètre Kelvin, elle exprime le flux thermique s'écoulant, en régime permanent sous l'effet d'un gradient thermique entre deux isothermes du liquide. Elle décroît lorsque la température et la masse volumique augmentent. Pour la plupart des liquides, on utilise la formule de Gragöe[8] :

$$\lambda = \frac{0.117(1-at)}{d_{15}} \quad (1.3)$$

où :

d_{15} : la densité à 15°C.

$a = 5,4.10^{-4}$

4. POINT D'ECLAIR :

Le point d'éclair indique la température de l'huile pour laquelle les vapeurs émises au contact de l'air et approchées d'une flamme prennent feu sans atteindre la combustion totale. Sa détermination renseigne quant à la présence d'hydrocarbures gazeux dissous dans l'huile par suite de défauts électriques ou thermiques[14].

5. LE POINT DE FEU :

C'est le point d'inflammation de l'huile. Il indique la température minimale pour laquelle la combustion des vapeurs d'un liquide sera entretenue[8].

6. LA COULEUR :

La coloration d'une huile neuve est une indication de son degré de raffinage, les huiles très raffinées sont transparentes incolores, une variation de la couleur dans des livraisons d'huile neuve peut indiquer un changement dans l'origine de l'huile ou dans la fabrication, sans que cela soit absolument certain. Dans le cas d'une huile usagée, une coloration foncée indique sa détérioration[12].

1.2.4.3.3. PROPRIETES CHIMIQUES :

1. L'ACIDITE :

L'acidité d'une huile peut avoir des origines diverses, elle est minérale, et peut provenir des opérations de raffinage et organique dite de pétrole[7]. L'acidité d'une huile en service résulte des produits d'oxydation. Pour une huile neuve, l'acidité est de l'ordre de 0,5[mg KOH/g][10].

2. TENEUR EN EAU :

Les propriétés électriques des liquides sont affectées par leur teneur en eau. Cette dernière dépend de la température et de la pression partielle d'eau dans l'atmosphère au dessus du liquide. La teneur à saturation et à température ambiante est de l'ordre de 50 ppm pour les huiles minérales[10].

3. OXYDATION DES HUILES MINERALES ISOLANTES :

On distingue par oxydation, un ensemble de réactions chimiques complexes et lentes au cours desquelles les hydrocarbures réagissent avec l'oxygène. Il se produit tout d'abord des produits de décomposition solubles (aldéhydes, cétones, acides organiques), puis par condensation et polymérisation, des dépôts insolubles[8].

Les conséquences de l'oxydation de l'huile pour les isolations sont :

- Une mauvaise évacuation de la chaleur par suite de l'accroissement de la viscosité et de la présence de dépôts.
- Une diminution des propriétés électriques de l'huile (la conductivité et les pertes diélectriques augmentent).
- L'apparition de l'acidité dans l'huile.

I.3. FACTEURS INFLUANT SUR LA RIGIDITE DIELECTRIQUE DES LIQUIDES :

1. HUMIDITE :

la présence d'un peu d'humidité dans le liquide diminue considérablement la rigidité diélectrique de ce dernier.

La figure suivante illustre bien cet effet sur la rigidité diélectrique à 50Hz de l'huile de transformateur[5] :

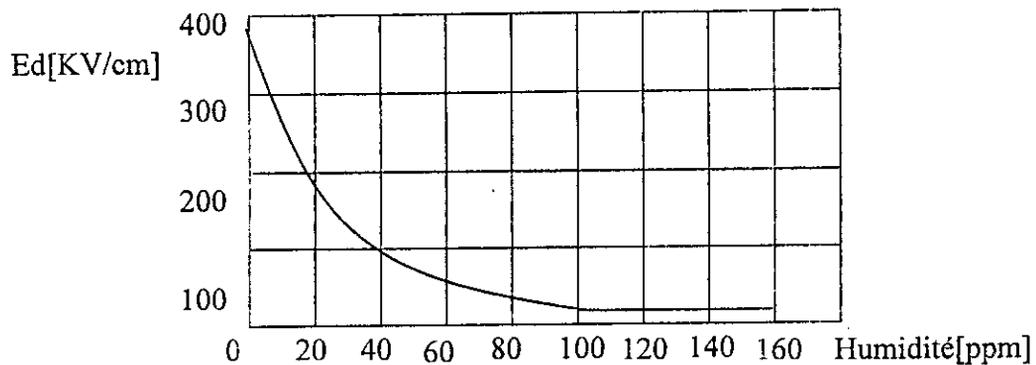


Fig.1.1. Influence de l'humidité sur la rigidité diélectrique de l'huile de transformateur

2. PRESSION HYDROSTATIQUE :

La tension de claquage croît de façon remarquable quand la pression augmente sans indication de saturation en polarité négative, cette croissance est bien moins rapide et tendrait à disparaître en polarité positive[15].

Cette forte dépendance de pression en polarité négative indique que le mécanisme des bulles est dominant, sans pour autant exclure les mécanismes gazeux[16].

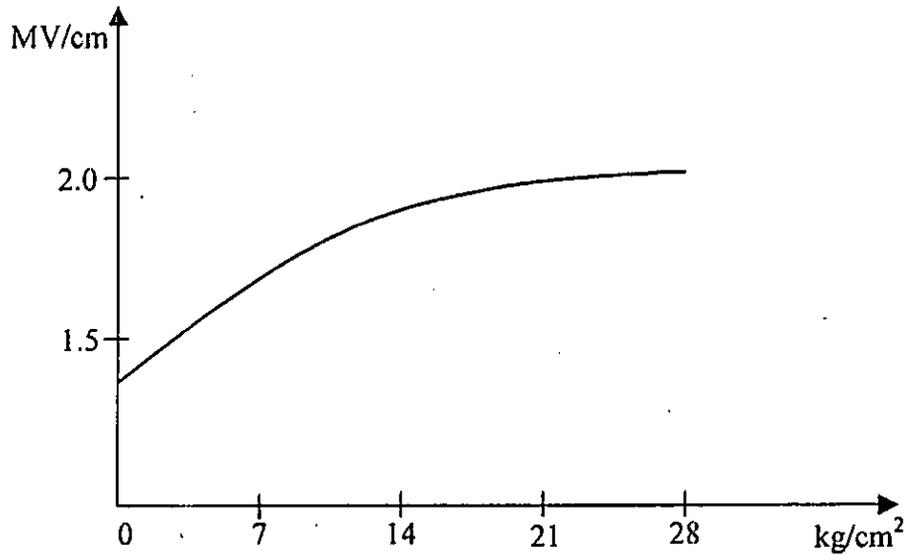


Fig.I.2. Influence de la pression hydrostatique sur la rigidité diélectrique

3. DISTANCE ENTRE ELECTRODES :

En augmentant la distance entre les électrodes, la tension de claquage augmente, mais la rigidité diélectrique du liquide diminue.

Dans le cas de champ électrique non uniforme, lorsque la tension appliquée dépasse un certain seuil, l'effet de couronne apparaît (formation des streamers)[5].

4. TEMPERATURE :

L'influence de la température est double. En augmentant la température, la concentration de gaz et de l'humidité diminue dans le liquide. Cependant, les conditions pour avoir un claquage thermoélectrique s'améliorent[10].

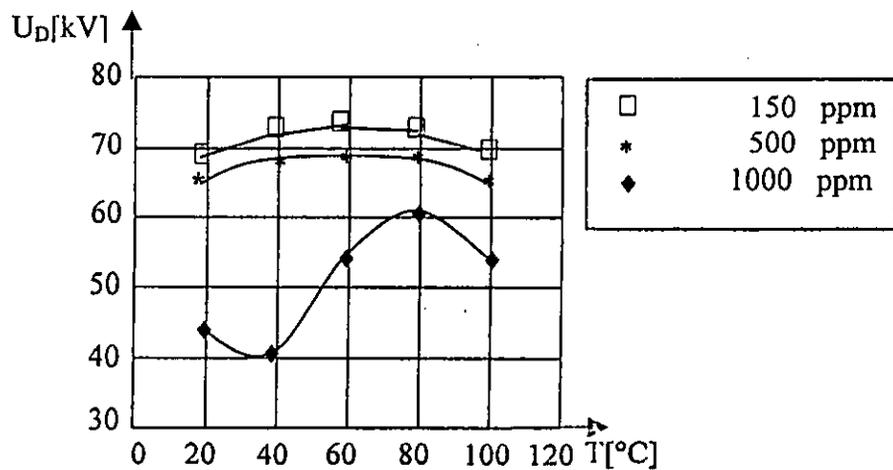


Fig.I.3. La tension de claquage U_D en fonction de la température pour différentes valeurs de l'humidité

5. DUREE D'APPLICATION DE LA TENSION :

La figure ci-dessous présente la caractéristique $U_c(t)$, tension de claquage en fonction du temps, de l'huile de transformateur. Plus la durée d'application augmente, plus l'effet des différentes impuretés se fait ressentir et la tension de claquage diminue[5].

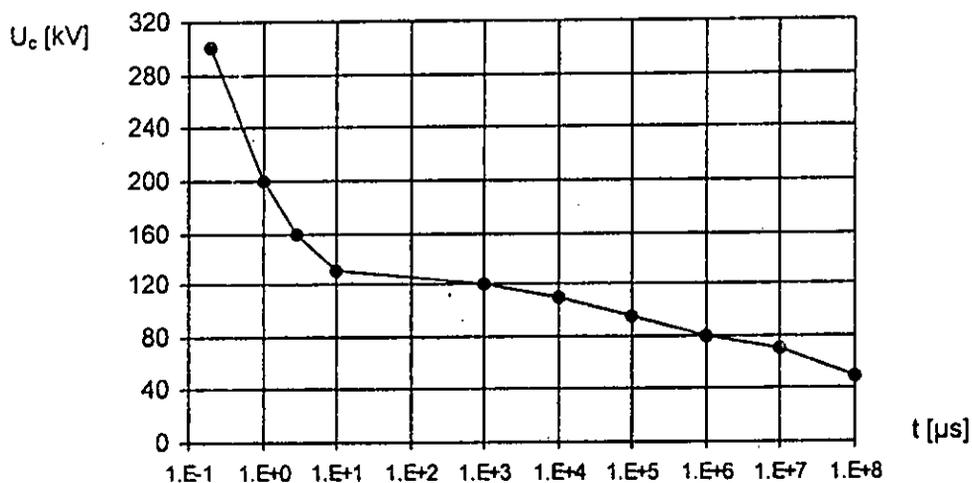


Fig.I.4. Tension de claquage en fonction de la durée d'application de la tension

I.4. INFLUENCE DE LA TEMPERATURE SUR LES PROPRIETES DE L'HUILE DE TRANSFORMATEUR :

Nous présentons certains résultats obtenus dans le cas de l'huile neuve, usagée, régénérée, dans les laboratoires de la SONELGAZ/HAMMA et de l'ENICAB/GUE DE CONSTANTINE [19].

1- PERTES DIELECTRIQUES :

Dans ce cas de type d'huile (neuve, usagée et régénérée), le facteur de pertes diélectriques augmente en exponentiel en fonction de la température (Fig.I.5). Cette évolution est due en même temps à la réduction de la viscosité causant une augmentation de la mobilité des porteurs de charges et à l'augmentation du nombre des porteurs à cause d'une plus grande dissociation thermique.

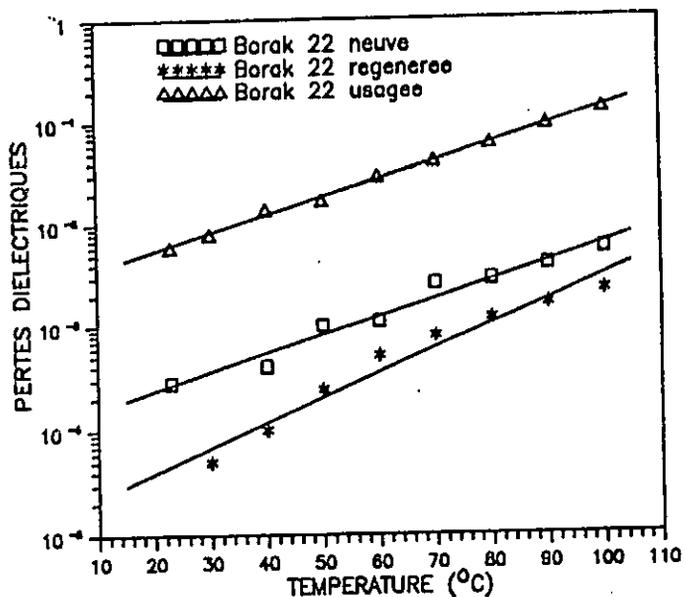


Fig.I.5. Variation de $tg(\delta)$ en fonction de la température

Nous remarquons qu'une régénération de l'huile usagée engendre une nette amélioration du facteur des pertes diélectriques.

2- PERMITTIVITE RELATIVE :

Nous remarquons que la permittivité relative décroît linéairement avec la température dans tous les cas de type d'huile, et les permittivités de l'huile neuve et usagée sont confondues. Ceci permet de dire que la permittivité est un paramètre qui serait affecté par la régénération.

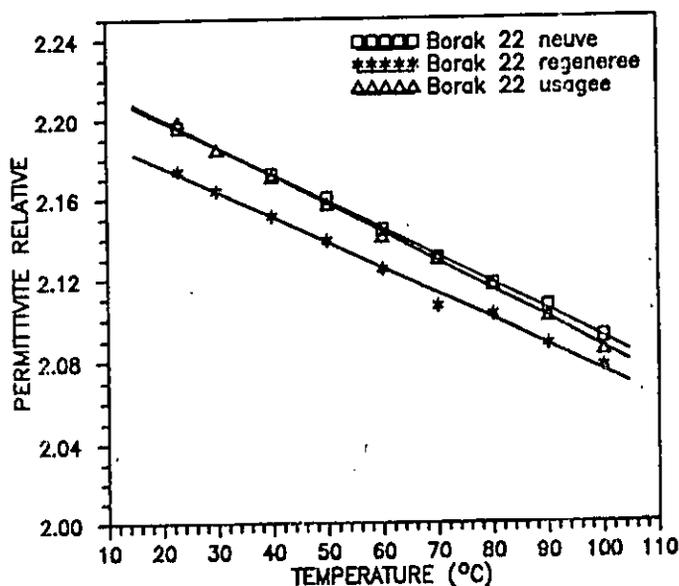


Fig.I.6. La permittivité relative en fonction de la température

3- RESISTIVITE TRANSVERSALE :

La résistivité décroît en fonction de la température selon la loi du type :

$$\text{Log}(\rho) = A.T \text{ (Fig.I.7).}$$

Une forte amélioration de cette propriété est obtenue après la régénération de l'huile.

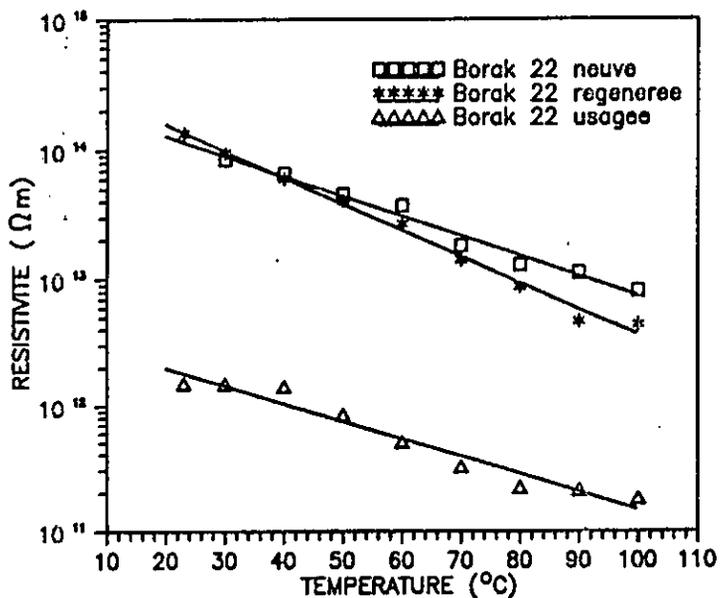


Fig.1.7. Variation de la résistivité en fonction de la température.

4 - VISCOSITE :

Nous remarquons que dans tous les cas de type d'huile, la viscosité diminue quand la température augmente (Fig.I.8), et à haute température, l'huile usagée ne requiert pas sa viscosité après sa régénération.

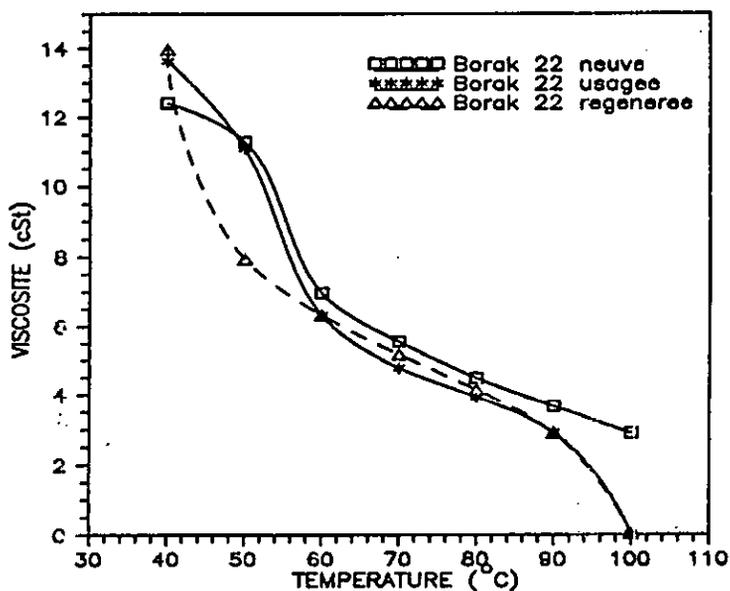


Fig.1.8. Variation de la viscosité en fonction de la température

I.5. VIEILLISSEMENT DES HUILES DE TRANSFORMATEURS :

Tout au long de leur utilisation, les huiles minérales isolantes peuvent être soumises à plusieurs contraintes électriques, thermiques, chimiques...etc. [17]

Sous l'action combinée ou séparée de ces contraintes pendant la mise en service des transformateurs, les huiles isolantes se dégradent au cours du temps, ce qui est connue sous le nom du vieillissement.

On appelle vieillissement tout phénomène se traduisant par une évolution lente et irréversible des propriétés du matériau.

Le terme du vieillissement n'implique pas un mécanisme précis, on est souvent en présence de phénomènes complexes faisant intervenir simultanément plusieurs mécanismes tels que l'oxydation à température modérée, hydrolyse, migration d'adjuvants[18].

I.5.1. VIEILLISSEMENT THERMIQUE DE L'HUILE DE TRANSFORMATEUR :

On convient d'appeler vieillissement thermique tout phénomène traduit une évolution qui ne fait intervenir que[18] :

- Le matériaux.
- L'atmosphère.
- La température.

L'étude du phénomène de vieillissement thermique d'un isolant liquide tel que l'huile de transformateur à pour deux buts principaux[3] :

- Le premier but est d'évaluer la durée de vie de ces isolants sous l'effet des contraintes thermiques.
- Le second but est de trouver une éventuelle corrélation entre le processus de vieillissement et les contraintes qui le provoquent.

I.5.2. LOIS DU VIEILLISSEMENT DES ISOLANTS :

Les premiers travaux ont porté sur le vieillissement des matériaux isolants. En 1930, V.M.Montsinger, après avoir étudié la diminution des caractéristiques mécaniques des papiers isolants imprégnés d'huiles minérales sous l'influence de la température, propose une loi générale pour traduire le vieillissement qui s'exprime par la relation[9] :

$$L(h) = C.exp(-b\theta) \quad (1.4)$$

Avec :

C et b sont des constantes.

L (h) : durée de vie.

θ (°C) : température.

Des études expérimentales sont venues confirmer que, dans un très grand nombre de réactions chimiques et tout au moins en première approximation, les constantes de vitesse k varient avec la température conformément à l'équation d'Arrhenius :

$$k = k_0 \exp\left(-\frac{w}{RT}\right) \quad (1.5)$$

avec :

- k_0 : constante.
- $T(k)$: constante de gaz parfaits.
- $w (J)$: énergie d'activation.

En s'appuyant sur cette équation, W.Dakin (1948) à montré que la durée de vie L d'un matériau isolant ou d'un système d'isolation en fonction de la température T peut se mettre sous la forme simple :

$$\log(L) = A + \frac{B}{T} \quad (1.6)$$

A et B étant des constantes.

L6. INFLUENCE DU VIEILLISSEMENT THERMIQUE SUR LES PROPRIETES DE L'HUILE :

Nous présentons dans ce paragraphe certains résultats, sur l'évolution des propriétés suivantes : le facteur des pertes diélectrique, la permittivité de l'huile de transformateur sous l'influence du vieillissement thermique.

Ces essais sont effectués au DCME de la SONELGAZ/HAMMA (ALGER), sous deux températures différentes : 80°C et 120°C, et la durée de vieillissement est de 1500 heures[3].

L6.1. FACTEUR DE PERTES DIELECTRIQUES :

Les figures (1.9) et (1.10) présentent les variations du facteur de pertes diélectriques $\text{tg}(\delta)$ en fonction du temps de vieillissement [3].

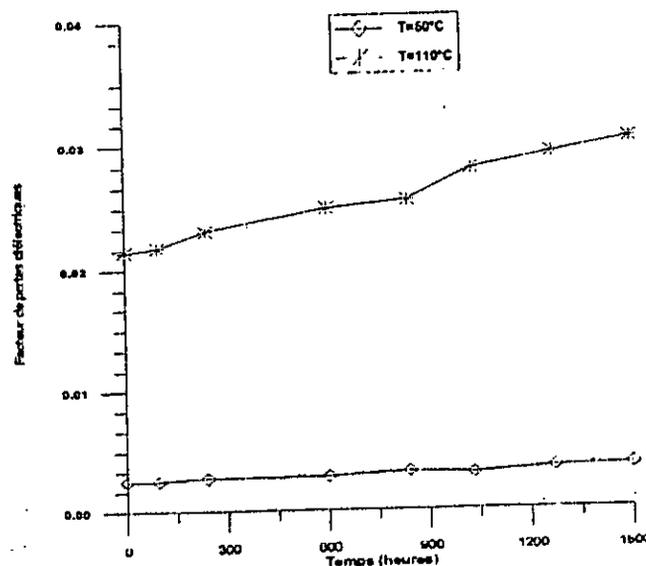


Fig.I.9. Facteur du pertes diélectriques durant le vieillissement à T=80°C

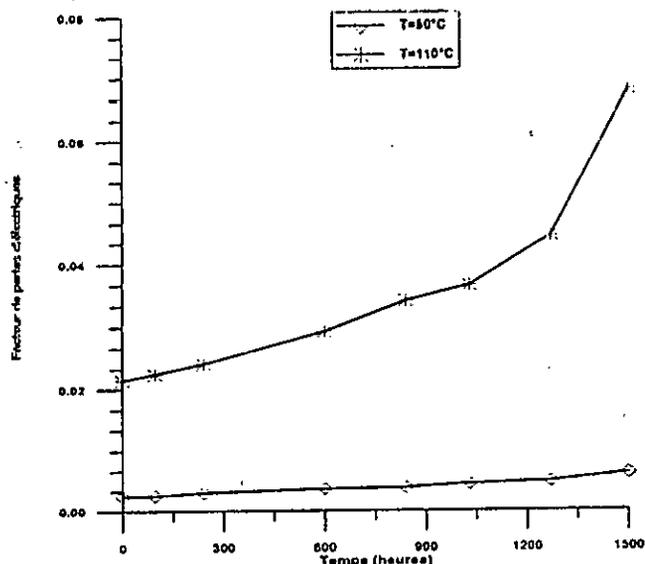


Fig.I.10. facteur des pertes diélectriques durant le vieillissement à T=120°C

D'après ces résultats, on remarque que le facteur de pertes diélectriques $tg(\delta)$ augmente au cours du vieillissement thermique, ceci est dû à l'accroissement du nombre des porteurs de charges par dissociation thermique.

Le facteur de pertes diélectriques augmente plus rapidement dans le cas de la température de 120°C. Cette forte augmentation s'explique par la formation des produits d'oxydation accélérés par des températures élevées au cours du vieillissement.

I.6.2. PERMITTIVITE RELATIVE (ϵ_R) :

Les résultats concernant les variations de la permittivité relative durant le vieillissement thermique à 80°C et à 120°C sont représentés respectivement aux figures(I.11) et (I.12).

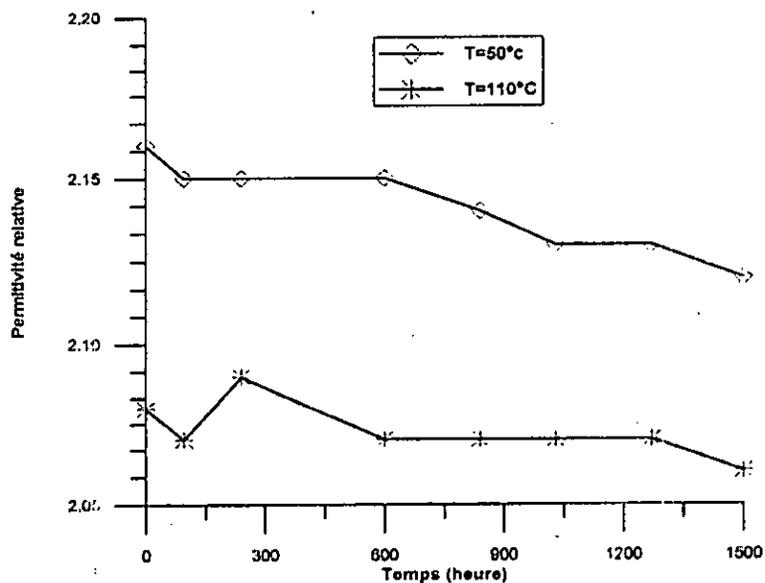


Fig.I.11. Evolution de la permittivité relative (ϵ_r) durant le vieillissement à T=80°C

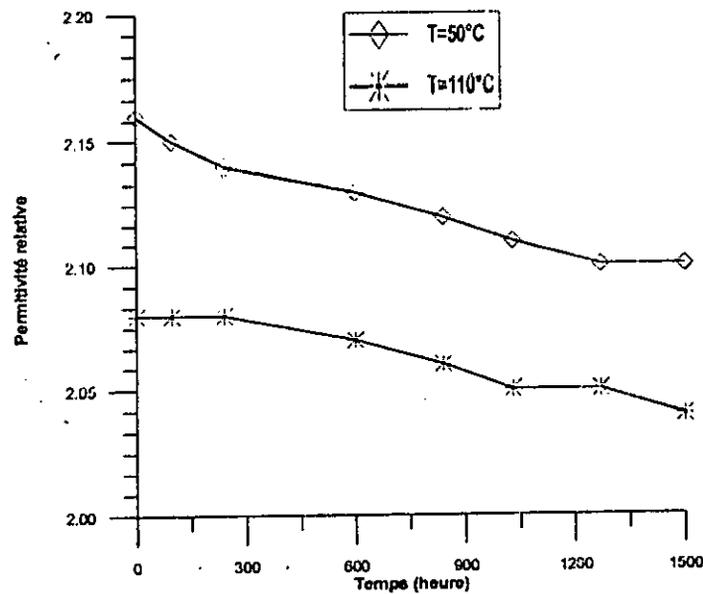


Fig.I.12. Evolution de la permittivité relative (ϵ_r) durant le vieillissement à $T=120^\circ\text{C}$

D'après ces deux courbes, on remarque que la permittivité relative ϵ_r subit une très faible variation durant le vieillissement.

I.6.3. AUTRES PROPRIETES :

Etat de l'huile	Neuve	Viellie à 80°C	Viellie à 120°C
Tension de claquage(kV)	51	51	49
Rigidité diélectrique(kV/cm)	235	253	235
Indice d'acidité(mg KOH/g)	0.026	0.046	0.108
Teneur en eau(p.p.m)	30	31	35
Indice de couleur	1(blanc pur)	1	3(citron pâle)
Viscosité cinématique(CST)	12.54	12.54	12.49

Tableau I.1. Différentes propriétés de l'huile avant et après vieillissement

On remarque d'après le tableau I.1 que :

- la tension de claquage diminue pour $T=120^\circ\text{C}$, cela est due à la vaporisation de l'eau sous température élevée, ce qui n'est pas le cas pour l'huile vieillie à 80°C .
- l'indice d'acidité augmente avec le vieillissement de l'huile, ceci est lié aux conditions suivantes : température de vieillissement, pression, contact avec l'air, présence d'eau. Ces conditions donnent naissance à des produits acides.
- Le vieillissement thermique de l'huile a provoqué une augmentation de la teneur en eau, qui est passée de 30ppm pour l'huile neuve à 31ppm pour l'huile vieillie à 80°C , et à 35ppm pour l'huile vieillie à 120°C .

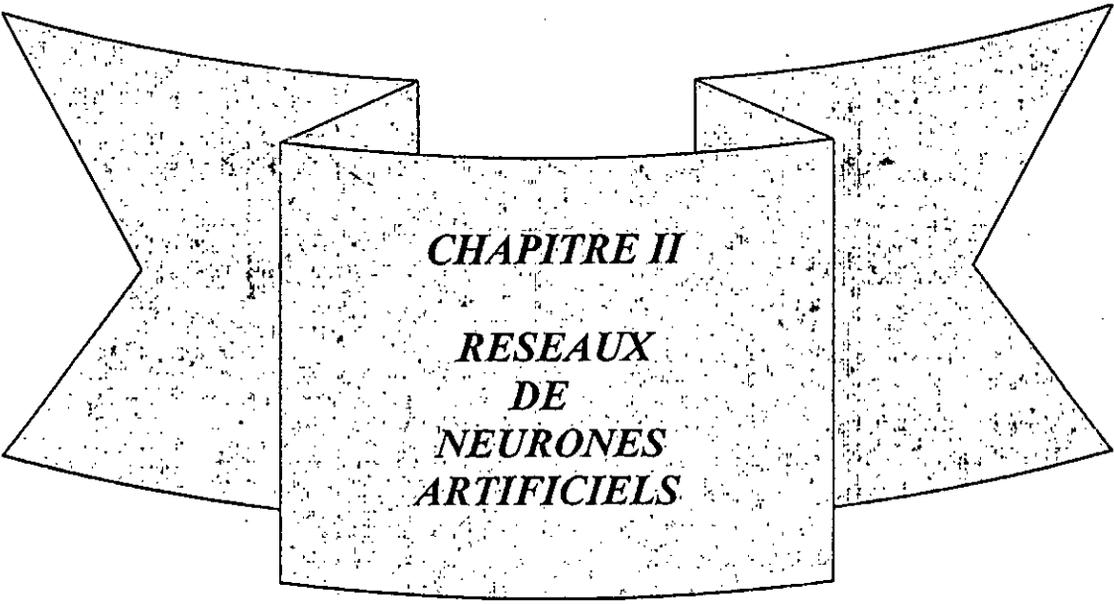
- L'évolution de l'indice de couleur de 1 à 3 pour l'huile vieillie à 120°C est due à l'oxydation de l'huile en service qui a pour conséquence la formation des produits acides.
- La viscosité cinématique est presque constante durant le vieillissement. La pollution de l'huile a pour effet de ralentir l'écoulement de l'huile.

I.7. REMPLACEMENT ET REGENERATION DE L'HUILE ISOLANTE :

Le remplacement de l'huile peut s'imposer quand celle-ci est polluée au point de rendre tout traitement plus coûteux que le remplacement. Une opération de régénération de l'huile est nécessaire lorsque elle est sévèrement polluée par des produits de composition. Si ces derniers sont instables, ils provoquent des dépôts dont l'élimination nécessite un décuivage, et s'ils sont, par contre, solubles, ils contaminent le diélectrique en jouant parfois un rôle de catalyseur dans la poursuite de la décomposition et ne peuvent être éliminés que par une opération chimique ou physico-chimique spécifique assimilable à un nouveau raffinage.

I.8. CONCLUSION :

- Les propriétés électriques de l'huile (pertes diélectriques, permittivité relative, résistivité) dépendent de la température.
- La permittivité relative ne varie pas d'une façon considérable en fonction de la température. Par contre, nous remarquons une importante diminution de la viscosité.
- Quel que soit l'état de pureté de l'huile, le facteur de pertes diélectriques augmente avec la température.
- La régénération de l'huile usagée a permis une amélioration des propriétés électriques étudiées (pertes diélectriques, résistivité, permittivité relative).
- En ce qui concerne le vieillissement thermique de l'huile, les résultats confirment une variation des caractéristiques électriques et quelques caractéristiques physico-chimiques de l'huile vieillie, et surtout pour celle vieillie à la température la plus élevée. Ces variations se traduisent notamment par l'élévation de $\text{tg}(\delta)$, l'indice d'acidité, et un changement de couleur.



CHAPITRE II
RESEAUX
DE
NEURONES
ARTIFICIELS

II.1. INTRODUCTION :

Malgré la constante augmentation des calculateurs, malgré les approches théoriques de plus en plus sophistiquées, un certain nombre de tâches résistent encore aux algorithmes et aux méthodes classiques de traitement des signaux et des données. Ces tâches relèvent typiquement du traitement, en temps réel, de très grand nombre de données souvent multidimensionnelles et arriva à des cadences élevées. Le grand nombre de données, leur variabilité, le fait qu'elles ne répondent pas à des modèles physiques connus nous laissent souvent démunis devant des tâches de caractérisation, de reconnaissance et de prise de décision.

Or, toutes ces opérations se trouvent réalisées de manière naturelle chez les êtres vivants. Un visage est reconnu en quelques dixièmes de seconde, une voix est reconnue au milieu du bruit ambiant et le discours est perçu. Le contrôle simultané d'une centaine de muscles en fonction de la position et de la survenue d'un obstacle dans le champ visuel est réalisé en quelques dixièmes de secondes. Pour une machine, réaliser de telles tâches serait une véritable prouesse, et nous, nous les réalisons sans même y penser ! c'est là que réside l'attrait des « neurones artificiels ».

En 1943, Mc Culloch et Pitts inventèrent le premier neurone formel de l'histoire, un produit scalaire entre un vecteur d'entrée et un vecteur poids, suivi d'un élément à seuil répondant 0 ou 1 selon que le résultat du produit est supérieur ou non au seuil. Ainsi après la publication de la théorie du psychanalyste J.Hebb sur la plasticité synaptique, Rosenblatt inventa le « perceptron », identique au neurone formel, mais doué de la faculté d'apprentissage[20].

Vinrent ensuite Minsky et Papert en 1969[21], qui démontrèrent qu'avec un perceptron à une couche de neurones, on ne pouvait séparer que des données linéairement séparables et que pour des opérations plus générales de classification, il fallait mettre plusieurs couches. Il y eut un « trou » dans les années 70 où rien ne se passa, car les chercheurs et les investisseurs se désintéressèrent des réseaux de neurones pour se tourner vers l'approche symbolique de l'intelligence artificielle, qui semblait beaucoup plus prometteuse.

Le renouveau actuel des réseaux de neurones est dû à des contributions originales, comme celle de Hopfield en 1982, qui en montrant l'analogie des réseaux de neurones avec certains systèmes physiques, a permis de leur appliquer un formalisme riche et bien maîtrisé. Récemment, en 1985, des nouveaux modèles mathématiques ont permis de dépasser les limites du perceptron[22].

Les réseaux de neurones, avec leurs aptitudes en classification, mémorisation, filtrage et approximation, sont devenus un moyen très efficace et ont conquis plusieurs domaines. Nous citons le contrôle et modélisation des systèmes, le traitement de signal et notamment de l'image.

II.2. LE NEURONE BIOLOGIQUE :

Les cellules nerveuses, appelées neurones, sont les éléments de base du système nerveux central. Celui-ci en posséderait environ cents milliards. Les neurones possèdent de nombreux points communs dans leur organisation générale et leur système biochimique avec les autres cellules. Ils présentent cependant des caractéristiques qui leur sont propres et se retrouvent au niveau des cinq fonctions spécialisées qu'ils assurent[22] :

- Recevoir des signaux en provenance de neurones voisins.
- Intégrer ces signaux.
- Engendrer un flux nerveux.
- Le conduire.
- Le transmettre à un autre neurone capable de le recevoir.

II.2.1. STRUCTURE DES NEURONES :

Un neurone est une cellule constituée principalement de trois parties (Fig.II.1) qui, vis-à-vis des transferts d'information, ont un rôle fonctionnel bien défini ; ce sont les dendrites, le soma et l'axone[20][22].

- **LES DENDRITES :**

Chaque neurone possède une « chevelure » de dendrites. Celles-ci sont de fines extensions tubulaires, de quelques dixièmes de microns de diamètre et d'une longueur de quelques dizaines de microns. Elles reçoivent l'information au niveau de points de contact avec les autres neurones. Ces points de contact sont appelés « synapses ». L'information est acheminée vers le corps cellulaire ou soma.

- **LE SOMA :**

Outre son rôle concernant le métabolisme de la cellule, le soma recueille et concentre l'ensemble des informations reçues par les dendrites et en effectue une sommation dite « spatio-temporelle », si le potentiel somatique dépasse un certain seuil, il y a émission d'un potentiel d'action. Ce signal, très bref (1ms), correspond à une oscillation de relaxation du potentiel membranaire du neurone, il est transmis sans atténuation le long de l'axone.

- **L'AXONE :**

L'axone qui est à proprement parler la fibre nerveuse, sert de moyen de transport pour les signaux émis par le neurone. Il se distingue des dendrites par sa forme et par les propriétés de sa membrane externe. En effet, il est généralement plus long (sa longueur varie d'un millimètre à plus d'un mètre) que les dendrites, et se ramifie à son extrémité, là où il communique avec d'autres neurones, alors que les ramifications des dendrites se produisent plutôt près du corps cellulaire.

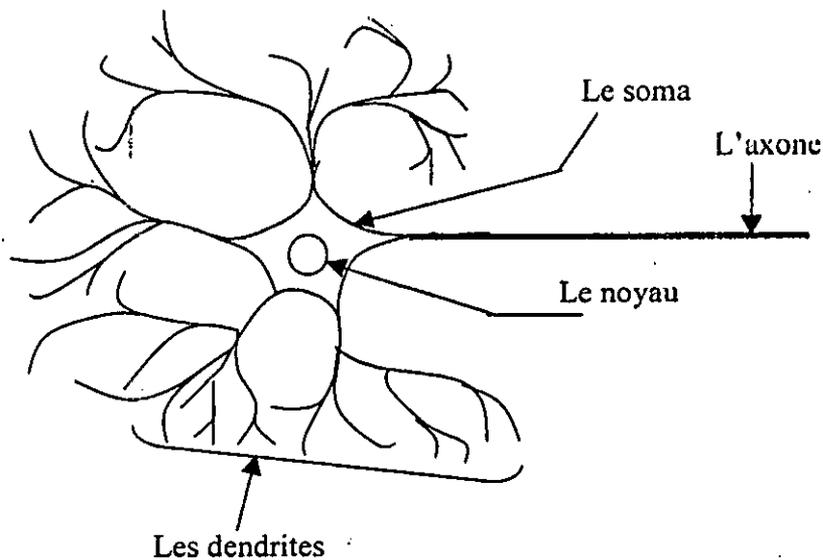


Fig.II.1. Le neurone biologique.

II.3. LE NEURONE FORMEL :

Le premier neurone formel a été modélisé par Mc Culloch et Pitts, en s'inspirant de leurs travaux sur les neurones biologiques, le modèle proposé (Fig.II.2) est le 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 dont la valeur est celle de son activation. Si le neurone n'est pas activé, il ne transmet rien[22]. Le fonctionnement du neurone formel est alors :

$$s = f\left(\sum_1^n w_i e_i\right) \quad (\text{II.1})$$

tel que :

e_i : entrées du neurones formel.

s : sortie

w_i : paramètres de pondération.

$$f : \text{fonction de seuillage} \quad \begin{cases} f(x) = 1 & \text{si } x > \beta \\ f(x) = 0 & \text{si } x < \beta \end{cases}$$

x : la somme pondérée des entrées du neurone formel.

β : le seuil du neurone formel.

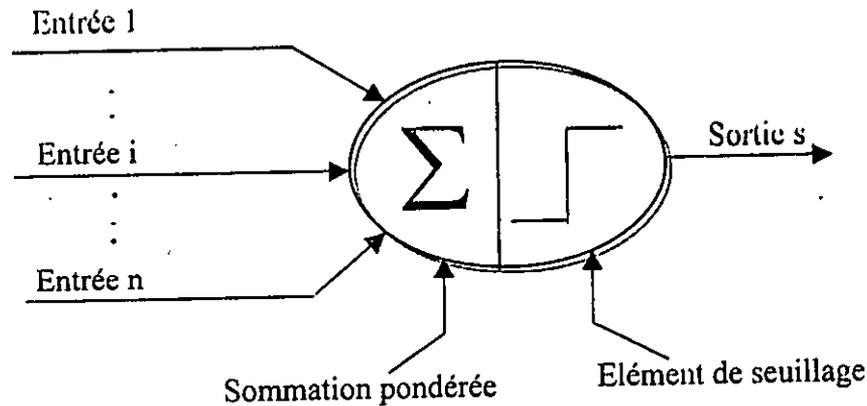


Fig.II.2. le modèle de Mc Culloch et Pitts.

II.3.1. MODELE MATHEMATIQUE GENERAL DU NEURONE :

II.3.1.1. FONCTION DISCRIMINANT :

Cette fonction, appelée aussi fonction de base, 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 utilisée.

Afin d'améliorer les performances des réseaux, d'autres formes de fonctions discriminant non linéaires ont été élaborées. Nous citons ci-dessous les plus utilisées :

a) Fonction de Base Linéaire LBF :

Cette fonction fait la sommation pondérée des entrées vers le neurone, sa forme est en générale définie par :

$$\Phi(z, w) = z^T \cdot w \quad (\text{II.2})$$

Où w représente la matrice des poids et z les entrées qui viennent de l'extérieur ou de la part des autres neurones vers le neurone en question.

b) Fonction de Base Radiale RBF :

La forme de cette fonction est définie par :

$$\Phi(\lambda, w) = -\frac{\|x - w\|^2}{2} \quad (\text{II.3})$$

Où les poids w sont présentés comme étant les centroïdes de chaque classe de l'espace d'entrée. le discriminant calcule le rayon entre les entrées x et le centre de chaque classe.

I.3.1.2. FONCTION DYNAMIQUE LINEAIRE :

Cette fonction détermine la dynamique du réseau. Dans le cas générale, elle est régie par l'équation différentielle de premier ordre suivante :

$$\alpha_0 \dot{u}_i(t) + \alpha_1 u_i(t) = v_i(t) \quad (\text{II.4})$$

Où u_i représente l'activité du $i^{\text{ème}}$ neurone. Et v_i représente l'entrée du système dynamique décrit par l'équation (II.4).

I.3.1.3. FONCTION NON DYNAMIQUE NON LINEAIRE :

C'est la fonction dite d'activation. Elle a pour objectif de rendre l'activité du neurone bornée. La première fonction qui a été proposée est la fonction seuil, mais elle délivre une sortie binaire qui n'est mathématiquement pas adaptée à certaines opérations notamment la différentiation qui est nécessaire pour l'amélioration des paramètres internes du réseau. Pour cela, la fonction d'activation doit être continue, dérivable et monotone. Parmi ces types de fonctions on retrouve la sigmoïde ou la tangente hyperbolique ou la log-sigmoïde.

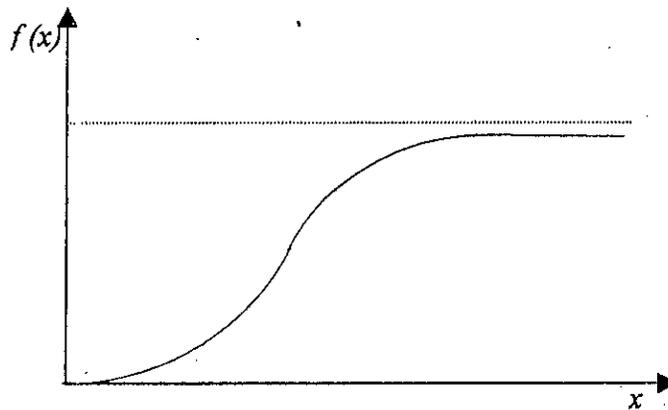


Fig.II.3. Allure de la fonction sigmoïde.

Il existe une autre fonction d'activation qui n'est pas monotone, c'est la fonction gaussienne (Fig.II.4) Elle se distingue des autres, par sa non linéarité significative et son comportement impulsif. C'est en effet, ce comportement local qui la rend très efficace.

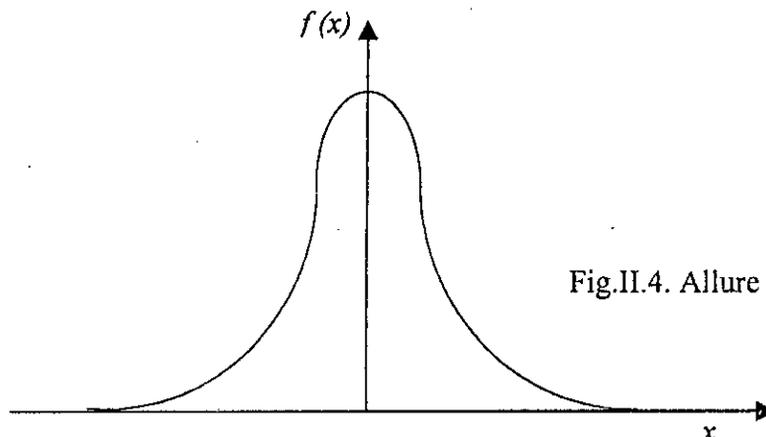


Fig.II.4. Allure de la fonction gaussienne

II.4. ARCHITECTURE DES RESEAUX DE NEURONES :

II.4.1. RESEAUX STATIQUES :

Dans ce type de réseaux, chaque neurone d'une couche reçoit ses entrées à partir des neurones de la couche précédente ou de l'entrée du réseau, sans qu'il y ait retour d'information « Feed-back ». Ces réseaux sont utilisés dans la classification et l'approximation des fonction[23].

II.4.2. RESEAUX DYNAMIQUES :

L'introduction de Feed-back entre les neurones rend les réseaux dynamiques (Fig.II.5). Cela pour les exploiter en identification et en commande des systèmes dynamiques[24].

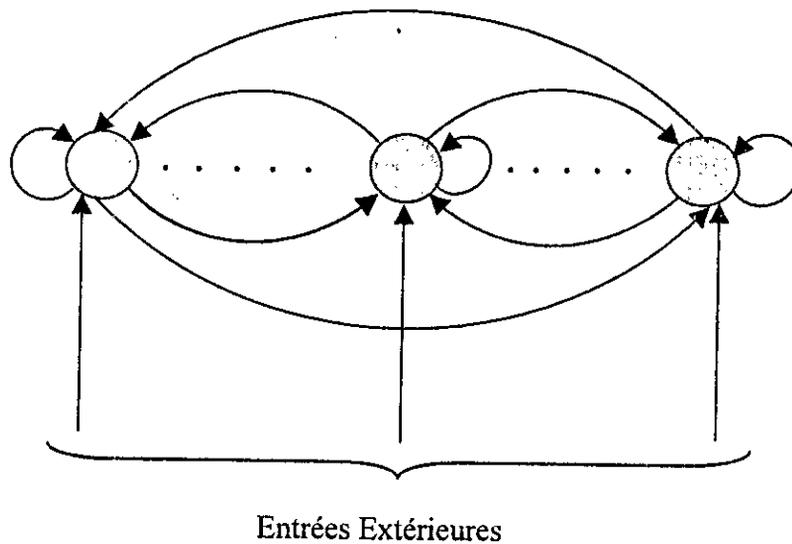


Fig.II.5. Architecture générale d'un réseau dynamique.

II.5. APPRENTISSAGE DES RESEAUX DE NEURONES :

L'apprentissage d'un réseau de neurones est défini comme étant n'importe quel changement opéré dans la mémoire du réseau, cela a comme objectif l'amélioration des performances futures du réseau, sur la base d'une connaissance acquise au fur et à mesure des expériences passées. Il existe deux types d'apprentissage, suivant la tâche pour laquelle ce réseau est utilisé : l'apprentissage supervisé et l'apprentissage non supervisé.

II.5.1. APPRENTISSAGE SUPERVISE :

Il se fait en présence d'un superviseur qui dirige le comportement du réseau en lui présentant les couples d'entrées et leurs sorties désirées. Cet apprentissage se fait par l'intermédiaire d'un critère à optimiser définissant la performance du réseau à chaque étape.

II.5.2. APPRENTISSAGE NON SUPERVISE :

Cet apprentissage nécessite seulement la présence des entrées sans l'intervention d'un superviseur. L'entraînement se fait sur la base d'informations locales existant au niveau des neurones.

II.6. RESEAUX DE NEURONES A APPRENTISSAGE NON SUPERVISE :

Ces réseaux traitent des échantillons en entrée sans avoir d'informations sur les sorties que ceux ci doivent générer. L'objectif de ces réseaux n'est pas, en effet, de faire une approximation. Le rôle de ces réseaux est, en général, basé sur la mémorisation et la classification de leurs entrées.

II.6.1. MEMOIRES ASSOCIATIVES :

La mémorisation est une fonction importante dans tout système de traitement d'information.

L'idée de base des mémoires associatives tourne au tour des concepts mathématiques classiques en traitement de signal. L'introduction des réseaux de neurones dans ce domaine vise à rendre les opérations plus souples, plus efficaces et plus simple à implémenter.

La mise en mémoires des informations à l'intérieur de ces réseaux se fait grâce à un apprentissage non supervisé.

Pendant cette étape appelée Encoding Step, le réseau calcule les poids synaptiques w , représentant un codage de l'information à stocker. Après cela, des exemples parfois incomplets ou modifiés, sont présentés au réseau. Grâce à la présentation interne déduite de l'apprentissage, les neurones calculent leurs activations. A la sortie, le réseau doit se rappeler de ce qu'il a appris. Cette étape est appelée Recal Step[25].

Suivant le mode de fonctionnement, on distingue deux types de mémoires associatives :

a) Mémoire auto-associatives :

Elle sont destinées à stocker des vecteurs prototypes

$$P^k = p_1^k, p_2^k, \dots, p_n^k$$

tel que $k = 1, \dots, M$ représente les différentes classes.

b) Mémoires hétéro-associatives :

Elles stockent les couples prototypes (X^k, Y^k) , où :

$$Y^k = y_1^k, y_2^k, \dots, y_n^k \quad \text{avec } k = 1, \dots, M$$

A chaque vecteur X^k en entrée, cette mémoire associe le vecteur Y^k en sortie et vice versa.

II.6.1.1. RESEAU DE HOPFIELD :

Ce réseau est une mémoire auto-associative constitué par une couche de neurones entièrement interconnectés (Fig.II.6).

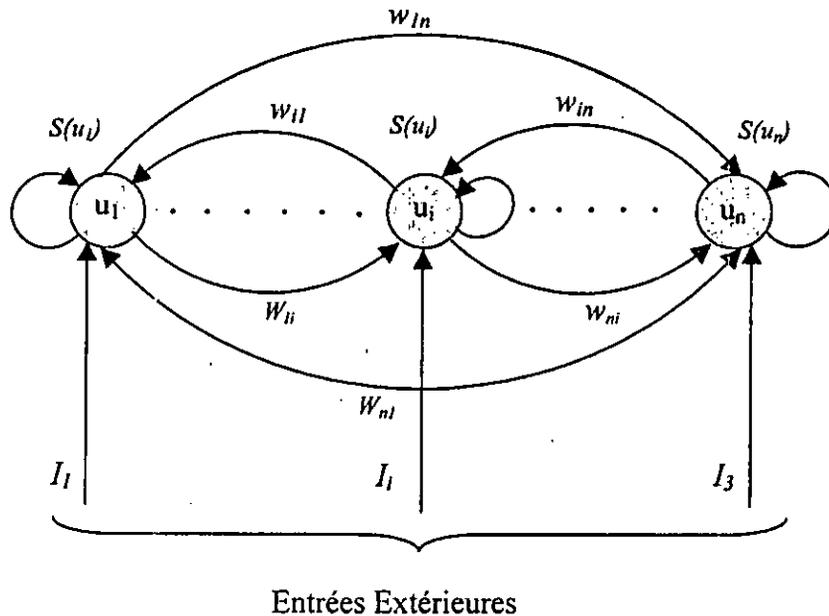


Fig.II.6. Architecture du réseau de Hopfield.

L'originalité du réseau de Hopfield réside notamment dans son modèle électrique très connu, ce modèle traduit le comportement dynamique intégrateur des neurones, celui-ci est défini par le système d'équations suivant

$$\dot{u}_i = \sum_{j=1}^n w_{ij} S_j(u_j) - b_i u_i + I_i \quad i = 1, \dots, n \quad (\text{II.5})$$

Où I_i représentent les entrées extérieures, w_{ij} des paramètres définissant les forces des liens synaptiques représentées par les conductances de ces connexions, b_i représente la conductance équivalente à l'entrée du neurone, et $S_j(u_j)$ représente la sortie du neurone[26].

La sortie du neurone est, comme son entrée, binaire. La fonction de sortie est donc la fonction seuil définie par l'équation (II.5).

II.6.1.2. MEMOIRES ASSOCIATIVES BIDIRECTIONNELLES (BAM) :

Elles représentent une extension du réseau de Hopfield vers une mémoire hétéro-associative, ces réseaux sont composés de deux couches de neurones F_x et F_y (Fig.II.7). L'information à l'intérieur de ce réseau peut circuler dans les deux sens (de F_x vers F_y ou de F_y vers F_x)[25].

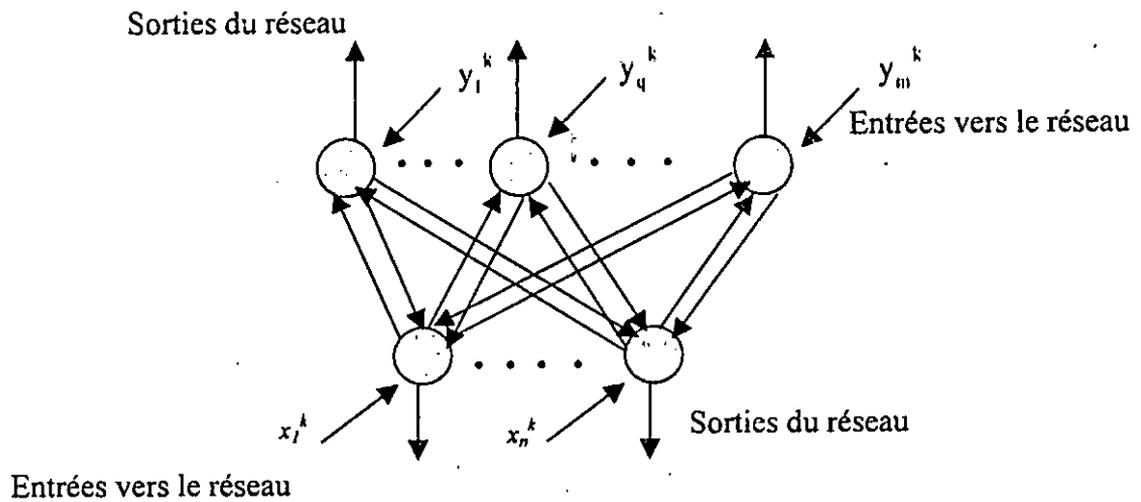


Fig.II.7. Architecture d'une Mémoire Associative Bidirectionnelle

En mémorisant les couples d'exemples (X^k, Y^k) , $k=1, \dots, M$, ce réseau associe au vecteurs $X^k = [x_1^k, x_2^k, \dots, x_n^k]$ les vecteurs $Y^k = [y_1^k, y_2^k, \dots, y_m^k]$ et vice versa.

II.6.1.3. MEMOIRE ASSOCIATIVES LINEAIRES OPTIMISEES (OLAM) :

Le réseau de Hopfield ne garantit un bon fonctionnement que si les exemples sont mutuellement orthogonaux, ce qui n'est pas toujours vérifié en pratique.

Les mémoires associatives optimisées constituent une solution pour les problème d'orthogonalité, ces réseaux utilisent la règle de projection pour leur apprentissage, mais leur architecture est la même que de celle de Hopfield[23].

En injectant une entrée égale à un des prototypes stockés ou proche de lui, on doit retrouver ce prototype en sortie. Ceci revient à résoudre l'équation suivante :

$$Wx^k = x^k \quad ; \quad k=1, \dots, M \quad (II.6)$$

une solution de cette équation est :

$$W = x \cdot x^T \quad (II.7)$$

Où x est une matrice regroupant tout les vecteurs prototypes à mémoriser et x^T représente la pseudo-inverse de cette matrice.

II.6.2. RESEAUX COMPETITIFS :

Ces réseaux s'adaptent au fur et à mesure au cours de leur utilisation, contrairement aux réseaux que nous avons étudiés qui sont à poids fixes. Ces réseaux contiennent, en plus de la couche d'entrée, une couche compétitive, où les neurones entrent en compétition entre eux, afin de déterminer lequel est le plus représentatif de la classe de l'espace d'entrée.

Nous présentons maintenant les modèles des principaux réseaux compétitifs :

II.6.2.1. COUCHE DE GROSSBERG :

Elle est constituée de neurones entièrement interconnectés, sa dynamique est régie par l'équation différentielle suivante[27] :

$$\dot{u}_i = -\mu \cdot u_i + \delta \sum_{j=1}^n S(u_j) w_{ij} + I_i \quad i = 1, \dots, n \quad (\text{II.8})$$

Où :

- u_i : représente l'activité du i^{eme} neurone.
- I_i : est l'entrée extérieure vers ce neurone.
- $S(u_j)$: représente la sortie du neurone.
- w_{ij} : les poids synaptiques.

μ et δ sont des constantes positives.

lors de la présentation d'une entrée X^k au réseau, la dynamique de l'équation (II.8) fait que le neurone de la classe à laquelle cet exemple est le plus proche est activé au maximum. A l'équilibre c'est le seul qui reste activé et le seul à se voir réadapter ses poids. L'apprentissage se fait de la manière suivante[27] :

$$\dot{w}_{ij} = S(u_i) [-\alpha w_{ij} + \beta \cdot S(u_j)] \quad (\text{II.9})$$

Où :

- α : est un paramètre contrôlant le régime transitoire.
- β : est le taux d'apprentissage.

II.6.2.2. ADAPTIVE RESONANCE THEORY (ART) :

Le fonctionnement de ce réseau constitue une analogie avec la résonance physique. En effet, à partir d'un faible signal, l'information fait des « aller-retour » entre la couche d'entrée et celle de sortie. Constituant des revibrations, qui continuent jusqu'à ce que celle-ci entre en résonance. Soit en adhérant à une classe déjà existante, soit en créant une nouvelle classe, suivant la finesse de la classification .

Ce réseau est constitué de deux couches : une couche d'entrée ordinaire et une couche de sortie appelée Maxnet, dont les neurones sont entièrement interconnectés (FigII.8).

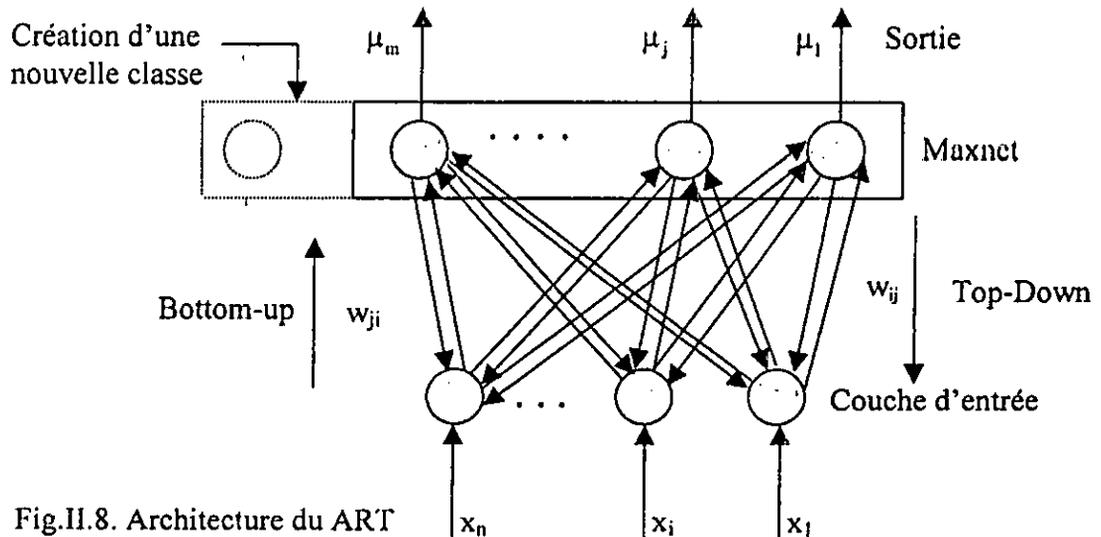


Fig.II.8. Architecture du ART

Chaque neurone de la couche de sortie représente une classe. Lorsque une entrée x est présentée au réseau, celle-ci est transmise vers la couche supérieure en passant par des liens synaptiques allant dans le sens direct appelés Bottom-Up « Bas-Haut ». Chaque neurone du Maxnet calcule son activation. Le neurone ayant ses poids les plus corrélés avec l'exemple en entrée, présente le plus grand résultat en sortie, celui-ci sera candidat pour être désigné comme étant la classe à laquelle cet exemple appartiendrait. Avant de prendre cette décision, l'exemple x est renvoyé vers la couche d'entrée, en passant par les liens synaptiques qui vont dans le sens inverse appelés Top-Down « Haut-Bas ». Là, un test, dit de vigilance, est appliqué afin d'estimer à quel point cet exemple ressemble à ceux appartenant à la classe gagnante. Si celui-ci réussit le test, l'exemple est accepté. Sinon le neurone gagnant est disqualifié et une autre classe est recherchée[28].

II.7. RESEAUX DE NEURONES A APPRENTISSAGE SUPERVISE :

L'apprentissage supervisé a pour rôle de ramener le réseau vers le comportement désiré imposé par un superviseur, ce qui se fait par la recherche du vecteur synaptique localement optimal w^* .

Il existe deux approches qui peuvent être faites pour ce type d'apprentissage. la première est basée seulement sur la validité de la décision du réseau. Ce type de réseaux est appelé (DBNN) réseaux basés sur les décisions, la deuxième catégorie est basée sur l'optimisation d'un critère qu'il faut optimiser.

II.7.1. RESEAUX DE NEURONES BASES SUR LA DECISION DBNN :

LE PERCEPTRON :

Le terme perceptron désigne aujourd'hui les réseaux de neurones statiques constitués d'une couche de neurones appelée couche d'entrée reliée à un ou plusieurs neurones, appelés neurones de sortie. L'information à l'intérieur de ces réseaux, ne circule que dans un seul sens, de l'entrée vers la sortie (Fig.II.9).

Réadaptation des poids

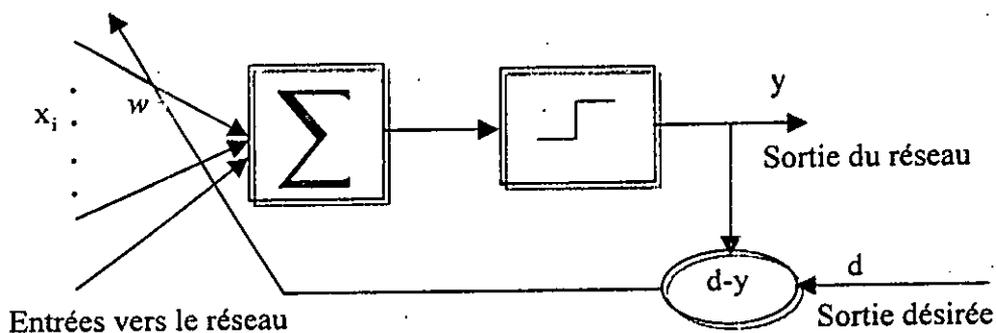


Fig.II.9. Structure du perceptron linéaire

L'activité du neurone dans ce type de réseaux est définie par l'équation linéaire suivante :

$$net_i = \sum_{j=1}^n x_j w_{ij} + \theta_i \quad ; \quad i = 1, \dots, m \quad (II.10)$$

Où :

- x_j : représentent les entées du neurone.
- w_{ij} : le poids synaptique en provenance du $j^{ème}$ neurone de la couche d'entrée vers le $i^{ème}$ neurone de sortie.
- θ : est un paramètre qui agit sur le seuil en sortie.

La sortie est définie par la fonction seuil :

$$S_i(net_i) = \begin{cases} 1 & net_i > 0 \\ 0 & net_i \leq 0 \end{cases}$$

L'adaptation des poids de ces réseaux se fait comme suit :

$$w(t+1) = w(t) + \Delta w(t) \quad (II.11)$$

tel que : $\Delta w(t) = \eta (d(t) - y(t)).x(t) \quad (II.12)$

Où :

- η : est un paramètre qui contrôle le taux d'apprentissage.
- $x(t)$: représente le vecteur présenté à l'entrée du réseau.
- $y(t)$: est la sortie du réseau.
- $d(t)$: est la sortie désirée .

II.7.2. ADALINE , MADALINE :

L'architecture du réseau ADALINE est la même que celle du perceptron, à savoir une couche d'entrée connectée à une ou plusieurs sorties. Dans le cas de plusieurs sorties, il est alors appelé MADALINE.

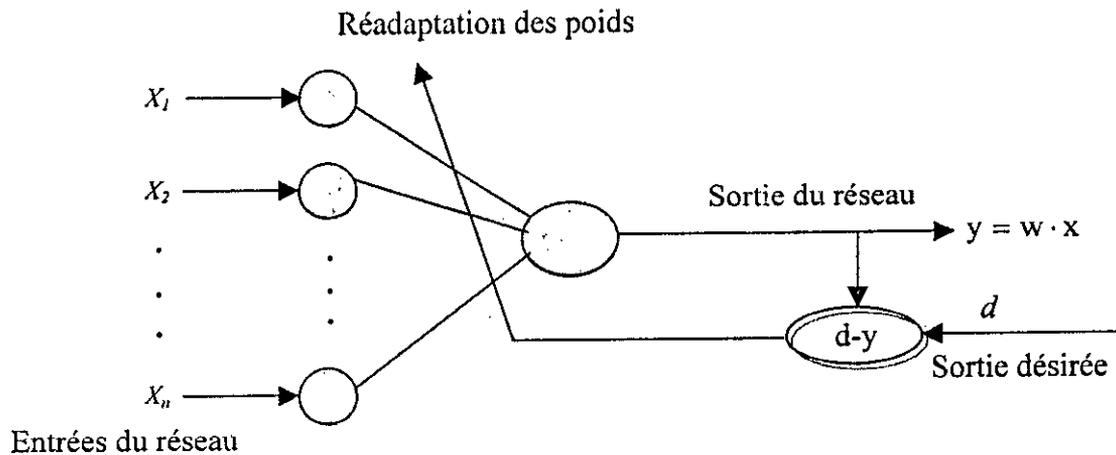


Fig.II.10. Structure du réseau ADALINE

II.7.2.1. APPRENTISSAGE :

A la présentation d'une entrée x^k , chaque sortie y_j^k le produit scalaire de cette entrée avec un vecteur poids w_j (Fig.II.11):

$$y_i^k = \sum_{j=1}^n x_j w_{ij} + \dots ; \quad i = 1, \dots, m \quad (II.13)$$

x_j : l'entrée du $j^{\text{ème}}$ neurone.

w_{ij} : poids synaptiques du $i^{\text{ème}}$ neurone de la couche de sortie provenant du $j^{\text{ème}}$ neurone de la couche d'entrée.

$$e_j^k = d_j^k - y_j^k \quad (II.14)$$

tel que :

e_j^k : est l'erreur instantanée de la sortie y_j .

d_j^k : est la réponse désiré du $j^{\text{ème}}$ élément pour la $k^{\text{ème}}$ entrée.

Il y a plusieurs méthodes qui nous permettent d'ajuster les poids w afin que, en moyenne, l'écart entre la sortie du réseau et celle désirée soit nul pour toutes les entrées.

II.7.2.1.1. METHODE DE DESCENTE DE GRADIENT :

Cette méthode est utilisée pour déterminer les adaptations qu'il faut apporter aux vecteurs des poids synaptiques à chaque présentation d'une entrée, afin de faire décroître

l'erreur de sortie jusqu'à l'obtention de la solution optimale. Pour cela, il suffit donc d'aller dans la direction opposée du gradient.

L'erreur quadratique instantanée est[28] :

$$\zeta^k = \frac{1}{2}(e^k)^2 = \frac{1}{2}(d^k - y^k)^2 \tag{II.15}$$

Afin de progresser dans le sens opposé du gradient,

$$\Delta W^k = -\eta_k \frac{\partial \zeta^k}{\partial W^k} \tag{II.16}$$

A partir des équations (II.15) et (II.16), on obtient :

$$W_{ij}^{k+1} = W_{ij}^k + \eta_k (d_i^k - y_i^k) x_j^k \tag{II.17}$$

Où :

η_k est le taux d'apprentissage.

On remarque dans ce type de réseaux de neurones qu'une augmentation du nombre de couches, n'a absolument aucune utilité, n'importe quel nombre de couche sera équivalent à une seule dont la matrice des poids est le produit des matrices poids de toutes ces couches. L'amélioration des performances de ce réseau nécessite, donc, l'introduction de non linéarité qui le rendra sensible à l'augmentation du nombre de couches.

II.7.3. RESEAUX MULTICOUCHES STATIQUES :

L'entrée dans ce type de réseaux est suivie par une succession de couches dites couches cachées, pour finalement aboutir à la sortie (FigII.13), ce passage par ces couches contribue à la richesse du traitement de l'information à l'intérieur du réseau afin d'obtenir le résultat désiré en sortie[29].

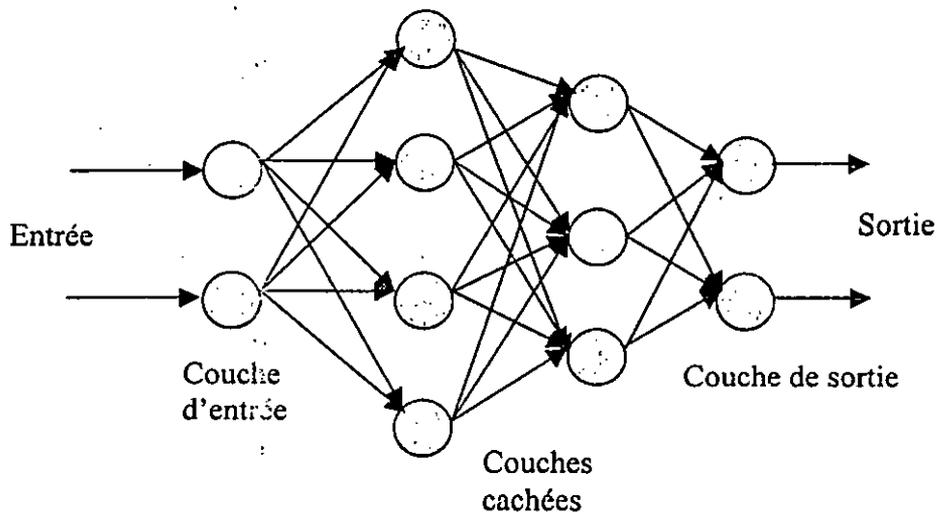


Fig.II.11. Architecture d'un réseau statique multicouche

II.7.3.1. RESEAUX A FONCTION DE BASE LINEAIRE LBF :

Le réseau LBF a un discriminant linéaire, il est régi par l'équation dynamique suivante :

$$u_i^l(t) = \sum_{j=1}^{N_{l-1}} w_{ij}^l(t) y_j^{l-1}(t) + \theta_i^l(t) \quad (II.18)$$

$$y_i^l(t) = f(u_i^l(t)) \quad (II.19)$$

tel que :

l : le nombre de couches.

$u_i^l(t)$: représente l'activité du $i^{\text{ème}}$ neurone dans la $l^{\text{ème}}$ couche

$y_i^l(t)$: la sortie du $i^{\text{ème}}$ neurone dans la $l^{\text{ème}}$ couche.

w_{ij}^l : représente le poids synaptique du $i^{\text{ème}}$ neurone de la $l^{\text{ème}}$ couche, en provenance du $j^{\text{ème}}$ neurone de la couche précédente.

θ_i^l : est un paramètre appelé Bias, il a pour objectif de donner plus de souplesse à la fonction de sortie du neurone en approximation des fonctions.

La fonction d'activation f du neurone est monotone non décroissante. Elle a pour rôle de limiter l'activité du neurone. Pour les neurones de la couche d'entrée, cette fonction n'est autre que la fonction identité. Ainsi dans la couche d'entrée, chaque neurone fait passer une composante du vecteur entrée vers sa sortie sans aucun traitement, c'est à dire $y_i^1(t) = x_i(t)$; où $x_i(t)$ représente la $i^{\text{ème}}$ entrée du réseau à l'instant t .

II.7.3.1.1. ALGORITHME DE RETROPROPAGATION DU GRADIENT :

Les réseaux de neurones à retropropagation ont été découverts indépendamment par plusieurs chercheurs de disciplines assez différentes. la première approche du gradient décroissant pour l'entraînement des réseaux de neurones multicouches est de source mathématique et a été développée par Amari (1967) qui a introduit une seule couche de neurones cachés pour réaliser une classification non linéaire. L'approche de Amari était dans le bon sens mais cela n'empêche que c'était une description non complète de la manière de développer une architecture multicouche[31].

Brypon et Ho en 1969 ont développé un algorithme très similaire jusqu'à ce que Werbos en 1974 ait découvert, en travaillant sur sa thèse de doctorat en statistiques, l'algorithme de rétro-propagation à plusieurs variantes appelé l'algorithme du retour dynamique[31].

En 1986, Rumelhart, Hinton et Williams[30] groupe d'informaticiens, exploitent le pouvoir et le potentiel de la rétro-propagation, ceci a poussé la communauté scientifique à s'y intéresser de plus en plus, ce qui a fait de lui l'algorithme connexionniste le plus utilisé qu'on désigne par « *backpropagation* ». L'erreur se rétro-propage à travers les couches du réseau (Fig.II.12).

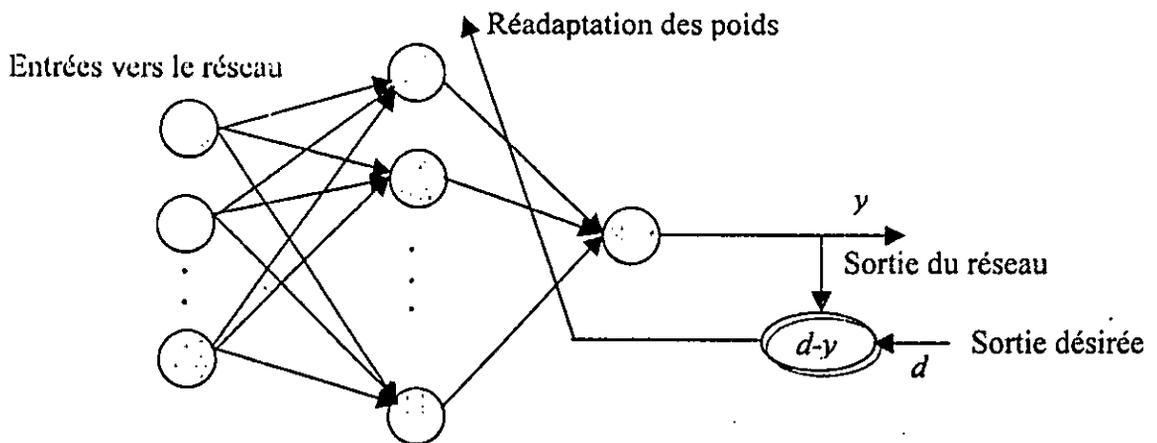


Fig.II.14. Principe de l'entraînement du réseau par rétropropagation de l'erreur

Cet algorithme s'applique comme suit :

Au début, il faut initialiser les poids synaptiques à des petites valeurs aléatoires comprises en général entre -1 et 1.

- 1- Présenter un nouvel exemple x^p à l'entrée du réseau.
- 2- calculer la sortie du réseau et l'erreur en sortie :

$$E_p = \sum_{i=1}^{n_L} (d_i^p - y_i^L)^2 \tag{II.20}$$

n_L est le nombre de neurones dans la couche de sortie.

- 3- calculer les valeurs de réadaptation des poids.

$$[\Delta w'_{ij}]_p = \eta \cdot f'(u'_i(t)) \cdot \delta'_i y'_j \tag{II.21}$$

Où :

η est le paramètre qui contrôle la vitesse de l'apprentissage appelé taux d'apprentissage.

$f(u'_i(t))$ est la sortie du $i^{ème}$ neurone de la $l^{ème}$ couche.

$\Delta w'_{ij}$ est la valeur rajoutée aux poids.

$$\text{Avec : } \delta'_i = \begin{cases} (d_i^p - y_i^L) & \text{pour les neurones de la couche de sortie} \\ \sum_{k=1}^{n_{l+1}} \delta'_{ki} w'_{ki} & \text{pour les neurones des autres couches} \end{cases} \tag{II.22}$$

4 -

- S'il s'agit d'un Data Learning, réadapter les poids.

- Poser :

$$\Delta w_{ij}^l(t) = [\Delta w_{ij}^l(t)]_p \quad (II.23)$$

- Aller à 5.

- S'il s'agit d'un block learning :

- Si $p \neq M$ (M étant le nombre d'exemples d'entraînement), retourner à 1.

- Sinon :

$$\Delta w_{ij}^l(t) = \sum_{p=1}^{p=M} [\Delta w_{ij}^l(t)]_p \quad (II.24)$$

5 - Réadapter les poids.

$$w_{ij}^l(t+1) = w_{ij}^l(t) + \Delta w_{ij}^l(t) \quad (II.25)$$

6 - Répéter de 1 à 5 autant de fois que nécessaire jusqu'à la convergence vers le seuil d'erreur fixé.

D'après cet algorithme, on peut opérer de deux manières différentes : Data Adaptive Learning ou Block Adaptive Learning.

a) Data Adaptive Learning :

Dans cette méthode, les poids sont réadaptés au passage de chaque exemple. De cette manière, le processus devient sensible à chaque exemple individuellement, ce qui le rend influençable par les bruits que peuvent contenir ces entrées durant l'entraînement.

b) Block Adaptive Learning :

Dans ce cas, on ne réadapte les poids qu'après passage de tous les exemples d'entraînement. le réajustement se fait suivant la moyenne de tous les exemples et la méthode est donc beaucoup moins sensible aux bruits que peuvent contenir des exemples singulier pouvant se présenter.

II.7.3.1.1.1. ETUDE DU TAUX D'APPRENTISSAGE :

La version théorique de la backpropagation réclame une variation infinitésimale des valeurs des poids à chaque itération. Cette variation est contrôlée par le taux d'apprentissage η . Un facteur plus grand génère des changements plus importants dans les valeurs des poids synaptiques, pouvant accélérer ainsi l'apprentissage. Malheureusement, ce n'est le plus souvent, pas le cas. En effet, une valeur importante de ce facteur engendre des oscillations et une instabilité dans la recherche du minimum, ce qui compromet la convergence[23].

Une technique très simple est utilisée, afin d'augmenter la rapidité de convergence, elle consiste à mettre en œuvre un algorithme qui a pour rôle de contrôler l'erreur d'entraînement à chaque étape (Fig.II.13).

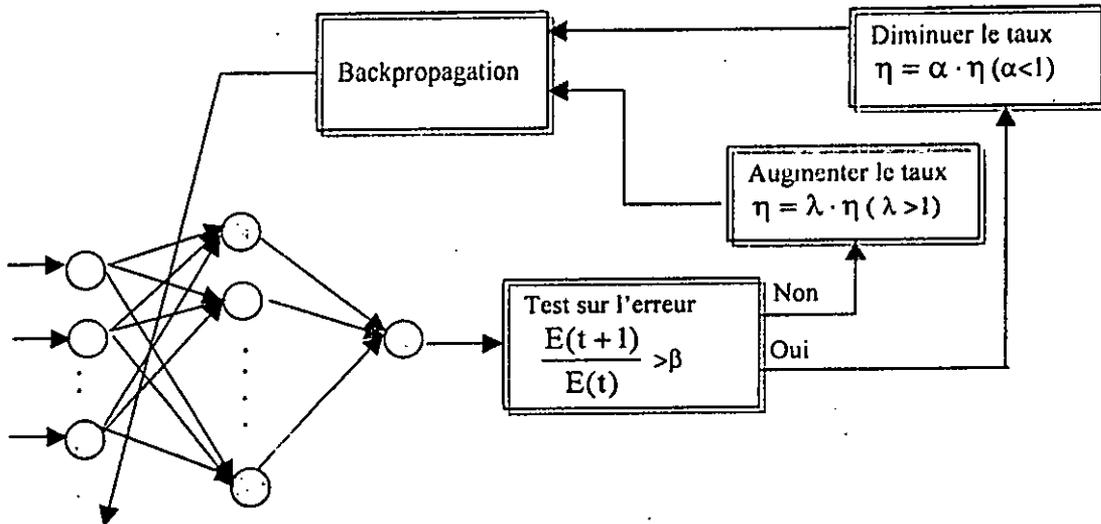


Fig.II.15. Schéma d'adaptation du taux d'apprentissage

Pour un Block Learning, à chaque fois que l'erreur actuelle dépasse la précédente par un certain seuil précédemment fixé, on rejette les poids générés, on revient au point précédent, et on diminue le taux d'apprentissage. Si par contre l'erreur diminue, les poids générés sont retenus et le taux d'apprentissage est augmentée. Cette méthode donne une convergence très rapide avec une bonne précision d'entraînement.

II.7.3.1.1.2. BACKPROPAGATION AVEC MOMENTIUM:

D.E.Rumelhart a proposé une solution très efficace pour accélérer la convergence, qui consiste à utiliser les changements précédents des poids pour la réadaptation des poids actuels[30]. L'équation d'adaptation devient :

$$w_{ij}^l(t+1) = w_{ij}^l + \eta \Delta w_{ij}^l(t) + \alpha \Delta w_{ij}^l(t-1) \tag{II.26}$$

Le terme ajouté est appelé momentum (élan), en analogie avec le système mécanique classique, où un objet en mouvement garde l'élan acquis grâce à la quantité de mouvement qui lui a été communiquée précédemment pour accélérer son mouvement (Fig.II.14). Le paramètre α est utilisé pour pondérer l'effet de ce terme. Sa valeur est généralement prise entre 0.8 et 0.9.

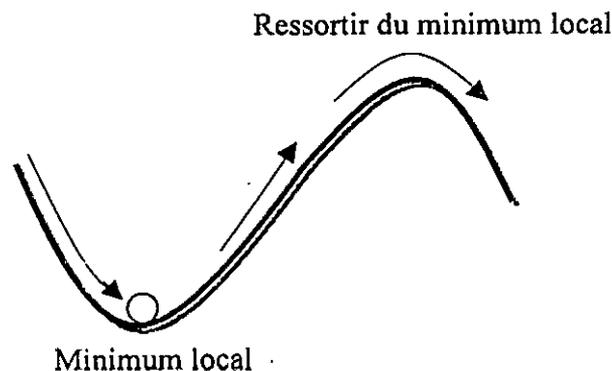


Fig.II.14. Effet du momentum pour échapper d'un minimum local

L'utilisation de cette méthode permet de faire sortir les poids des minimums locaux, afin de chercher d'autres optimums, ce qui donne beaucoup de chances d'aboutir à un minimum global.

II.7.3.1.2.METHODE D'OPTIMISATION ALEATOIRE ROM :

En 1965, J.Mathias a présenté cette méthode, elle a été développée par F.J.Sollis et J.B.Wetts en 1981, qui ont démontré la convergence de l'algorithme vers un minimum global[23].

Nous présentons ci-dessous l'algorithme de cette méthode :

1. Initialiser aléatoirement tous les poids W du réseau. Initialiser la variance v et la moyenne de la séquence $b(0)=0$.

2. Générer un nouvelle séquence aléatoire $\zeta(k)$ de moyenne $b(k)$ et de variance $v(k)$.

3. Calculer les erreurs $E = \sum_{p=1}^M E_p$ à la sortie du réseau pour chaque cas :

- $E(w(k))$
- $E(w(k) + \zeta(k))$
- $E(w(k) - \zeta(k))$

4.

- Si $E(w(k)+\zeta) < E(w(k))$, alors :

$$\begin{cases} w(k+1)=w(k)+ \zeta(k) \\ b(k+1)=0.4\zeta(k)+0.2b(k) \end{cases}$$

- Si $E(w(k)+\zeta(k)) \geq E(w(k))$ et $E(w(k)-\zeta(k)) < E(w(k))$, alors :

$$\begin{cases} w(k+1)=w(k)-\zeta(k) \\ b(k+1)=b(k)-0.4\zeta(k) \end{cases}$$

- Sinon,

$$\begin{cases} w(k+1)=w(k) \\ b(k+1)=0.5b(k) \end{cases}$$

5. Poser $k=k+1$ et refaire de 2 à 4 jusqu'à l'obtention de l'erreur désirée.

Contrairement à la méthode de Backpropagation où le calcul du jacobien est nécessaire, le grand avantage de cette méthode est son indépendance du critère à optimiser ; son utilisation dépend seulement de la variance du vecteur gaussien. Une séquence avec une grande variance génère de grands changements dans l'erreur mais jamais des oscillations. Il est intéressant de commencer l'entraînement avec une variance importante et de la faire baisser au fur et à mesure que l'entraînement avance.

II.7.3.1.3. RESEAUX LBF ET APPROXIMATION DE FONCTIONS :

En 1989, Hornik et White ont montré une architecture neurale à sortie sigmoïdale peut approximer n'importe quelle fonction avec la précision désirée, à condition d'avoir assez de neurones entre la couche d'entrée et celle de sortie[23].

La question posée est de déterminer le nombre de couches cachées à utiliser et le nombre de neurones nécessaires pour ces couches.

En 1990, Cybenko a prouvé qu'une couche cachée suffit pour approximer n'importe quelle fonction. Mais l'inconvénient est que le nombre de neurones que doit contenir cette couche peut être « astronomique »[23].

Cybenko et Funahashi ont émis un théorème en 1991 selon lequel, si $f(x)$ est une fonction continue croissante et réelle, et $t(x)$ est une fonction continue non croissante dans un sous-espace S dans R .

$\forall \varepsilon > 0, \exists k \in N, w_{ij} \in R, \theta_i \in R$ tel que :

$$\phi = \sum c_i f \left(\sum w_{ij} x_j + \theta_i \right) \quad (II.27)$$

Satisfaisant : $|\phi(X, w) - t(X)| < \varepsilon$ pour $t, X \in S$.

Ce théorème prouve que les réseaux LBF sont capables d'approximer n'importe quelle fonction continue dans l'espace[28].

II.7.3.1.4. DIMENSION D'UN RESEAU LBF :

La dimension d'un réseau permettant d'avoir un meilleur résultat est impossible à fixer. Il a été constaté qu'un nombre important de neurones, permet d'obtenir une bonne précision d'entraînement, par contre, un nombre insuffisant entraînera sa dégradation.

Malheureusement un réseau trop chargé en neurones cachés souffre de manque de généralisation, échouant ainsi devant des exemples autres que ceux présents dans la base d'apprentissage. Un nombre réduit de neurones qui devrait être favorable à la généralisation rend la précision d'entraînement très difficile à obtenir. Ce problème constitue le dilemme précision-généralisation. Il faut donc réussir un compromis entre les deux.

Un autre problème important dont souffrent les réseaux LBF, est la difficulté de subir l'apprentissage avec de nouveaux exemples sans perdre les caractéristiques déjà acquises. En effet, la fonction d'activation des neurones, qui est non locale, favorise ce phénomène d'oubli et rend ces réseaux souvent incapables d'apprendre de nouveaux exemples après avoir déjà subi un premier apprentissage[32].

II.7.3.2. RESEAUX A FONCTION DE BASE RADIALE RBF :

Les réseaux de neurones à fonction de base radiale, outre les couches d'entrée et de sortie, ils comprennent une couche cachée de n unités qui ne réagissent significativement qu'à une partie restreinte dans l'espace d'entrée suivant une fonction d'activation de type gaussien[31].

A travers une combinaison linéaire de fonctions non linéaires de base radiale, le fonctionnement de ces réseaux repose sur le principe des estimateurs à noyau, et le généralise du monovarié vers le multivarié. Un estimateur à noyau considère des fonctions de R^+ vers R qui sont de la forme :

$$\phi_i(x) = \phi(\|x - \xi_i\|) \tag{II.28}$$

L'idée principale a été introduite par M.J.D.Powell, toute fonction $f(x)$ d'une variable $x \in R^n$ peut être approchée par une interpolation composée par la somme de P fonctions noyaux de forme fixée $\phi(\|x - \xi_i\|)$ [20]:

$$f(x) = \sum_{i=1}^P \lambda_i \phi(\|x - \xi_i\|) \tag{II.29}$$

Où ξ_i représente le nœud d'interpolation pour $i=1, \dots, n$.

λ_i sont les paramètres que l'on déterminera à partir des exemples connus x_i, γ_i en résolvant le système :

$$f(x) = \gamma_i \quad i=1, \dots, n \tag{II.30}$$

$\phi(\|x - \xi_i\|)$ est une fonction assurant la continuité aux nœuds et la dérivabilité d'ordre supérieur en ces points[33].

II.7.3.2.1. ARCHITECTURE ET FONCTIONNEMENT DES RESEAUX RBF :

Parmi les fonctions noyaux qui sont utilisées dans les réseaux RBF, on peut trouver les formes suivantes[32] :

Formes cubiques : $\phi(r) = r^3$

Formes mutiquadratiques : $\phi(r) = (r^2 + k^2)^{-1/2}$

Formes gaussiennes : $\phi(r) = \exp\left(-\frac{r^2}{\beta^2}\right)$

Où r est un réel quelconque, remplaçant le discriminant des neurones. Mais la fonction gaussienne est la plus utilisée.

L'architecture d'un réseau RBF est représentée à la figure (II.15) :

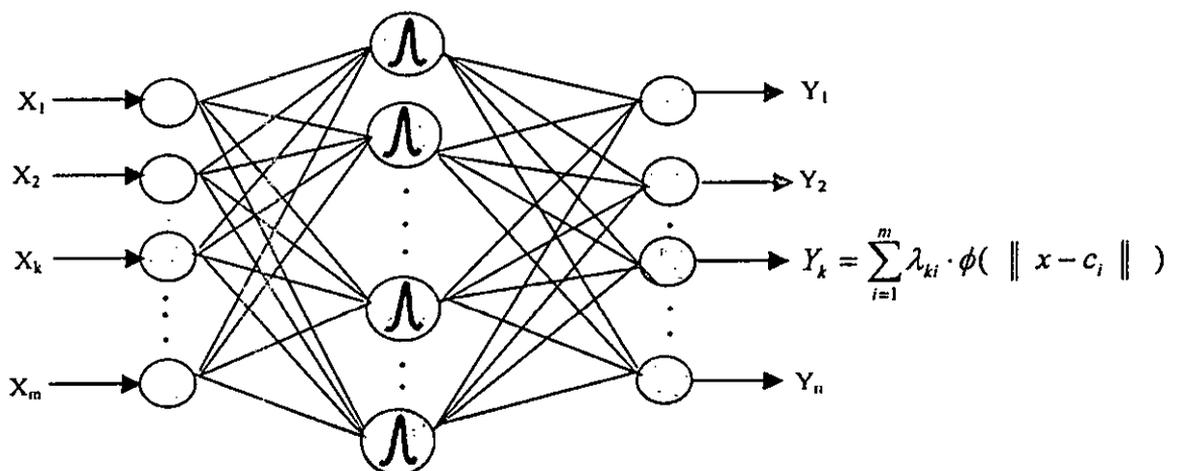


Fig.II.15. Architecture d'un réseau RBF

Il est composé de trois couches. La première reçoit les entrées injectées au réseau, la couche cachée est constituée de neurones, dont le discriminant est à fonction de base radiale. Chacun de ses neurones est doté d'un vecteur c_i appelé centre, les valeurs de ce vecteur représente chacune le barycentre de la classe qui lui correspond. Lorsqu'un vecteur X est présenté à l'entrée du réseau, chaque neurone calcule, à travers son discriminant, la distance euclidienne de cet exemple par rapport aux centres de la classe qui le représente :

$$d_i = \|X - c_i\| \quad (\text{II.31})$$

la sortie de chaque neurone de cette couche est donnée comme suit :

$$s_i = \exp\left[\frac{(X - c_i)^T (X - c_i)}{2\sigma_i^2}\right] \quad (\text{II.32})$$

Où σ_i est un paramètre qui représente une mesure de la dispersion des données associées à chaque nœud. La couche de sortie est constituée de neurones linéaires effectue l'approximation définie par la somme pondérée suivante :

$$y_k(X) = \sum_{i=1}^m \lambda_{ki} \phi(\|X - c_i\|) \quad 1 \leq k \leq n \quad (\text{II.33})$$

Avec n le nombre de neurones de sortie.

II.7.3.2.2. APPRENTISSAGE DES RESEAUX RBF GENERALISES :

Ces réseaux permettant non seulement l'apprentissage des poids des connexions w , mais aussi celui des centres et des rayons sont appelés Réseaux à Base Radiale généralisés. L'ajustement des centres et des rayons de ces réseaux peut être aisément effectué au moyen d'un algorithme de type descente de gradient[32].

Nous présentons ci-dessous les étapes de cette technique d'apprentissage supervisé :

1. Initialiser les centres $c_i(0)$, les paramètres de la sortie de chaque neurone de la couche cachée $\sigma_i(0)$, $i = 1, \dots, m$, et les poids synaptiques $\lambda_{ij}(0)$, $j = 1, \dots, n$.
2. Choisir un taux d'entraînement $\alpha < 1$ variable ou fixe.
3. Présentation successive de P exemples d'entraînement .

Calcul de l'erreur quadratique en sortie E entre la réponse désirée et celle fournie par le réseau.

$$E = \sum_{p=1}^M \sum_{j=1}^n (d_j^p - y_j^p)^2 \quad (\text{II.34})$$

Avec :

M représentant le nombre d'exemples d'entraînement

y_j^p d_j^p représentent respectivement la $j^{\text{ème}}$ sortie du réseau et celle désirée pour le $p^{\text{ème}}$ vecteur d'entrées.

4. Réajuster les centres et les paramètres des fonctions noyau :

$$\begin{aligned}c_i(t+1) &= c_i(t) + \alpha \Delta c_i(t) \\ \sigma_i(t+1) &= \sigma_i(t) + \alpha \Delta \sigma_i(t)\end{aligned}\quad (\text{II.35})$$

Avec :

$$\begin{aligned}\Delta c_i(t) &= -\frac{\partial E}{\partial c_i(t)} \\ \Delta \sigma_i(t) &= -\frac{\partial E}{\partial \sigma_i(t)}\end{aligned}\quad 1 \leq i \leq m \quad (\text{II.36})$$

5. Réajuster les poids synaptiques :

$$\lambda_{ji}(t+1) = \lambda_{ji}(t) + \alpha \Delta \lambda_{ji}(t) \quad (\alpha \text{ est le taux d'apprentissage}) \quad (\text{II.37})$$

Avec :

$$\Delta \lambda_{ji}(t) = -\frac{\partial E}{\partial \lambda_{ji}(t)}, \quad 1 \leq j \leq n ; \quad 1 \leq i \leq m \quad (\text{II.38})$$

6. Répéter les étapes de 3. à 5. , le nombre de fois nécessaire jusqu'à l'obtention de la précision désirée en sortie.

II.7.3.2.3. RESEAUX RBF ET APPROXIMATION DE FONCTION :

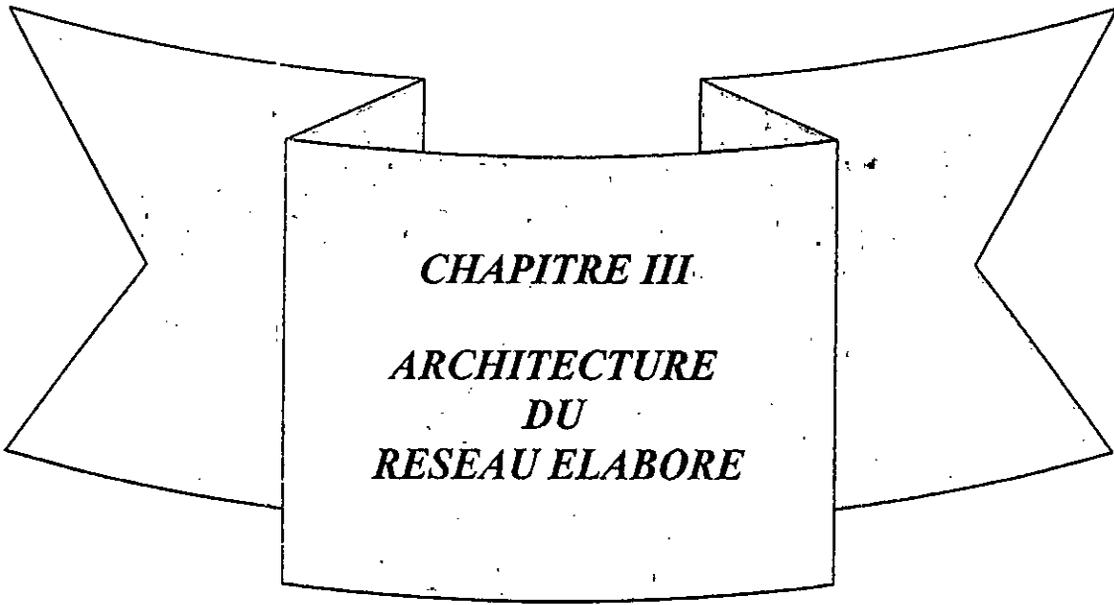
Les réseaux RBF constituent des modèles de réseaux très efficaces notamment pour l'approximation des fonctions. L'utilisation de la fonction Gaussienne permet de bénéficier de sa caractéristique locale pour faciliter l'apprentissage et améliorer l'approximation. Par ailleurs, la procédure d'entraînement, basée sur le fonctionnement de ce réseau, qui consiste en une classification suivie d'une optimisation rend l'apprentissage beaucoup moins difficile et plus rapide que celui des réseaux LBF. Par ailleurs, ces réseaux sont utilisés avec une seule couche cachée, ceci libère l'utilisateur du choix du nombre de couches.

Néanmoins ces réseaux présentent quelques inconvénients. En effet, lorsque le nombre d'exemples d'apprentissage est important, la couche cachée risque d'être surchargée en neurones, compliquant ainsi les calculs et ralentissant l'apprentissage.

II.8. CONCLUSION :

- La méthode de Backpropagation constitue l'algorithme d'apprentissage supervisé le plus utilisé, néanmoins, elle souffre de certaines limitations. La vitesse de convergence est lente, et de plus cet algorithme se bloque souvent dans des minimums locaux.
- L'algorithme d'optimisation aléatoire ROM, constitue une alternative très intéressante. En effet, cet algorithme ne dépend pas du système pour lequel le réseau est utilisé. De plus sa technique de recherche permet d'atteindre un minimum global.
- Les réseaux LBF souffrent de l'indétermination du nombre optimal d'élément dans chaque couche . Mais concernant la dimension du réseau , il est clair qu'un réseau à deux couches constitue un approximateur universel.

- Les réseaux RBF se distinguent par leur caractéristique de meilleurs approximateurs. En effet ces réseaux sont dotés de la caractéristique de calcul et d'adaptation paramétrique locaux. Ceci fait que ces réseaux soient à l'abri de plusieurs problèmes que connaissent les LBF, comme le sur-apprentissage, et la difficulté d'introduction de nouveaux exemples sans toucher à la précision d'apprentissage des exemples précédemment appris.



CHAPITRE III
ARCHITECTURE
DU
RESEAU ELABORE

III.1. INTRODUCTION :

L'objectif de ce chapitre est d'élaborer un programme qui nous permet de prédire la suite de certaines propriétés électriques d'une huile isolante (BORAC 22) en fonction du temps de vieillissement sous une température constante. Comment savoir que notre réseau donnera une bonne prédiction ?

D'abord nous allons prendre de l'intervalle du temps d'une propriété, une partie de durée relativement courte, dont on fait l'apprentissage, puis on effectue la prédiction dans le reste de l'allure. finalement à partir de ce travail, on peut vérifier la qualité de prédiction en comparant les points prédits avec ceux de la courbe réelle. Nous allons voir les détails dans les paragraphes qui suivent.

III.2. ARCHITECTURE DU RESEAU ELABORÉE :

Comme nous avons vu dans le deuxième chapitre, que les réseaux RBF constituent des modèles de réseaux très efficaces pour la généralisation des fonctions. L'utilisation de la fonction gaussienne permet de bénéficier de sa caractéristique locale pour faciliter l'apprentissage et améliorer la généralisation. En outre, l'avantage le plus important c'est que ces réseaux sont toujours utilisés avec une seule couche cachée, ce qui nous libère du choix de nombre de couches.

Les réseaux de neurones à base radiale peuvent aisément et naturellement satisfaire aux contraintes suivantes[32] :

- Contrainte de capacité d'approximation universelle :

Toute fonction $f(x)$ d'une variable $x \in \mathbb{R}^n$, peut être approchée par une interpolation composée par la somme de P fonctions noyaux de forme fixée $\phi(x)$:

$$f(x) = \sum_{j=1}^P \lambda_j \Phi(\|x - \zeta_j\|) \quad (\text{III.1})$$

Où :

- ζ_i : représentant les nœuds de l'interpolation pour $i=1, \dots, n$.
- λ_j : sont les paramètres que l'on déterminera.

- Contrainte de couverture :

Elle s'exprime par le fait que, pour chaque sortie y_i associée à un vecteur d'entrée x , il existe au moins un paramètre w_j tel que la valeur $\left| \frac{\partial y_i(x, w)}{\partial w_j} \right|$ soit suffisamment grande dans le voisinage de x .

- Contrainte de généralisation locale :

Si la valeur $\left| \frac{\partial y_i(x, w)}{\partial w_j} \right|$ est grande dans le voisinage de x , elle doit être relativement proche de zéro une fois que l'on s'éloigne de ce voisinage.

Tout réseau à base radiale, il comprend outre les couches d'entrée et de sortie, une couche cachée de n unités qui réagissent suivant une fonction d'activation de type Gaussien (Fig.III.1), la valeur de sortie de chaque neurone de la couche de sortie est donnée par :

$$y_i(X) = \sum_{j=1}^L w_{ji} \rho_j(X) \quad i=1, \dots, n \quad (III.2)$$

Où :

$$\rho_j(X) = \exp\left(-\frac{1}{2} \sum_{k=1}^m \frac{(x_k - c_{kj})^2}{\sigma_{kj}^2}\right) \quad (III.3)$$

m est le nombre de neurones dans la couche d'entrée.
 L est le nombre de neurones dans la couche cachée.
 n est le nombre de neurones dans la couche de sortie.

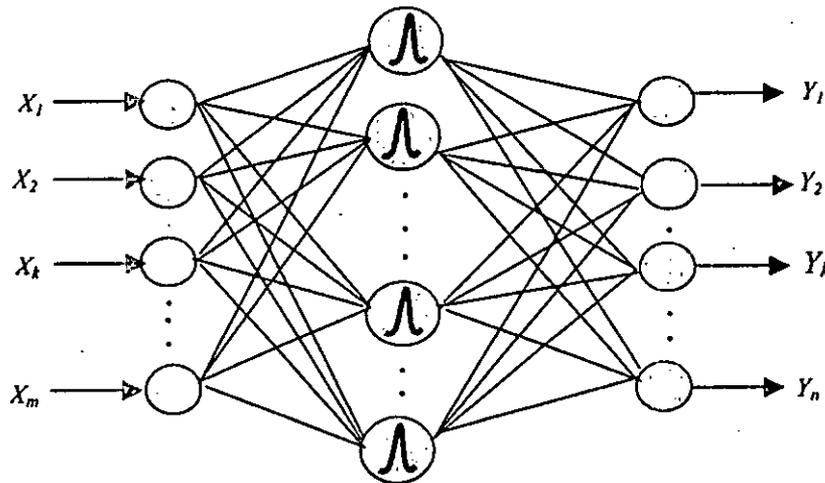


Fig.III.1. Architecture d'un réseau RBF

- c_{kj} : le centre du $j^{\text{ème}}$ nœud de la couche cachée correspondant à la $k^{\text{ème}}$ entrée.
- x_k : est la $k^{\text{ème}}$ entrée.
- X : est le vecteur d'entrée.
- $\rho_j(x)$: est la fonction d'activation du $j^{\text{ème}}$ nœud de la couche cachée pour le vecteur d'entrée X .

Les centres des fonctions gaussiennes sont choisis par plusieurs techniques afin de les mettre comme points représentatifs dans les zones de grandes concentrations de données.

Mais il existe une méthode plus simple qui consiste à disposer les centres en treillis régulier pour couvrir uniformément la partie utile de l'espace d'entrée. Les rayons des gaussiennes dans ce cas, sont fixés par les distances entre centres voisins afin d'assurer un certain degré de recouvrement entre les différentes gaussiennes. La (Fig.III.2) définit le taux de recouvrement τ d'un réseau à base radiale à treillis régulier unidimensionnel (la couche d'entrée contient un seul neurone), en supposant les gaussiennes d'écart-type constant(σ) et une distances entre centres de δ , le taux de recouvrement est donné par[32] :

$$\tau = \exp \left(- \frac{1}{2} \frac{\left(\frac{\delta}{2} \right)^2}{\sigma^2} \right) \quad (\text{III.4})$$

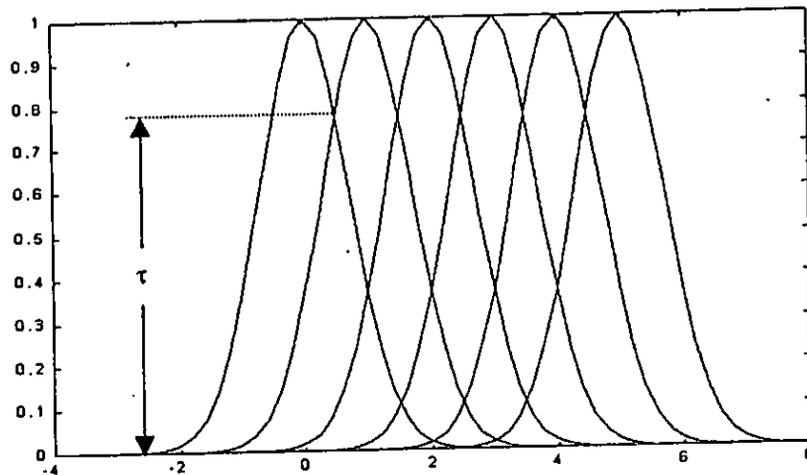


Fig.III.2. Schéma d'un treillis régulier unidimensionnel

Pour un réseau à treillis régulier et uniforme sur le domaine d'approximation, le nombre de centres (L) et le taux de recouvrement (τ) définissent entièrement la topologie du réseau .

Mais quelquefois, la disposition des centres à priori s'avère impraticable, car, le nombre de centres nécessaires est parfois trop élevé (dans le cas où l'on décide d'adopter une disposition en treillis régulier). Il y a une méthode très efficace qui consiste à utiliser un nombre beaucoup plus restreint de centres et de laisser libre la disposition de ces centres ; l'algorithme d'apprentissage ajuste, outre les poids w_j , les positions des centres ainsi que les valeurs des rayons associées. Le réseau s'auto-organisera pour placer un nombre élevé de centres là où la fonction est accidentée, et espacer les centres dans les régions où la fonction est plus lisse. Ces réseaux nous permettent non seulement l'apprentissage des poids des connexions w , mais aussi celui des centres et des rayons, ils sont appelés « Réseaux à Base Radiale Généralisés » .

L'ajustement des centres et des rayons d'un réseau à base radiale généralisé peut être aisément effectuée au moyen d'un algorithme de type « descente de gradient ».

Soit $j = \frac{1}{2}(y - y^*)$ le critère à minimiser.

Les connexions s'obtiennent par :

$$\begin{aligned}\Delta w_j &= -\eta_w \frac{\partial j}{\partial w_j} \\ \Delta c_{jk} &= -\eta_c \frac{\partial j}{\partial c_{jk}} \\ \Delta \sigma_{jk} &= -\eta_\sigma \frac{\partial j}{\partial \sigma_{jk}}\end{aligned}\quad (\text{III.5})$$

Tels que η_w , η_c et η_σ sont respectivement les taux d'apprentissage des poids w_j , des centres c_{jk} et des rayons σ_{jk} (voir l'annexe 1).

Dans notre étude nous allons utiliser un réseau à base radial normalisé, la $i^{\text{ème}}$ sortie du réseau s'exprime sous la forme suivante :

$$y_i = \frac{\sum_{j=1}^L w_{ji} \rho_j(x)}{\sum_{j=1}^L \rho_j(x)} \quad (\text{III.6})$$

Où $\sum_{j=1}^L \rho_j(x)$ est le terme de normalisation.

Ce réseau diffère du réseau non normalisé dans la propriété de calcul local : dans un réseau non normalisé, le neurone requiert uniquement des données locales pour le calcul de son activation. Par contre l'introduction du facteur de normalisation, qui est une qualité commune à chaque neurone, donne à celui-ci des données propre à lui et aussi des informations concernant les autres neurones.

Le réseau que nous avons utilisé est constitué d'un neurone dans la couche d'entrée (treillis régulier unidimensionnel), d'une couche cachée de nombre d'unités égal au nombre de centres des gaussiennes et d'une couche de sortie constituée d'un neurone. La fonction d'activation du $j^{\text{ème}}$ neurone dans ce cas, pour une entrée x_i s'exprime comme suit :

$$\rho_j(x_i) = \exp\left(\frac{-1}{2} \cdot \frac{(x_i - c_j)^2}{\sigma_j^2}\right) \quad (\text{III.7})$$

c_j et σ_j sont respectivement le centre et la variance du $j^{\text{ème}}$ neurone de la couche cachée.

III.2. CHOIX DE L'ALGORITHME D'APPRENTISSAGE :

Au début, nous avons essayé de travailler avec un réseau à base radiale généralisé et d'utiliser la méthode de « descente de gradient », mais nous étions confrontés au problème du choix des taux d'apprentissage des poids, centres et variances. S'ils sont trop petits, on avance trop lentement à la recherche de l'optimum, donc le temps d'exécution va être trop lent, par contre s'ils sont grands, les paramètres à ajuster oscillent autour de l'optimum. Ce type d'ajustement est aussi la source du problème des minimums locaux, car il dépend du gradient de l'erreur. Avec toutes ces contraintes, nous nous sommes orientés vers l'utilisation de la méthode d'optimisation aléatoire (ROM) qui est basée sur la génération

des séquences aléatoires, afin de déplacer les valeurs des poids synaptiques dans l'espace à la recherche de l'optimum. Cette technique ne nécessite pas de déterminer le gradient, elle permet donc d'échapper aux minimums locaux. L'algorithme de cette méthode est présenté dans le chapitre II. On peut représenter cette technique par le schéma bloc suivant[34] :

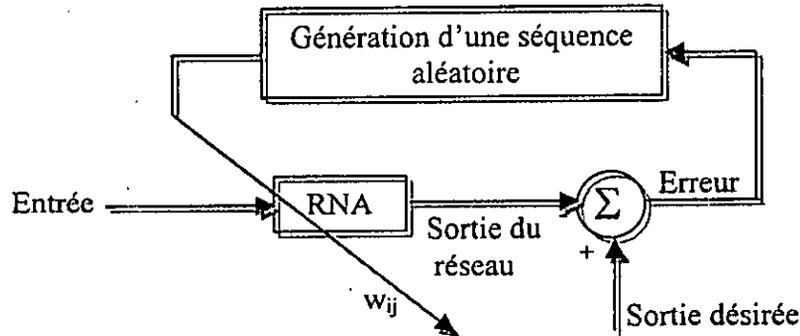


Fig.III.4. Schéma représentatif de la ROM

L'utilisation de la ROM dépend de la variation du vecteur gaussien, une séquence avec une grande variation génère de grands changements dans l'erreur, mais jamais des oscillations. Cependant, elle marque quelquefois des stagnations de la surface d'erreur. Dans notre travail, nous avons commencé l'entraînement avec une variance de valeur importante que nous multiplions par un paramètre réducteur à chaque fois que l'erreur se bloque sur un point.

III.3. APPLICATION DU RÉSEAU ÉLABORÉ DANS LE VIEILLISSEMENT THÉRMIQUE :

1. PHASE D'APPRENTISSAGE :

Nous avons essayé de tester deux réseaux, nommés RBF1 et RBF2. Leur méthode d'apprentissage se fait de deux manières différentes :

- Le premier comporte une entrée et une sortie de natures différentes, en effectuant l'apprentissage avec des exemples de type (t_i, y_i) . C'est à dire, on présente à l'entrée du réseau les valeurs de temps $t_i(h)$ et on ajuste sa sortie pour avoir les valeurs de la propriété correspondantes aux temps injectés.
- Le deuxième comporte une entrée et une sortie de même nature. Dans ce cas, l'apprentissage s'effectue avec des exemples de type (y_i, y_{i+1}) . C'est à dire, les paramètres du réseau sont adaptés de telle manière que lorsqu'on lui présente à l'entrée par exemple la valeur de la propriété correspondante au temps t_i , il nous donne à la sortie la valeur correspondante au temps t_{i+1} .

Nous avons utilisé dans les deux cas la même technique d'apprentissage « Block Adaptive Learning ». C'est à dire, les poids sont réadaptés après le passage de tous les

exemples d'entraînement, dans ce cas, le réseau n'est pas sensible à chaque exemple individuellement, ce qui est préconisé dans un problème de prédiction.

2. PHASE DE PRÉDICTION :

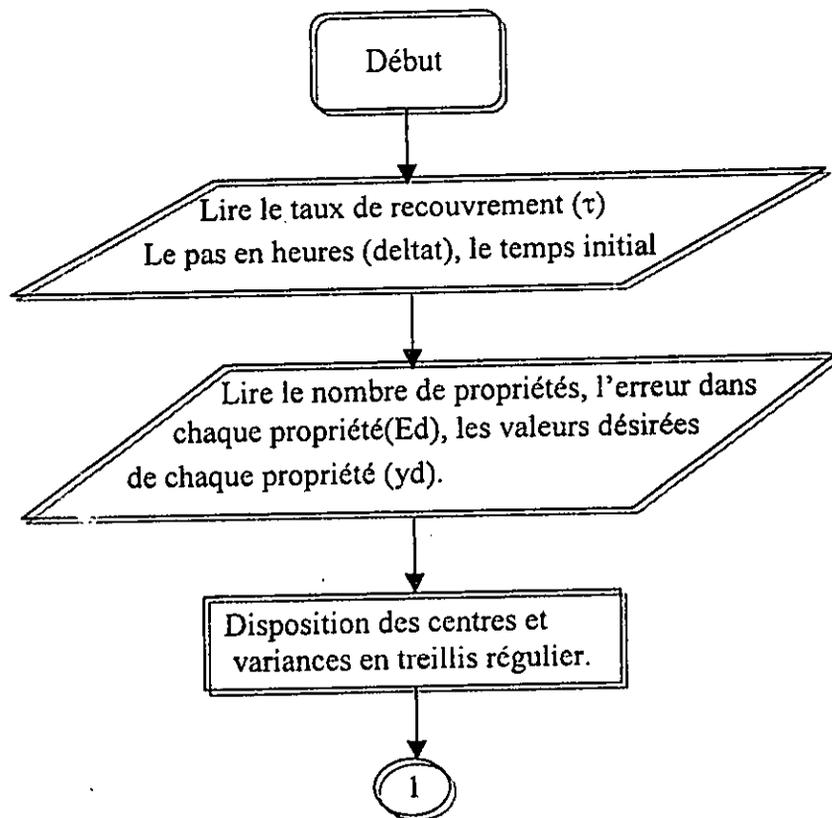
A) LE RESEAU RBF 1 :

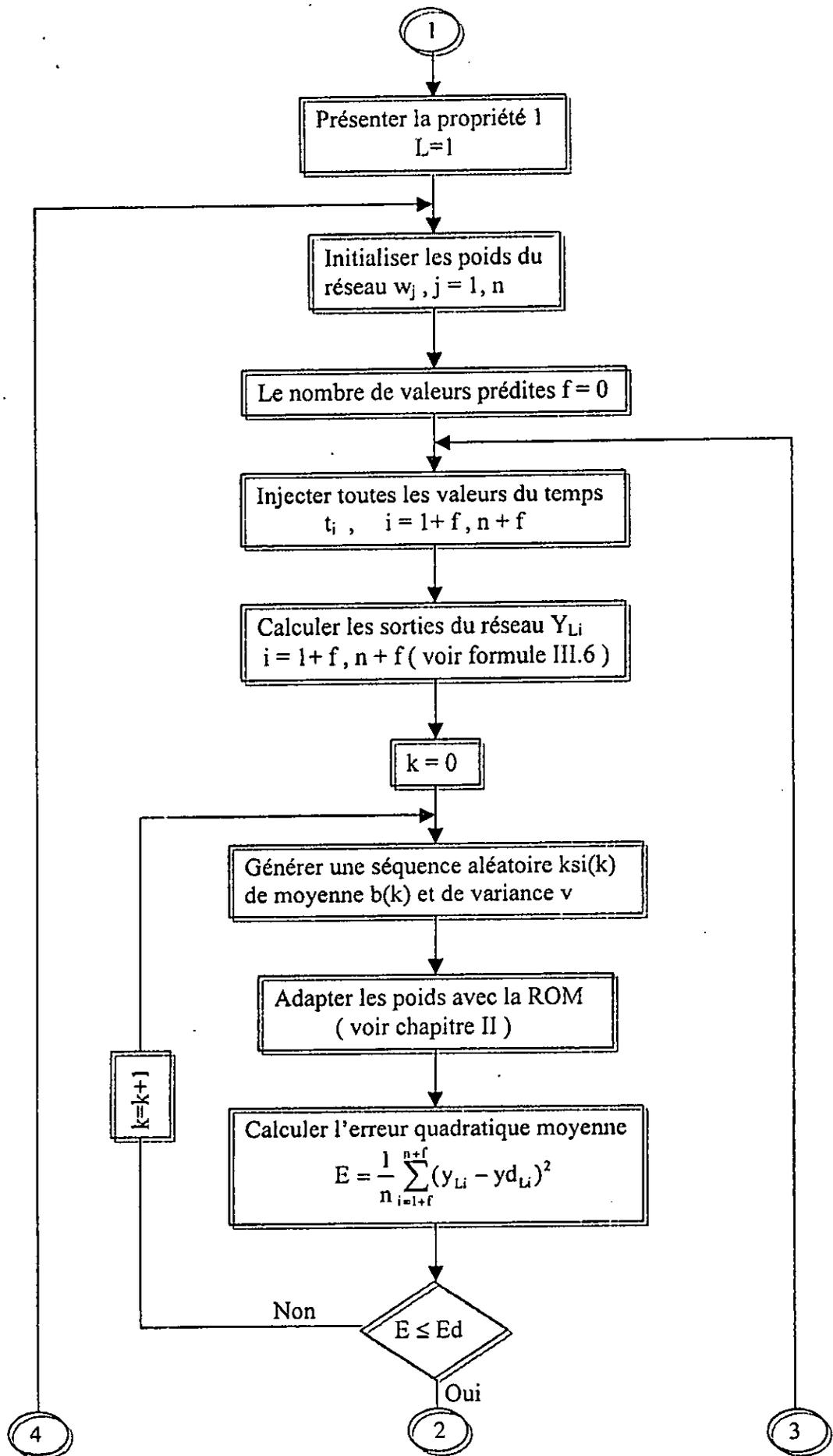
Après la présentation de tous les exemples d'entraînement $(t_1, y_1), \dots, (t_n, y_n)$ et l'adaptation des paramètres du réseau, on présente à l'entrée du réseau le temps t_{n+1} dont on ne connaît pas la sortie désirée y_{n+1} . On la considère comme sortie exacte (valeur prédite) et on refait l'apprentissage dès le début en présentant les exemples $(t_1, y_1), \dots, (t_{n+1}, y_{n+1})$. On refait les mêmes procédés, jusqu'à ce qu'on arrive à injecter le temps t_{n+m} (m est le nombre de valeurs à prédire) et on obtient à la sortie la dernière valeur prédite. Ensuite l'exécution du programme sera arrêté, et toutes les valeurs prédites vont être affichées sur le fichier de résultats.

Mais, vu le problème de sur-apprentissage au cours de la prédiction, car à chaque fois on augmente le nombre d'exemples d'entraînement, et vu aussi le problème des propriétés expérimentales qui ne prennent pas la même forme durant tout le temps de vieillissement, nous avons essayé de rendre la base d'apprentissage évolutive au cours du temps. C'est à dire à chaque fois qu'on reprend l'apprentissage, on élimine la première valeur de la base d'apprentissage et devient $(y_{1+f}, \dots, y_{n+f})$, tels que :

f : est la $f^{\text{ème}}$ valeur prédite ($0 \leq f < m$).
 m est le nombre de valeurs à prédire.

Nous présentons ci-dessous l'organigramme de ce réseau :





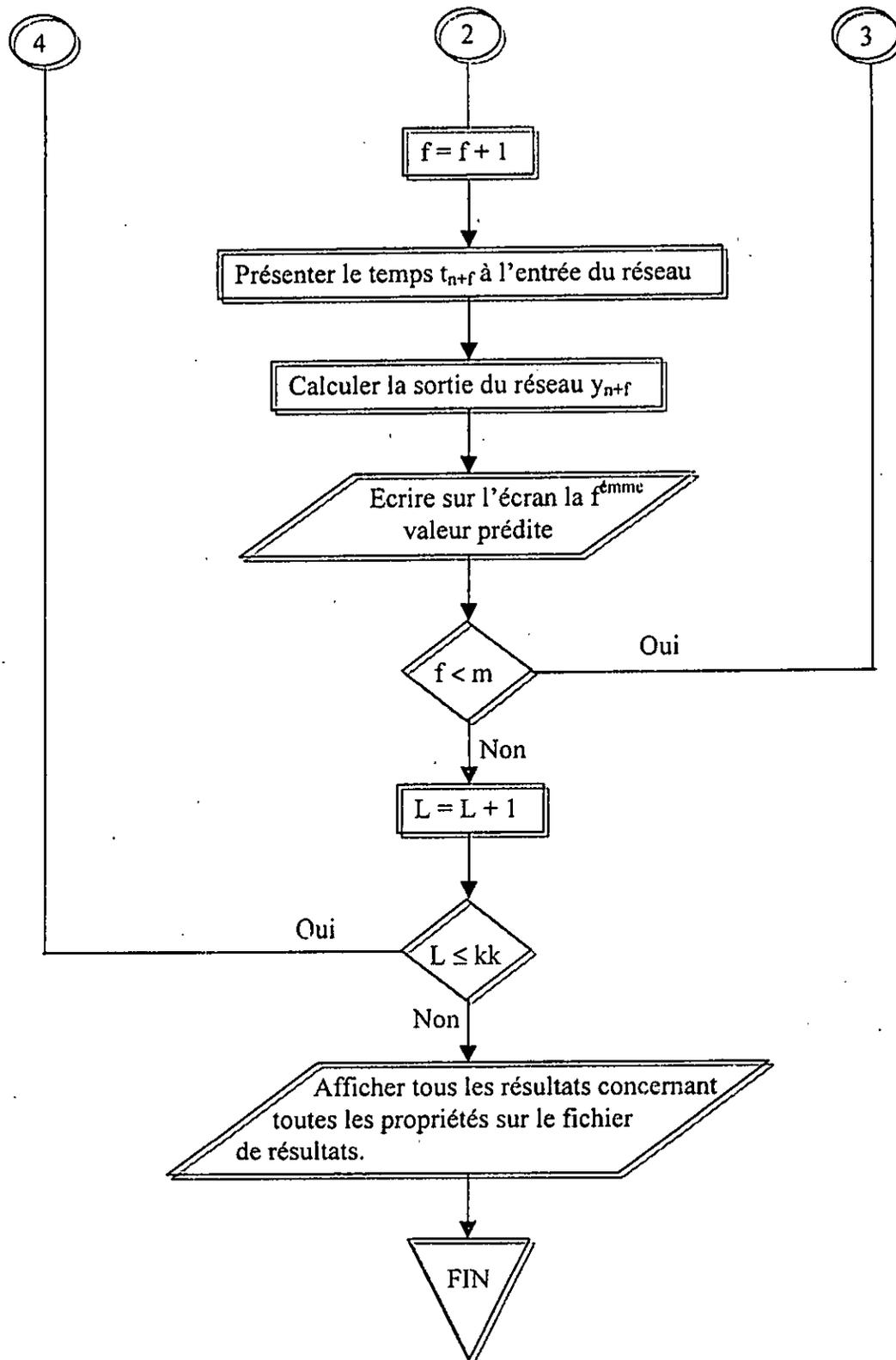
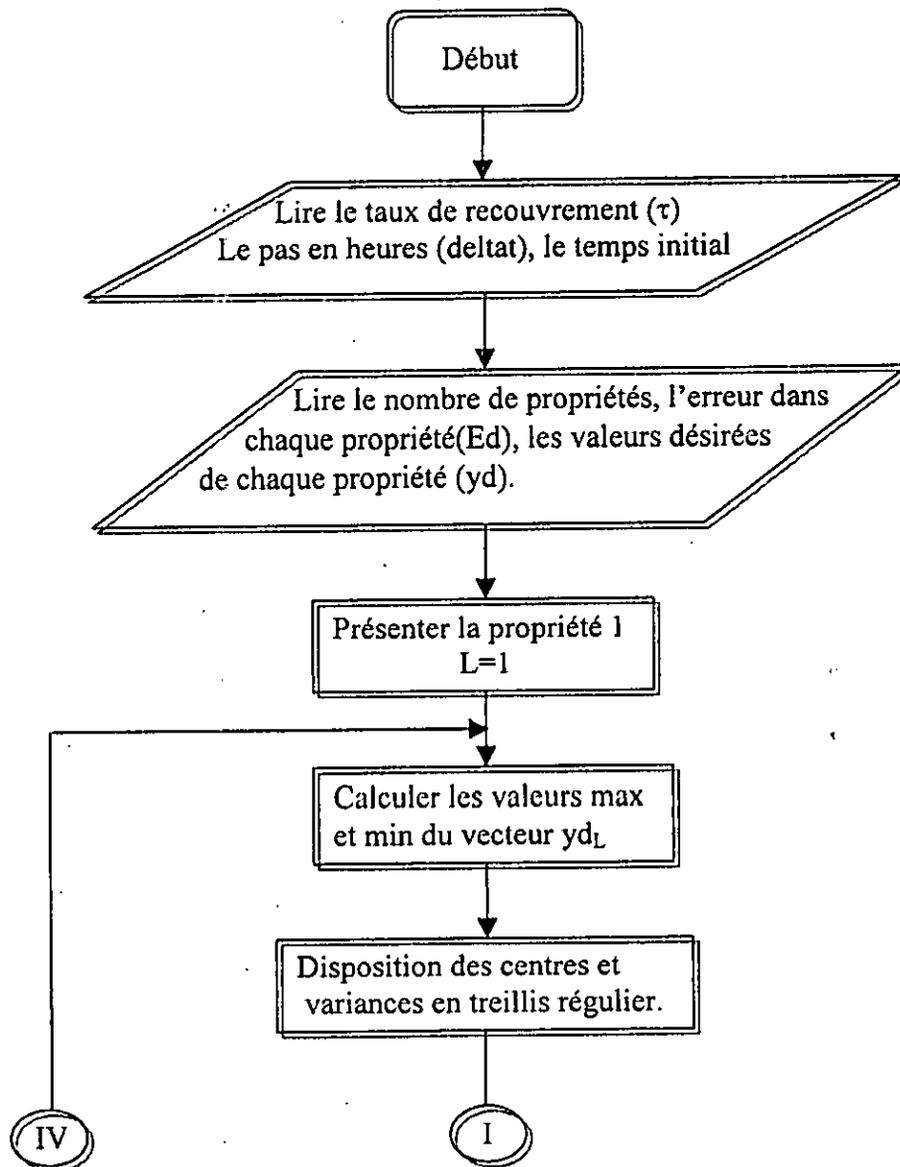


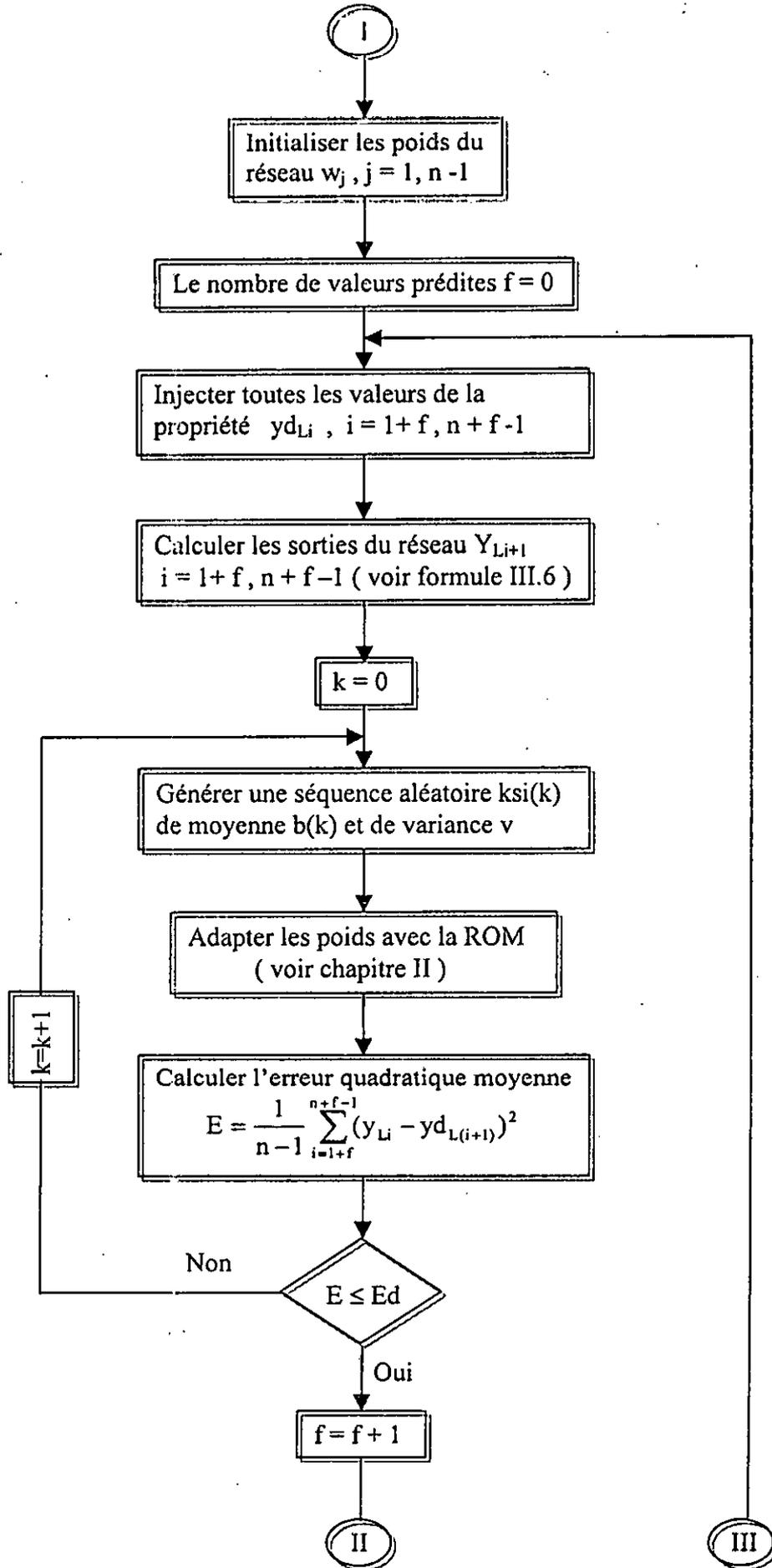
Fig.III.5. L'organigramme du réseau RBF 1

B) LE RESEAU RBF 2 :

On fait les même procédés que dans le premier cas. Mais cette fois ci, la base de temps est éliminée, et on travaille seulement avec les valeurs de la propriété. Les exemples d'entraînement seront $(y_1, y_2), (y_2, y_3), \dots, (y_k, y_{k+1}), \dots, (y_{n-1}, y_n)$. Après avoir terminé l'apprentissage, on présente, à l'entrée du réseau, la valeur y_n et on obtient à la sortie la $(n+1)^{\text{ème}}$ valeur qu'on considère comme valeur exacte. On ajoute cette dernière à la base de données et on refait l'apprentissage dès le début. On reprend ainsi les même étapes jusqu'à la $(n + m - 1)^{\text{ème}}$ valeur qu'on injectera, et la valeur reçu sera la dernière valeur qu'on doit prédire. Ainsi, tous les résultats obtenues seront affichés dans le fichier de résultats.

La structure générale de l'organigramme de ce réseau sera présentée dans la figure ci-dessous :





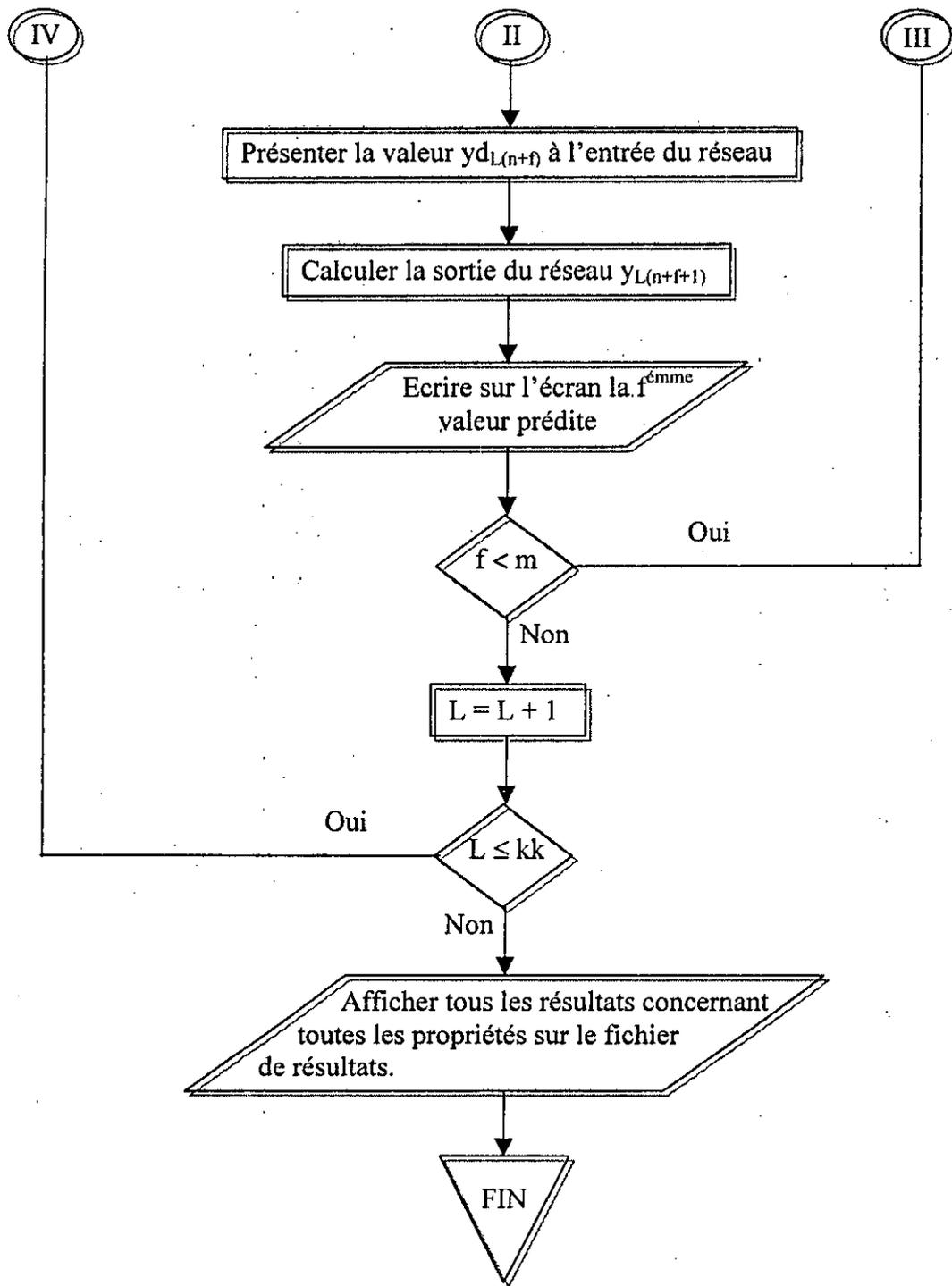
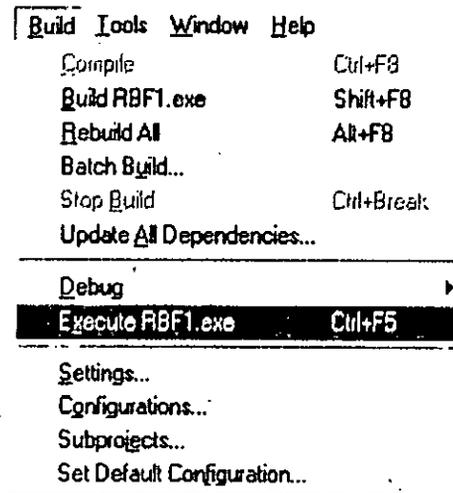


Fig.III.6. L'organigramme du réseau RBF 2

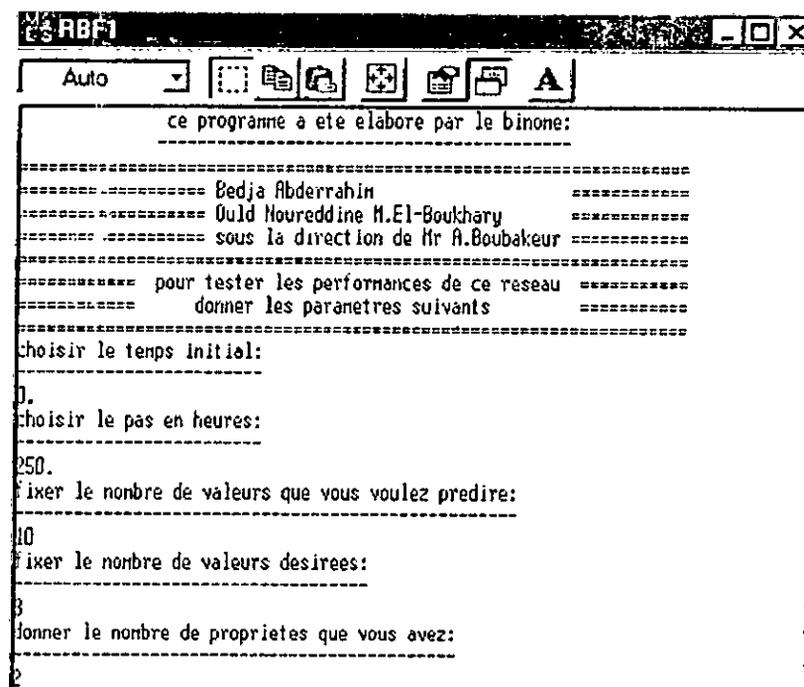
III.4. APPLICATION :

Afin d'illustrer les possibilités de prédiction de nos programmes, nous présentons un exemple du réseau RBF1 :

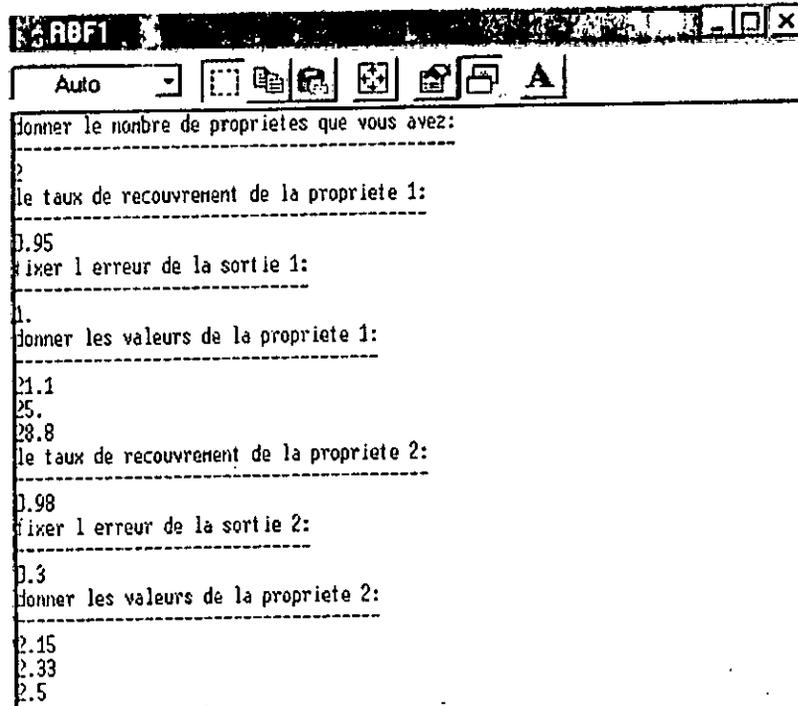
Pour lancer l'exécution du programme, l'utilisateur devra cliquer sur le menu « Build », et l'exécution se réalise en cliquant sur le sous menu « ExecuterBF1.exe ».



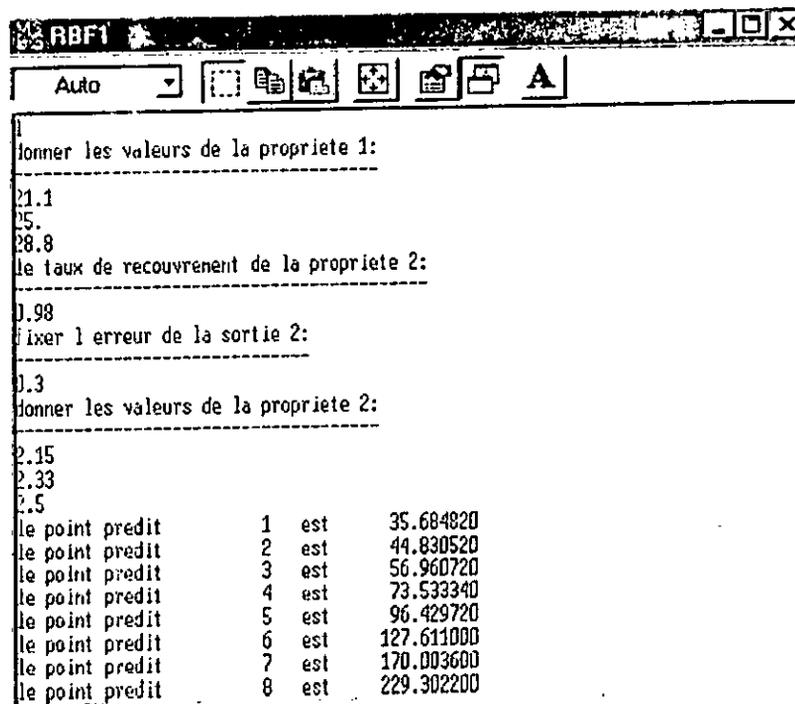
La fenêtre « DOS » apparaît sur votre écran, vous devez d'abord saisir les paramètres concernant le réseau :



En suite vous devez donner les valeurs des caractéristiques que vous avez :



Une fois que cela est fait, les valeurs des points prédits apparaissent sur l'écran noir, une à une :



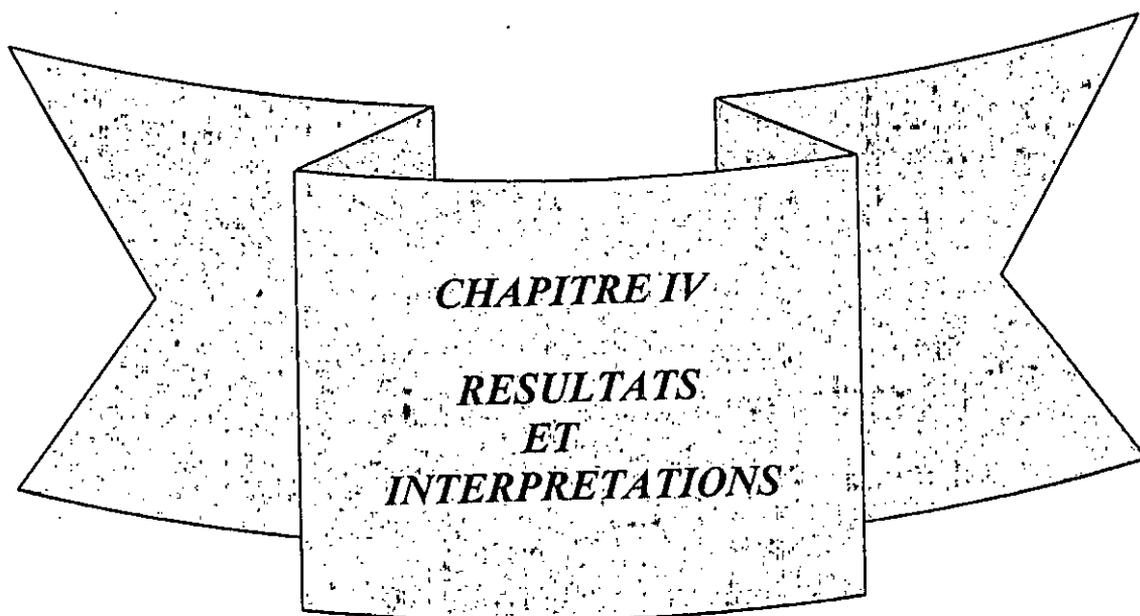
Après l'affichage de tous les résultats prédits, vous cliquez sur le fichier « fy.res » pour voir tous les résultats en détail concernant toutes les caractéristiques :

The screenshot shows a window titled "Microsoft Developer Studio - [fy.res]". The menu bar includes File, Edit, View, Insert, Build, Tools, Window, and Help. The toolbar contains various icons, and the text "fdelta" is visible. The main content area displays a table with three columns: "t (h)", "T=120°C", and "T=80°C". The table contains 11 rows of data, showing values for different time intervals from 0 to 3000 hours.

t (h)	T=120°C	T=80°C
.000000E+00	21.100000	2.150000
250.000000	25.000000	2.330000
500.000000	28.800000	2.500000
750.000000	35.684820	2.044403
1000.000000	44.830520	2.746620
1250.000000	56.960720	2.737622
1500.000000	73.533340	3.087525
1750.000000	96.429720	3.021047
2000.000000	127.611000	3.375192
2250.000000	170.003600	3.281958
2500.000000	229.302200	3.601253
2750.000000	314.512300	4.020838
3000.000000	438.370200	4.171863

III.5. CONCLUSION :

- Les réseaux RBF normalisés se distinguent par leur caractéristique de meilleurs approximateurs, et sont à l'abri de plusieurs problèmes que connaissent les autres réseaux.
- L'algorithme de « descente de gradient », qui dépend des taux d'apprentissage (η_w , η_c , η_σ) qu'on trouve certaines difficultés de les choisir pour être proche du minimum global, alors que la Méthode d'Optimisation Aléatoire est une technique qui dépend pas du gradient de l'erreur, d'où elle s'avère très efficace dans sa technique de recherche, pour atteindre le minimum global.
- Nous avons proposé une architecture avec une base d'apprentissage dynamique au cours de la prédiction, afin d'éviter certaines contraintes, comme le sur-apprentissage, et aussi d'obliger le réseau à se débarrasser de ses premiers souvenirs pour s'adapter avec les nouvelles valeurs prédites.
- Enfin, nous avons expliqué l'utilisation de nos programmes, l'un utilise le temps pour prédire les valeurs futures d'une propriété, et l'autre qui nie le temps et utilise seulement les valeurs de la propriété pour prédire les valeurs futures.

A decorative banner with a central rectangular section containing text. The banner has a wavy, ribbon-like appearance with pointed ends. The text is centered within the rectangular section.

CHAPITRE IV
RESULTATS
ET
INTERPRETATIONS

IV.1. INTRODUCTION :

Nous allons présenter dans ce chapitre quelques résultats de prédiction des deux programmes que nous avons élaborés et ce, en jouant sur les paramètres suivants :

- Le pas en heures.
- L'intervalle d'apprentissage.

Notre intérêt est porté sur la recherche des paramètres qui nous donnent une bonne prédiction tout en s'approchant de la pratique, c'est à dire, choisir un pas en heures où le manipulateur peut prendre des valeurs, et d'essayer d'optimiser le temps d'apprentissage de telle sorte qu'il soit le plus court possible, cela pour voir le comportement du réseau dans un intervalle élevé. Comme précisé dans le chapitre précédent, les résultats prédits vont être comparés aux valeurs de la courbe réelle. S'ils coïncident avec un certaine marge d'erreur, on juge la prédiction comme bonne, et que notre réseau a une certaine fiabilité de prédire des points à l'extérieur de la courbe expérimentale.

Notre étude est limitée sur une seule caractéristique liée aux pertes diélectriques en fonction du temps de vieillissement pour deux températures différentes (80°C et 120°C). Après l'exposition de tous les résultats, nous allons faire une comparaison entre les deux programmes et sélectionner lequel qui fait une bonne prédiction.

IV.2. RÉSULTATS LIÉS AU RÉSEAU RBF1 :

□ Vieillissement à $T = 120^{\circ}\text{C}$:

Nous avons pris trois pas en heures : 150 , 200 , 250 . Pour chaque pas, nous essayons de varier le temps d'apprentissage afin de l'optimiser.

➤ Pas de 150h :

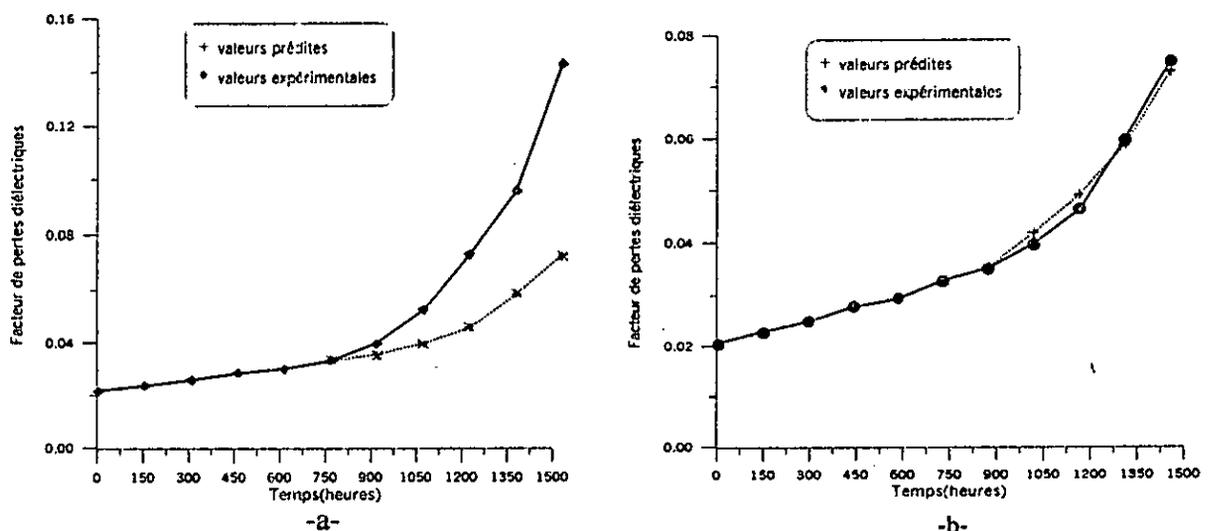


Fig.IV.1. Facteur de pertes diélectriques en fonction du temps de vieillissement à $T=120^{\circ}\text{C}$

- a- Temps d'apprentissage de 750h.
b- Temps d'apprentissage de 900h.

D'après ces résultats (Fig.IV.1), nous pouvons dire que la courbe (b) est celle qui correspond à une bonne approximation. Dans ce cas, notre réseau ne connaît la forme de la courbe réelle que lorsque la base de données soit élevée (900h), essayons maintenant de prendre un autre pas pour tester encore notre réseau.

➤ Pas de 200h :

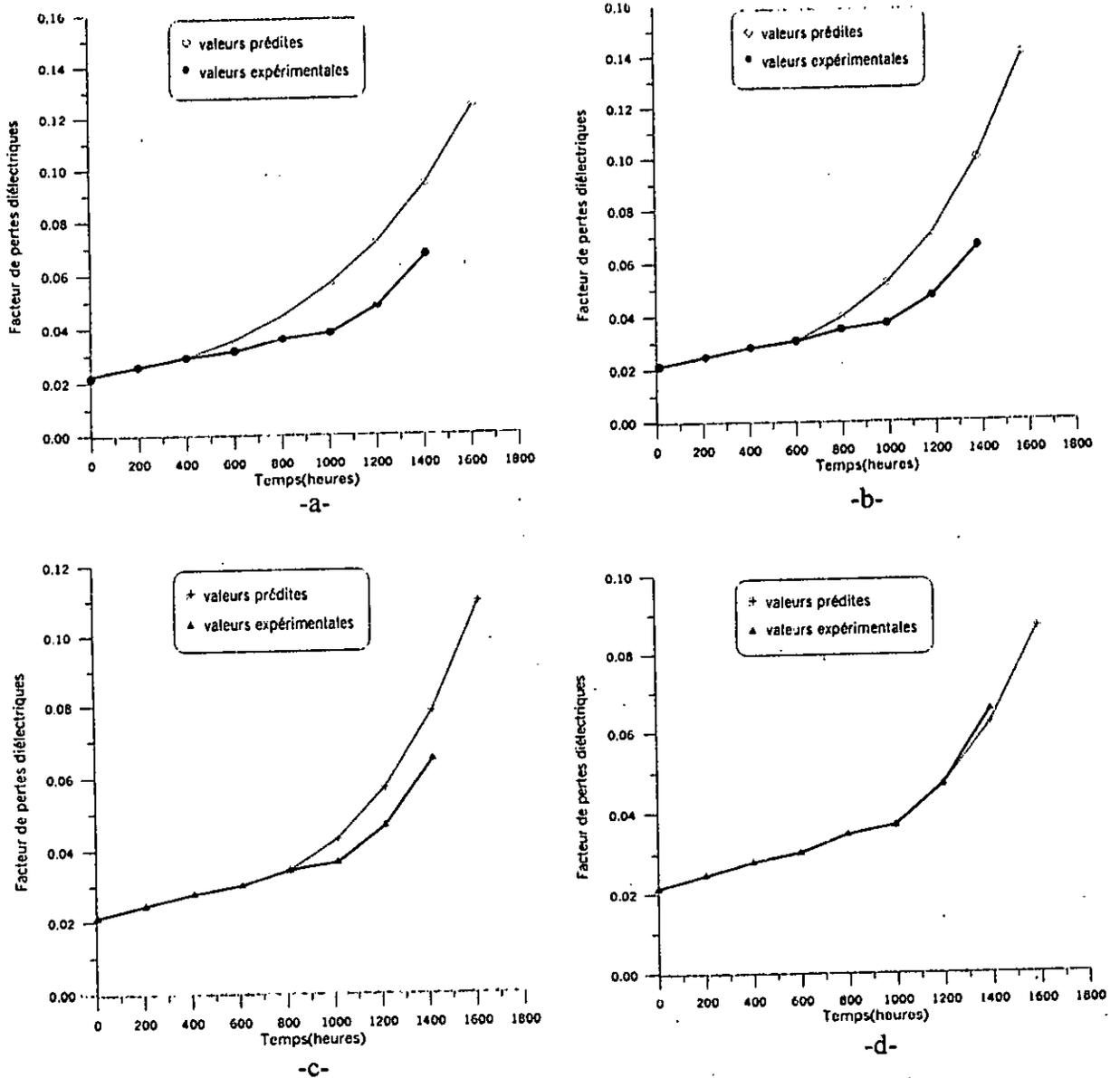


Fig.IV.2. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=120°C

- a- Temps d'apprentissage de 400h.
- b- Temps d'apprentissage de 600h.
- c- Temps d'apprentissage de 800h.
- d- Temps d'apprentissage de 1000h.

La même remarque que précédemment, le réseau suit convenablement la courbe réelle dans le cas de la figure(IV.2.d), où l'apprentissage a été effectué sur un temps de 1000h, par contre si nous diminuons ce temps comme le montre les autres figures, le réseau s'éloigne de la courbe réelle.

➤ Pas de 250h :

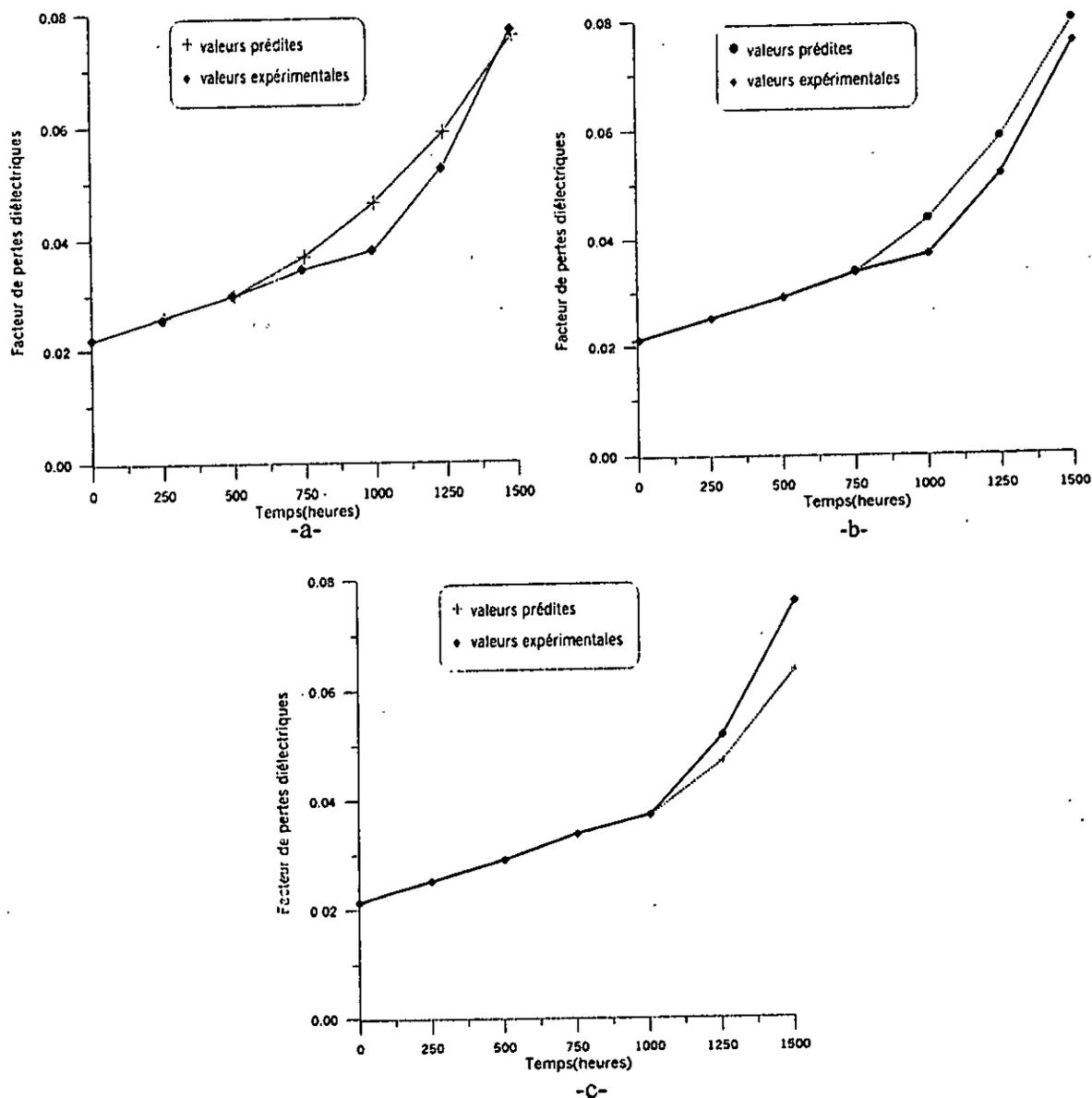


Fig.IV.3. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=120°C

- a- Temps d'apprentissage de 500h.
- b- Temps d'apprentissage de 750h.
- c- Temps d'apprentissage de 1000h.

Les courbes de la figure (IV.3), nous montrent qu'avec un temps d'apprentissage de 500h, les sorties du réseau suivent convenablement les résultats réels (figure(IV.3.a)).

□ Vieillessement à $T = 80^{\circ}\text{C}$:

Dans ce cas, nous avons pris trois pas en heures différents : 200 , 250 , 300.
 Pour chaque pas, nous essayons de chercher un temps d'apprentissage optimal.

➤ Pas de 200h :

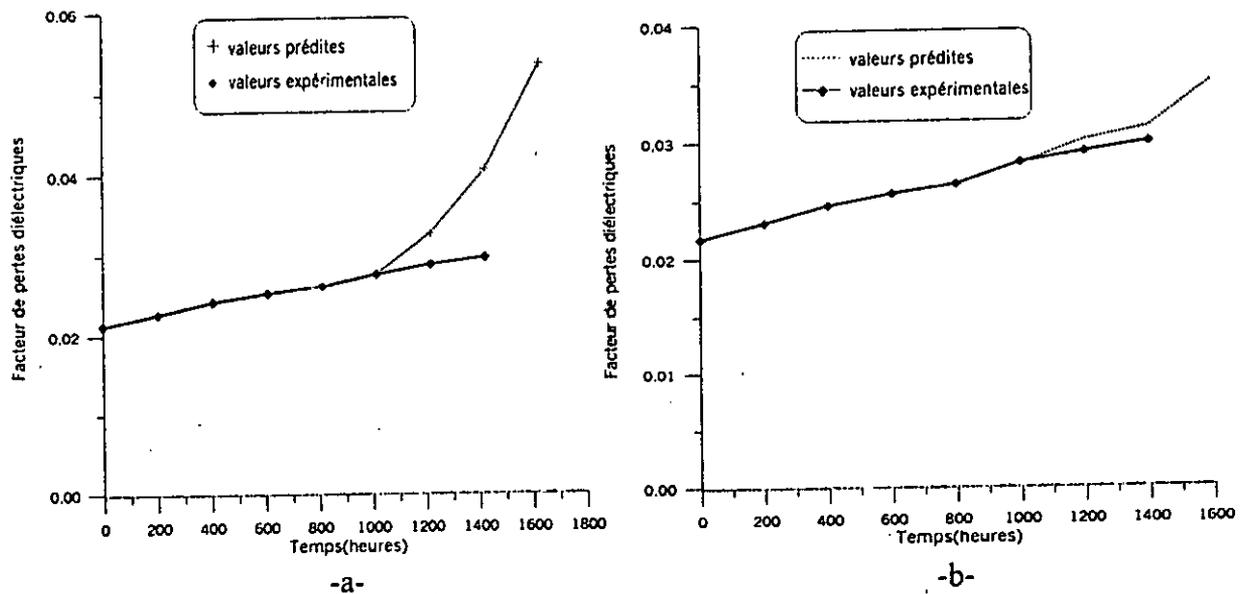


Fig.IV.4. Facteur de pertes diélectriques en fonction du temps de vieillissement à $T=80^{\circ}\text{C}$

- a- Temps d'apprentissage de 800h.
- b- Temps d'apprentissage de 1000h.

Les deux figures précédentes nous montre qu'il n'y a pas un temps d'apprentissage qui nous donne des résultats confondus avec la propriété réelle. Pour un intervalle de 1000h comme le montre la figure(IV.4.b), il y a une certaine approche mais pas suffisante pour juger la prédiction comme bonne.

➤ Pas de 250h :

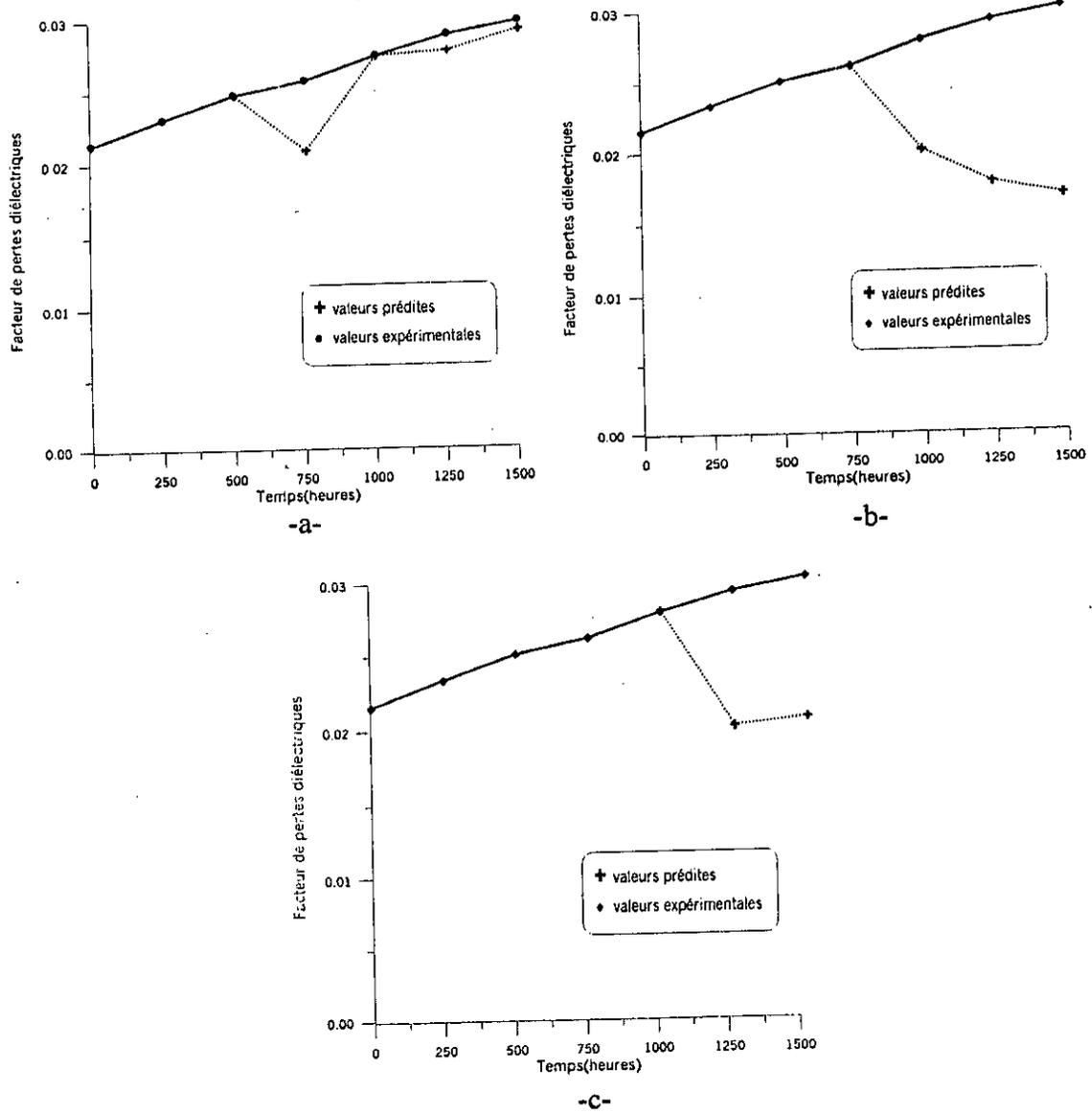


Fig.IV.5. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=80°C

- a- Temps d'apprentissage de 500h.
- b- Temps d'apprentissage de 750h.
- c- Temps d'apprentissage de 1000h.

D'après ces courbes (fig.IV.5), on constate une certaine dégradation des résultats obtenus dans le cas des courbes (b) et (c). Contrairement à la courbe (a), où le réseau a mal prédit la première valeur, mais il s'est amélioré lors de la prédiction des autres valeurs.

➤ Pas de 300h :

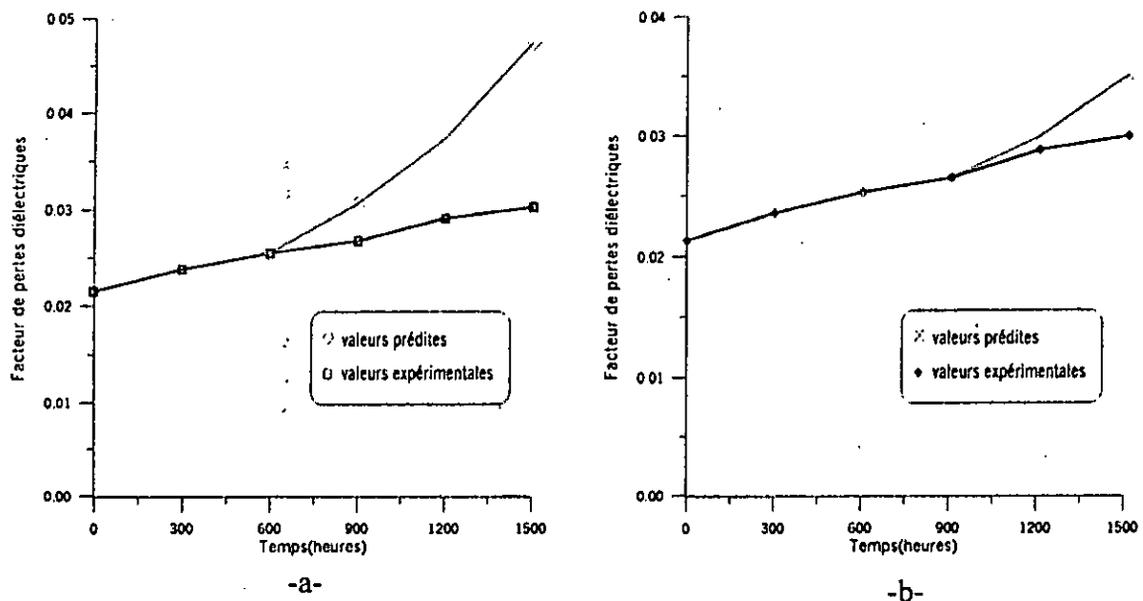


Fig.IV.6. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=80°C

- a- Temps d'apprentissage de 600h.
b- Temps d'apprentissage de 900h.

On remarque d'après ces courbes (Fig.IV.6) que le pas 300h n'a pas donné de bons résultats. Quelque soit le temps d'apprentissage pris, le réseau diverge toujours de la courbe réelle.

A partir de tous les résultats que nous avons vu, on peut dire que le réseau RBF1 est plus performant, pour les deux températures (80°C et 120°C), dans les cas suivants :

- ❖ Un pas en heures de 250.
- ❖ Un temps d'apprentissage de 500h.

Voici les résultats obtenus par le réseau RBF1 avec les paramètres cités dessus, mais la prédiction cette fois ci est prolongée jusqu'à 3000h :

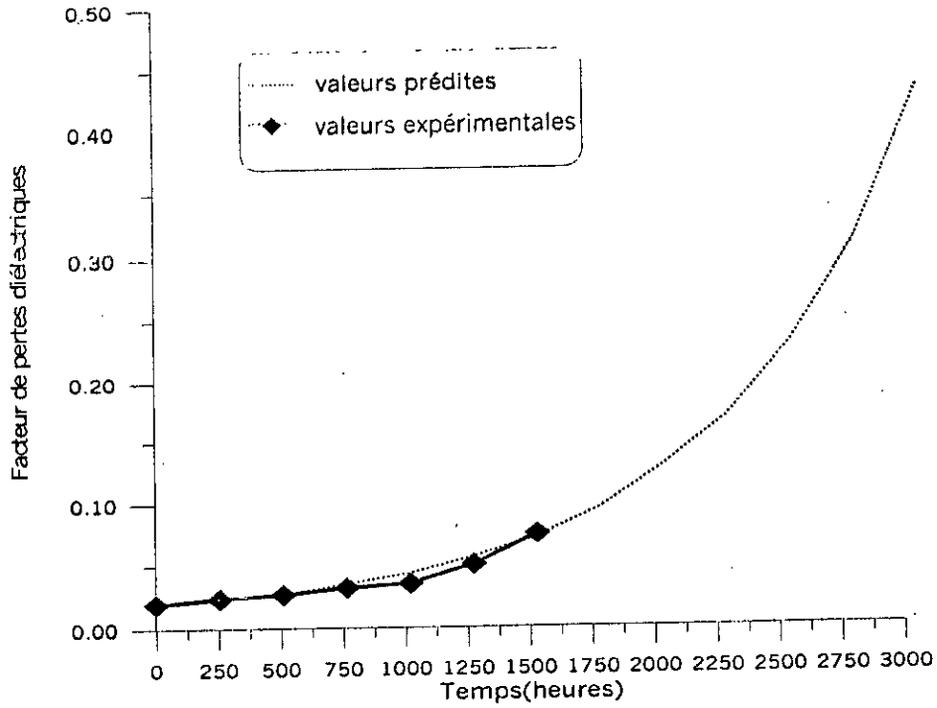


Fig.IV.7. Facteur de pertes diélectriques en fonction de temps de vieillissement (prédiction jusqu'à 3000h pour T= 120°C).

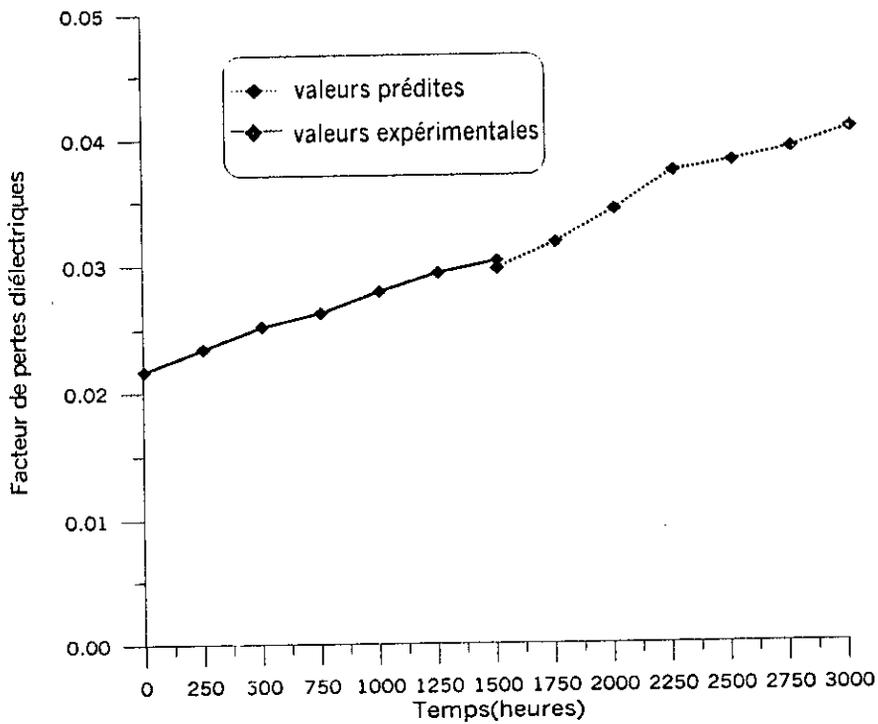


Fig.IV.8. Le facteur de pertes diélectriques en fonction du temps de vieillissement (prédiction jusqu'à 3000h pour T = 80°C).

IV.3. RÉSULTATS LIÉS AU RÉSEAU RBF2 :

Nous allons essayer d'utiliser les mêmes paramètres que dans le cas du réseau RBF1, cela pour pouvoir comparer entre les deux programmes.

□ Vicilissement à $T = 120^{\circ}\text{C}$:

➤ Pas de 150h :

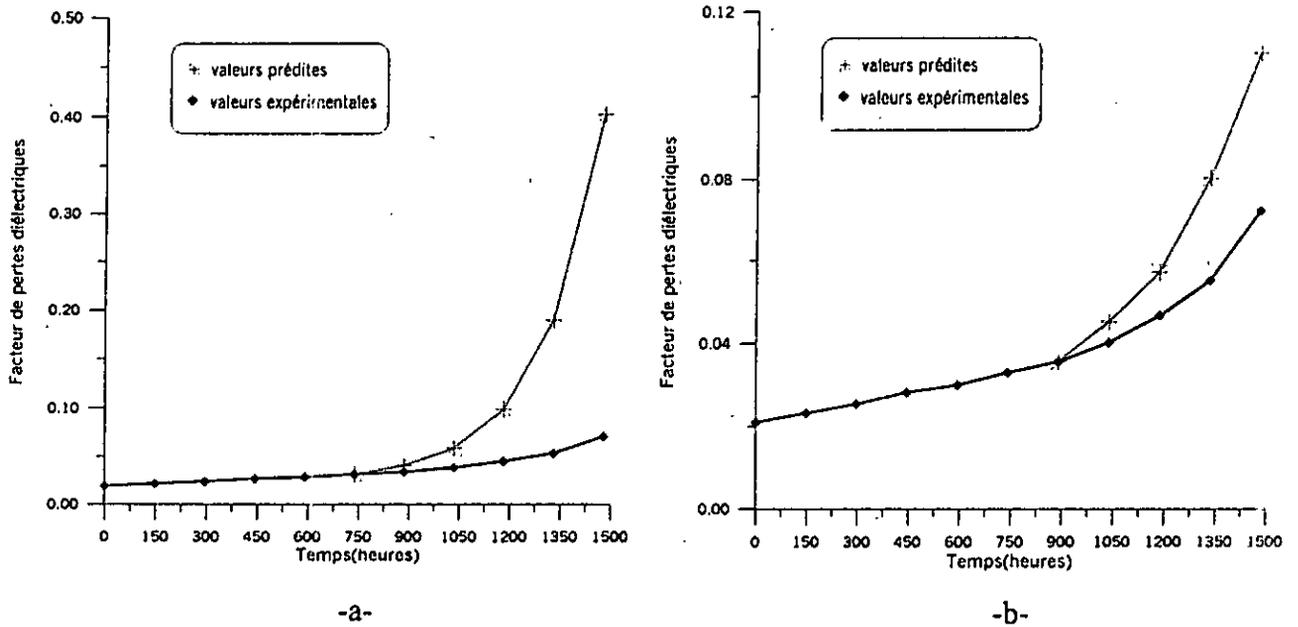


Fig.IV.9. Facteur de pertes diélectriques en fonction du temps de vieillissement à $T=120^{\circ}\text{C}$

a- Temps d'apprentissage de 750h.

b- Temps d'apprentissage de 900h.

On remarque d'après ces courbes (Fig.IV.9) une certaine divergence du programme, le réseau ne peut pas reconnaître l'allure exacte quelque soit l'intervalle d'apprentissage choisi.

➤ Pas de 200h :

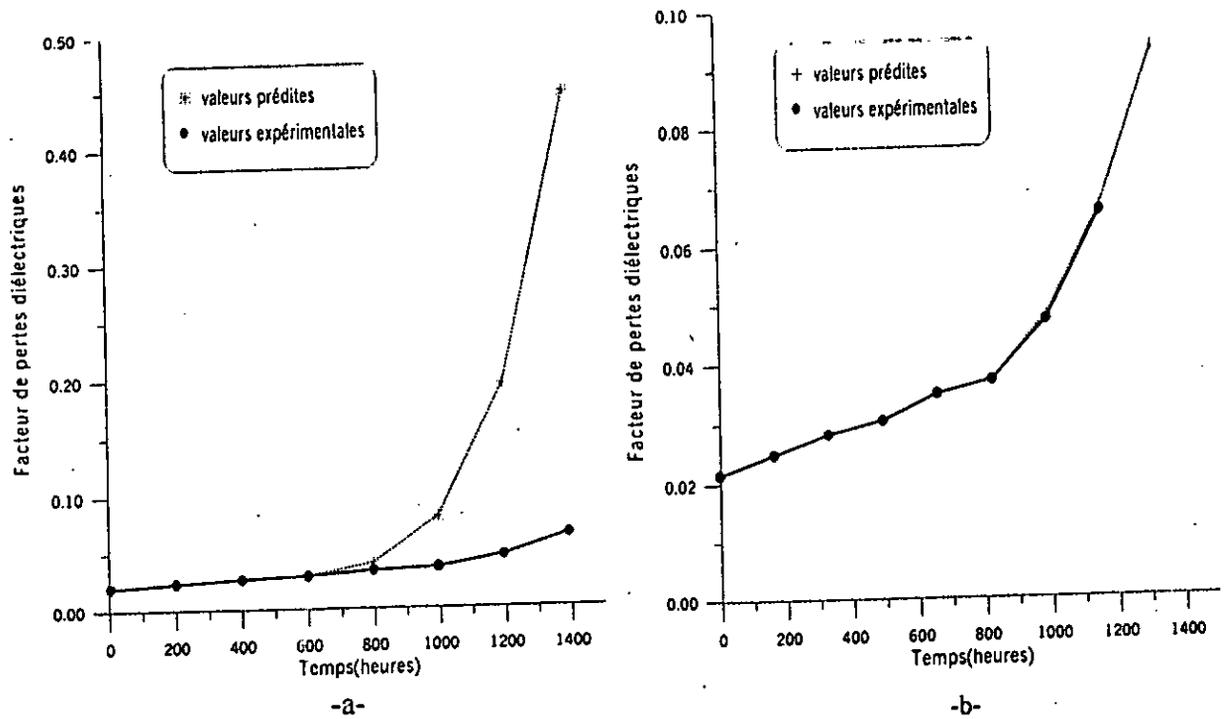


Fig.IV.10. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=120°C

- a- Temps d'apprentissage de 600h.
- b- Temps d'apprentissage de 1000h.

Pour un pas de 200h (Fig.IV.10), on constate que pour une base de données élevée (un temps de 1000h), comme le montre la courbe(b), les résultats obtenus par le réseau RBF2 coïncident avec ceux de la propriété réelle.

➤ Pas de 250 :

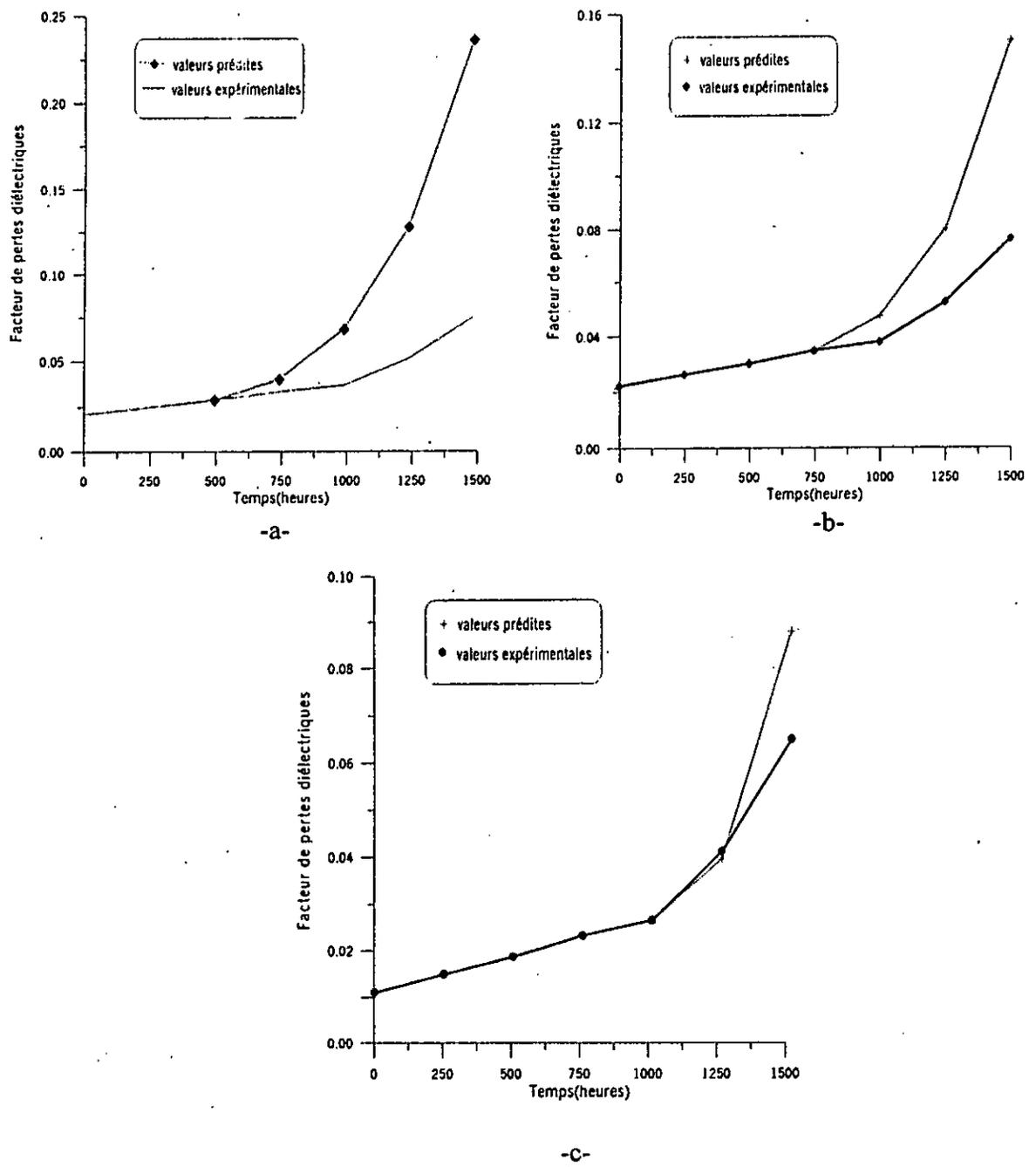


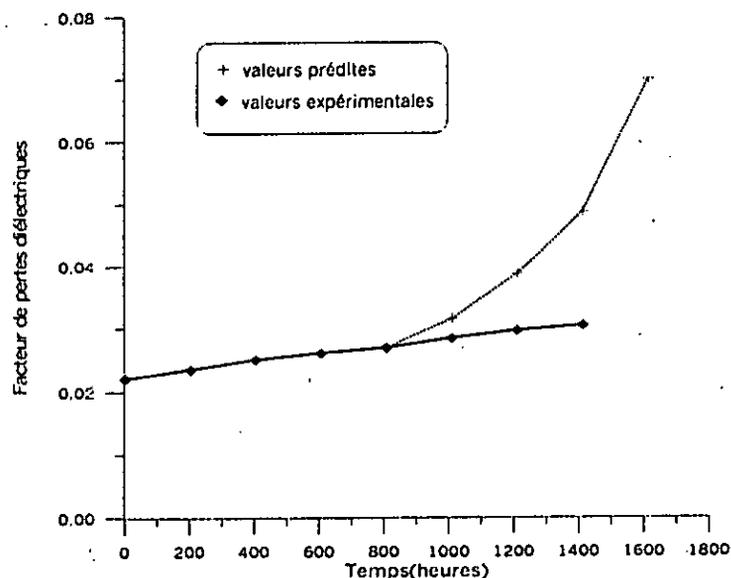
Fig.IV.11. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=120°C

- a- Temps d'apprentissage de 500h.
- b- Temps d'apprentissage de 750h.
- c- Temps d'apprentissage de 1000h.

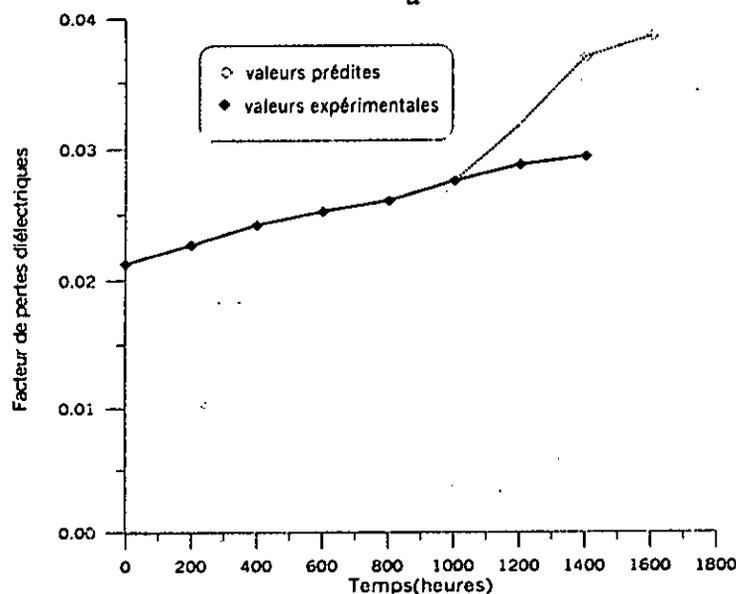
Comme nous remarquons dans ces courbes (Fig.IV.11), pour un pas de 250h on peut pas trouver une base de données permettant d'arriver à une bonne prédiction. Les résultats obtenus divergent toujours de la caractéristique réelle.

□ Vieillissement à $T = 80^{\circ}\text{C}$:

➤ Pas de 200h :



-a-



-b-

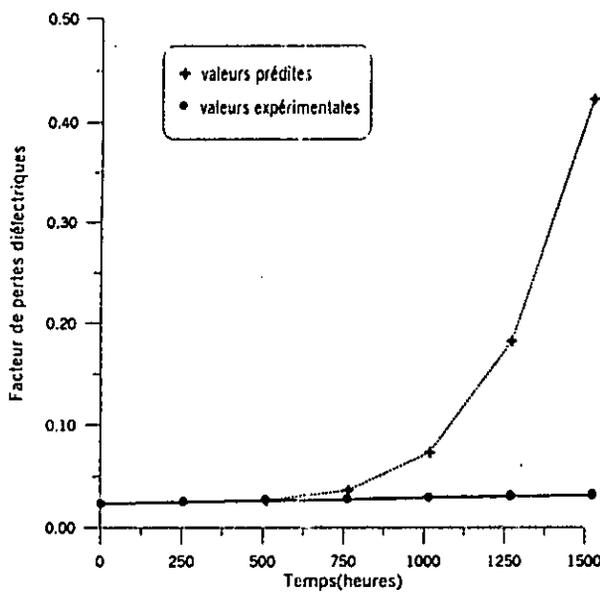
Fig.IV.12. Facteur de pertes diélectriques en fonction du temps de vieillissement à $T=80^{\circ}\text{C}$

a- Temps d'apprentissage de 800h.

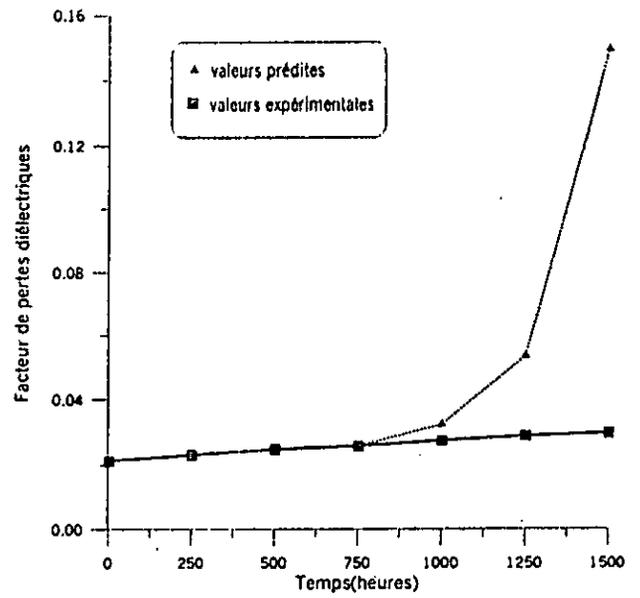
b- Temps d'apprentissage de 1000h.

La figure (Fig.IV.12) nous montre que le réseau RBF2 diverge de la propriété réelle quelque soit le temps d'apprentissage pris.

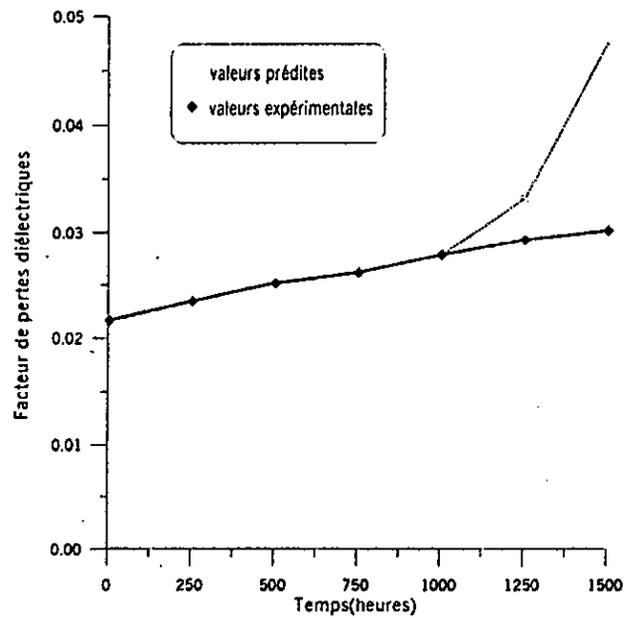
➤ Pas de 250h :



-a-



-b-



-c-

Fig.IV.13. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=80°C

- a- Temps d'apprentissage de 500h.
- b- Temps d'apprentissage de 750h.
- c- Temps d'apprentissage de 1000h.

Les courbes (a), (b) et (c) de la figure (IV.13), nous montre que pour un pas de 250h, il y a pas un temps d'apprentissage permettant d'avoir une convergence du réseau vers la sortie réelle.

➤ Pas de 300h :

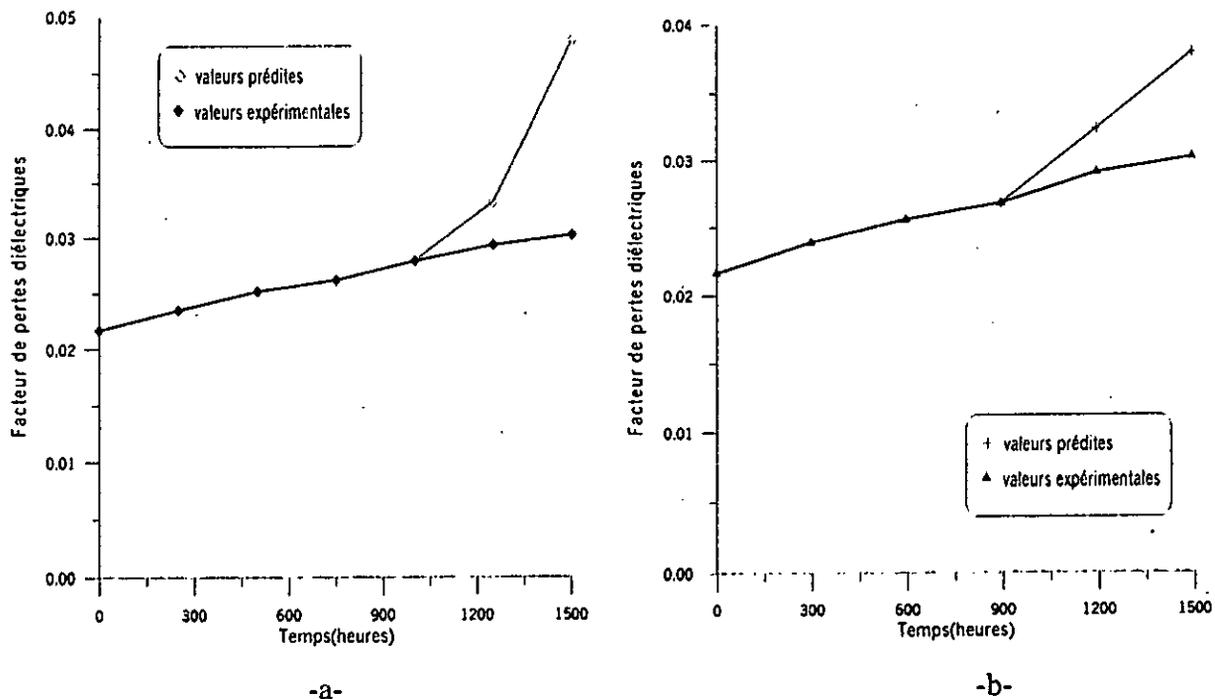


Fig.IV.14. Facteur de pertes diélectriques en fonction du temps de vieillissement à T=80°C

- a- Temps d'apprentissage de 600h.
- b- Temps d'apprentissage de 900h.

La même remarque que dans les cas précédents, le réseau diverge de la propriété réelle (Fig.IV.14). D'où, pour la température 80°C, le réseau RBF2 n'a pas donné des résultats encourageants pour pouvoir prolonger la prédiction jusqu'aux intervalles de temps élevés. Les performances du réseau sont limitées seulement pour la température 120°C et avec les paramètres suivants :

- ❖ Un pas en heures de 200h.
- ❖ Un temps d'apprentissage de 1000h.

Voici les résultats obtenus avec ces paramètres, et la prédiction est prolongée jusqu'à 3000h :

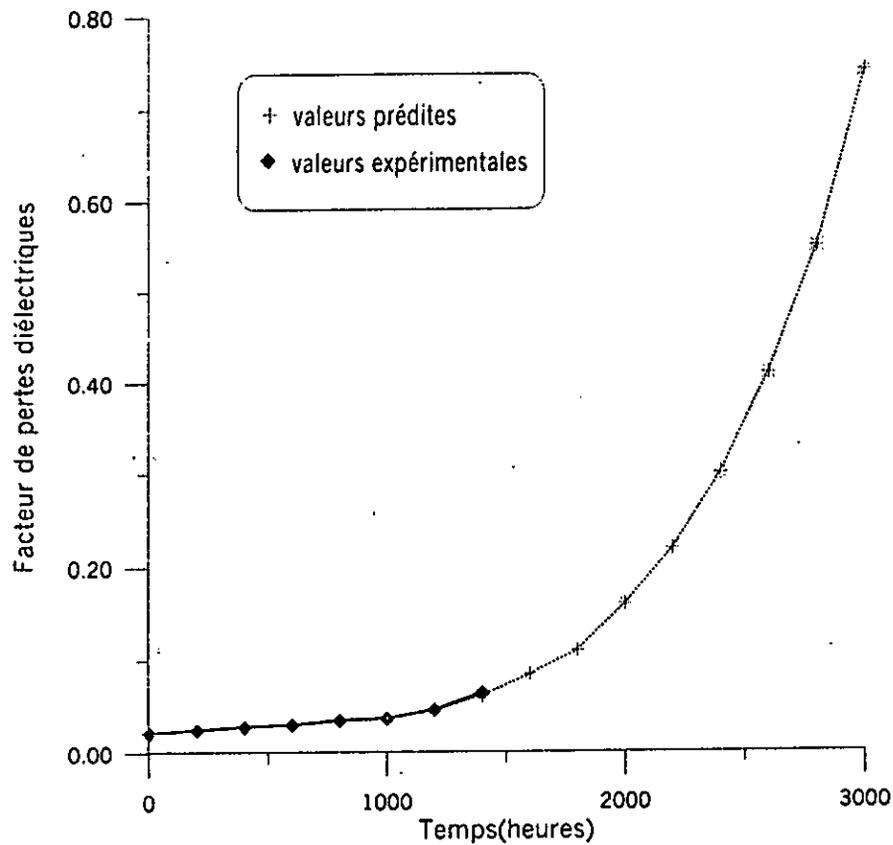


Fig.IV.15. Le facteur de pertes diélectriques en fonction de temps de Vieillessement (prédiction jusqu'à 3000h pour $T = 120^{\circ}\text{C}$).

IV.4. COMPARAISON ENTRE LES DEUX PROGRAMMES :

- Le réseau RBF1 nous a donné de bons résultats pour plusieurs paramètres, même s'il s'agit d'un temps d'apprentissage très court où on peut bien voir le comportement du réseau par rapport à la caractéristique réelle. Ce programme a convergé vers la sortie réelle dans les deux températures (80°C et 120°C) sans qu'il y ait de dégradation, on peut dire alors qu'avec le réseau RBF1 on trouve une certaine souplesse de choisir les paramètres convenables permettant d'avoir une bonne prédiction.
- Pour le réseau RBF2, un intervalle d'entraînement court, affecte les résultats de sortie de celui-ci. Par conséquent on n'a pas un intervalle de temps élevé, pour le tester en comparant son comportement par rapport à la caractéristique expérimentale. A chaque fois qu'on essaye de diminuer la base de données, les résultats prédits divergent de la sortie réelle et le réseau ne pourra pas connaître l'évolution de la propriété. Nous avons trouvé les performances de ce programme avec le temps d'apprentissage de 1000h et avec un pas de 200h pour la seule température $T=120^{\circ}\text{C}$ (Fig.IV.15).

- Voici les résultats rapportés sur le même graphe des deux programmes dont les paramètres sont choisis où ces derniers sont performants, et la prédiction est prolongée jusqu'à 3000h :

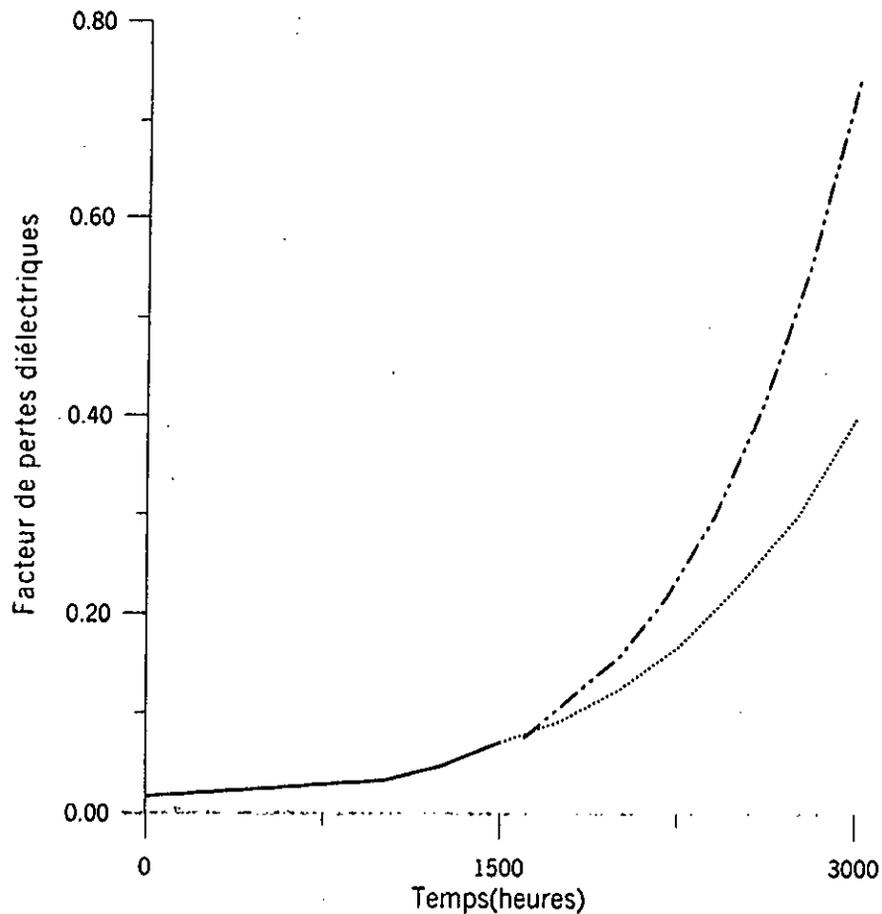


Fig.IV.16. Facteur de pertes diélectriques en fonction de temps de vieillissement dans le cas des deux programmes (RBF1 et RBF2).

- Courbe réelle
- ⋯ La sortie du réseau RBF1 avec un pas de 250h et un temps d'apprentissage de 500h.
- - - La sortie du réseau RBF2 avec un pas de 200h et un temps d'apprentissage de 1000h.

La figure (IV.16) nous montre que :

- Le réseau RBF2 a divergé, et a mal prédit la suite de la caractéristique réelle.
- La réseau RBF1 a bien suivi l'allure de la caractéristique expérimentale.

IV.5. INTERPRETATIONS :

Avec toutes les contraintes rencontrées dans le cas du réseau RBF2 pour prédire des intervalles de temps élevés, on peut déduire que le réseau RBF2 n'a pas trouvé une dépendance entre les valeurs de la propriété qu'il a appris pour bien prédire les valeurs futures. Par contre le réseau RBF1, avec l'introduction du facteur temps, il a pu faire une dépendance avec le temps et les valeurs de la caractéristique permettant ainsi d'améliorer la qualité de prédiction et s'approcher de la caractéristique réelle, pour bien justifier cela, prenons l'exemple suivant :

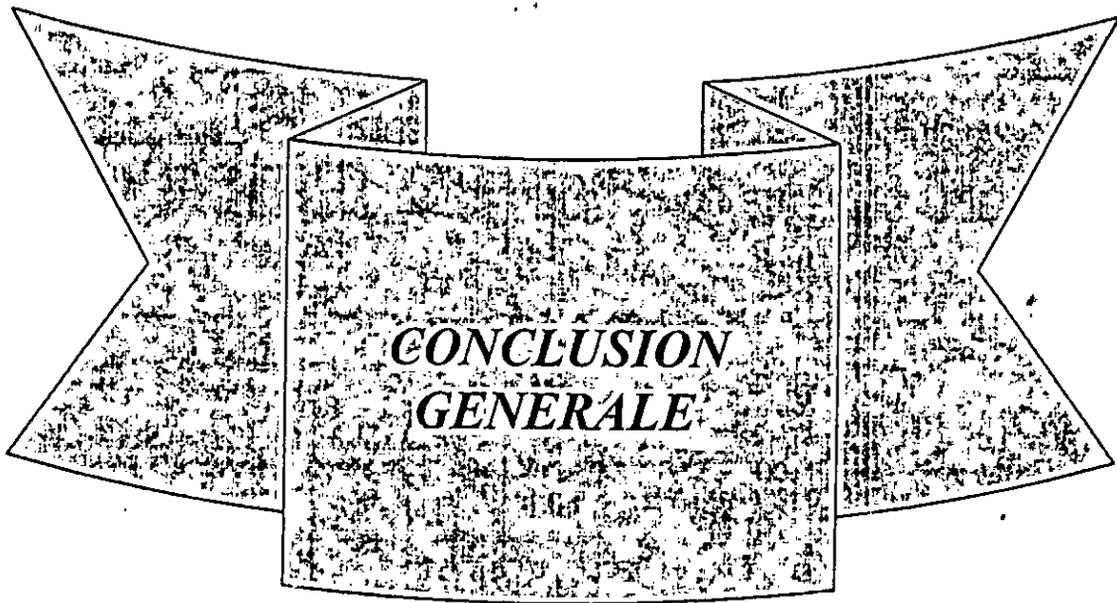
Nous considérons la fonction $y_i = t_i^2$ où les valeurs de la propriété y_i dépendent toujours du temps t_i , essayant maintenant de déterminer y_{i+1} en fonction de y_i :

$$\begin{aligned}y_{i+1} &= f(t_{i+1}) = f(t_i + A); \text{ avec } A \text{ est le pas en heures.} \\ &= (t_i + A)^2 \\ y_{i+1} &= y_i + 2At_i + A^2\end{aligned}$$

Ainsi la valeur y_{i+1} dépend non seulement de la valeur y_i , mais aussi du temps t_i . Et on ne peut pas déterminer ce temps en fonction de la valeur y_i , car, le temps n'a aucune dépendance de la propriété. Par conséquent, il n'y a pas une dépendance entre les valeurs de la propriété sans introduire le facteur temps.

IV.6. CONCLUSION :

- Après la discussions de tout les résultats obtenues par ces deux programmes, nous pouvons sélectionner le programme RBF1 comme meilleur approximateur de propriétés. Avec l'injection du temps à l'entrée du réseau et l'adaptation des sorties de ce dernier aux valeurs de la courbe réelle, on pourra bien avoir une approche de la caractéristique expérimentale en effectuant l'apprentissage avec un intervalle de temps relativement court.
- Comme nous avons observé dans les deux cas de températures (80°C et 120°C), ce programme (RBF1) peut prédire le triple de l'intervalle appris avec une marge d'erreur acceptable, est dans les environs de 3%, et les données doivent être espacées d'un pas de 150 à 300h.



CONCLUSION GENERALE

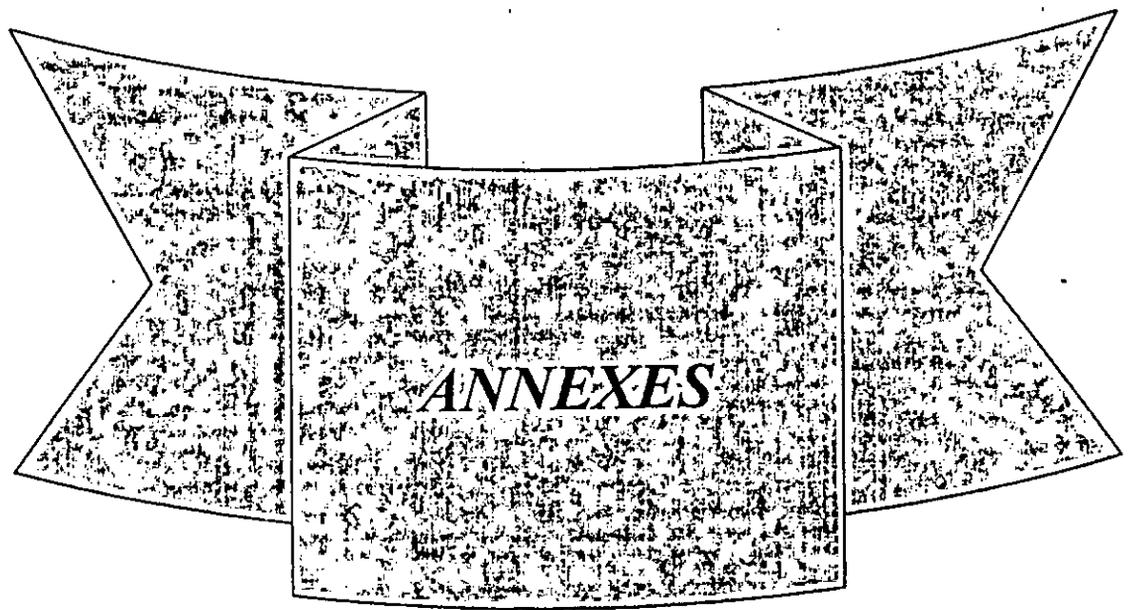
Le travail que nous avons effectué consiste en la prédiction de l'évolution de certaines propriétés d'une huile isolante, soumise à un vieillissement thermique.

Dans notre travail, nous avons mené une large recherche bibliographique qui nous a permis d'explorer les différents types de réseaux de neurones. Ceci afin d'aboutir à un réseau le plus robuste dont ses caractéristiques nous permettront de prédire les différentes propriétés de la « BORAC 22 » vieillie à 80°C et à 120°C. Nous avons testé deux programmes, l'un utilise le temps pour prédire les valeurs futures de la propriété, et l'autre ignore le temps et effectue la prédiction en connaissant seulement les valeurs de la caractéristique. Les résultats obtenus, nous ont permis d'aboutir aux conclusions suivantes :

- Les réseaux à base radiale normalisés se distinguent par leur caractéristique de meilleurs approximateurs de fonctions.
- La méthode d'optimisation aléatoire « ROM » est la meilleure technique d'apprentissage. En effet, elle ne dépend pas du système pour lequel le réseau est utilisé. De plus, sa technique de recherche permet d'atteindre un minimum globale en évitant ainsi toute divergence éventuelle.
- Pour les allures des propriétés traitées (80°C et 120°C), le programme **RBF1** peut prédire le triple de l'intervalle appris avec une marge d'erreur environs de 3%. Et les valeurs de la base de données doivent être espacées d'un pas de 150 à 300h, ce qui arrange ainsi le manipulateur de pouvoir faire les essais.
- Dans le cas du réseau **RBF2**, la qualité de prédiction se dégrade quand on essaye de diminuer le temps d'apprentissage, ceci nous oblige d'entraîner le système sur des intervalles de temps élevés pour que le réseau puisse trouver une dépendance entre les valeurs qu'il a appris. Ce qui affecte l'objectif de notre travail qui consiste à effectuer l'apprentissage sur des petits intervalles de temps, et voir le comportement du réseau dans un intervalle de prédiction élevé.

Les résultats obtenus par nos programmes sont toujours améliorables. Comme suite à notre travail, nous proposons dans les études à venir ce qui suit :

- Vu la dépendance existant entre les propriétés de l'huile, on pourra pas prédire un changement de l'allure d'une propriété sans connaître, en plus du temps, les autres propriétés avec lesquelles cette caractéristique est dépendante. Malheureusement, on a pas toutes les propriétés du vieillissement thermique qui dépendent des pertes diélectriques pour les injecter avec le temps. Dans ce cas, le réseau prédit non seulement l'évolution de ces propriétés, mais aussi acquiert une certaine dépendance entre toutes les propriétés. Il deviendra ainsi un prédicteur proche des caractéristiques réelles en détectant tout changement de forme des allures au long terme.



Annexe 1 : RBF Généralisé

Soit $J = \frac{1}{2}(y - y^*)^2$ le critère à minimiser, telle que y^* est donnée par l'équation suivante :

$$y^* = \frac{\sum_{j=1}^n w_j \rho_j(u)}{\sum_{j=1}^n \rho_j(u)} \quad \text{avec} \quad \rho_j(u) = \exp\left(-\frac{1}{2} \sum_{i=1}^m \frac{(u_i - c_{j,i})^2}{\sigma_{j,i}^2}\right)$$

Posons $R(u) = \sum_{j=1}^n \rho_j(u)$, ($R(u)$: est la facteur de normalisation), où n est le nombre de centres.

$$\frac{\partial J}{\partial w_j} = (y - y^*) \frac{\rho_j(u)}{R(u)}$$

$$\frac{\partial J}{\partial c_{j,i}} = \frac{\partial J}{\partial \rho_j} \frac{\partial \rho_j}{\partial c_{j,i}}$$

$$\frac{\partial J}{\partial \sigma_{j,i}} = \frac{\partial J}{\partial \rho_j} \frac{\partial \rho_j}{\partial \sigma_{j,i}}$$

$$\frac{\partial J}{\partial \rho_j} = (y - y^*) \left[\frac{w_j R(u) - \sum_{i=1}^n \rho_i(u) w_i}{R^2(u)} \right]$$

$$\frac{\partial \rho_j}{\partial c_{j,i}} = \rho_j \frac{u_i - c_{j,i}}{\sigma_{j,i}^2}$$

$$\frac{\partial \rho_j}{\partial \sigma_{j,i}} = \rho_j \frac{(u_i - c_{j,i})^2}{\sigma_{j,i}^3}$$

Les corrections s'obtiennent par :

$$\Delta w_j = -\eta_w \frac{\partial J}{\partial w_j}$$

$$\Delta c_{j,i} = -\eta_c \frac{\partial J}{\partial c_{j,i}}$$

$$\Delta \sigma_{j,i} = -\eta_\sigma \frac{\partial J}{\partial \sigma_{j,i}}$$

Annexe 2 : Le programme RBF1

```

Real yd(200,200), sigma(5), to(5), ed(5), t1(200), ydd(200,200)
Real c(200), sr, sb, b4(200,200), w(200,200), roy(200,200), ry(200)
Real xl(200,200), y(200,200), ea(100,100), eao(100,100)
Real ksi(200,200), s(150), gr, tb, gps(150), e2(200)
integer f, j1, a, lg
write(*,*) '      ce programme a ete elabore par le binome: '
write(*,*) '-----'
write(*,*) '===== Bedja Abderrahim ====='
write(*,*) '===== Ould Noureddine M.El-Boukhary '
write(*,*) '===== sous la direction de Mr A.Boubakeur ====='
write(*,*) '-----'
write(*,*) '=== pour tester les performances de ce reseau ==='
write(*,*) '===== donner les parametres suivants ====='
write(*,*) '-----'
write(*,*) 'choisir le temps initial:'
write(*,*) '-----'
read(*,*) ti
write(*,*) 'choisir le pas en heures:'
write(*,*) '-----'
read(*,*) deltat
write(*,*) 'fixer le nombre de valeurs a predire:'
write(*,*) '-----'
read(*,*) m
write(*,*) 'fixer le nombre de valeurs desirees:'
write(*,*) '-----'
read(*,*) n
write(*,*) 'donner le nombre de proprietes que vous avez:'
write(*,*) '-----'
read(*,*) kk
write(*,*) 'le taux de recouvrement de la propriete 1:'
write(*,*) '-----'
read(*,*) to(1)
write(*,*) 'fixer l erreur de la sortie 1:'
write(*,*) '-----'
read(*,*) ed(1)
write(*,*) 'donner les valeurs de la propriete 1:'
write(*,*) '-----'
do j=1,n
read(*,*) yd(1,j)
enddo
write(*,*) 'le taux de recouvrement de la propriete 2:'
write(*,*) '-----'
read(*,*) to(2)
write(*,*) 'fixer l erreur de la sortie 2:'
write(*,*) '-----'
read(*,*) ed(2)
write(*,*) 'donner les valeurs de la propriete 2:'
write(*,*) '-----'
do j=1,n
read(*,*) yd(2,j)
enddo
l=n
open(1, file='fy.res', status='unknown')

```

```
lk=1
21 continue
k=0
do 10 h=ti,ti+(n+m+1)*deltat,deltat
k=k+1
t1(k)=h
10 continue
f=0
350 continue
v=.3
valmax=t1(n+f)
valmin=t1(1+f)
deltal=(valmax-valmin)/l
sigma(lk)=sqrt(-(deltal)**2/(8*log(to(lk))))
g=valmax
do 20 j1=1,l
c(j1)=g+deltal
g=g-deltal
20 continue
k=0
do 30 k1=1,l
call random(sr)
w(lk,k1)=sr
30 continue
eao(lk,1)=0
300 continue
lg=1
do 40 k2=1,l
call random(sb)
ksi(lk,k2)=v*sb+b4(lk,k2)
40 continue
do 50 k3=1,3
if(k3.eq.1)then
a=0
else
if(k3.eq.2)then
a=1
else
a=-1
endif
endif
do 60 k4=1,l
w(lk,k4)=w(lk,k4)+a*ksi(lk,k4)
60 continue
do 70 i5=1+f,n+f+1
do 80 j5=1,l
roy(i5,j5)=exp(-(t1(i5)-c(j5))**2/(2*(sigma(lk))**2))
80 continue
70 continue
do 85 i6=1+f,n+f
ry(i6)=0
do 90 j6=1,l
ry(i6)=ry(i6)+roy(i6,j6)
90 continue
s(i6)=ry(i6)
85 continue
do 95 i7=1+f,n+f
x1(lk,i7)=0
do 96 j7=1,l
x1(lk,i7)=x1(lk,i7)+w(lk,j7)*roy(i7,j7)
96 continue
```

```

      y(lk,i7)=x1(lk,i7)/s(i7)
95  continue
      ! on definit l'erreur d'apprentissage par e
      e2(lk)=0
      do 100 k6=1+f,n+f
      e2(lk)=e2(lk)+((y(lk,k6)-yd(lk,k6))**2/n)
100 continue
      ea(lk,k3)=sqrt(e2(lk))
      50 continue
      if(ea(lk,2).lt.ea(lk,1))then
      do 150 i7=1,1
      w(lk,i7)=w(lk,i7)+ksi(lk,i7)
      b4(lk,i7)=.4*ksi(lk,i7)+.2*b4(lk,i7)
150 continue
      else if((ea(lk,2).ge.ea(lk,1)).and.(ea(lk,3).lt.ea(lk,1)))then
      do 200 i8=1,1
      w(lk,i8)=w(lk,i8)-ksi(lk,i8)
      b4(lk,i8)=b4(lk,i8)-.4*ksi(lk,i8)
200 continue
      else
      do 250 i9=1,1
      b4(lk,i9)=.5*b4(lk,i9)
250 continue
      endif
      endif
      if(ea(lk,1).gt.ed(lk))then
      k=k+1
      lg=lg+1
      eao(lk,lg)=ea(lk,1)
      if((eao(lk,lg)-eao(lk,lg-1)).eq..0)then
      if(v.ge..0)then
      v=v-.1
      else
      v=.01
      endif
      endif
      eao(lk,lg-1)=eao(lk,lg)
      goto 300
      else
      lm=0
      do 1 gk=valmax,valmax-2*delta1,-delta1
      lm=lm+1
1  gps(lm)=exp(-(gk-valmax)**2/(2*sigma(lk)**2))
      if(y(lk,n-1).le.y(lk,n))then
      do 2 j=1,3
2  w(lk,j)=w(lk,j)/gps(j)
      else
      do 3 j=1,3
3  w(lk,j)=w(lk,j)*gps(j)
      endif
      f=f+1
      do 1000 j5=1,1
      roy(n+f,j5)=exp(-(t1(n+f)-c(j5))**2/(2*(sigma(lk))**2))
1000 continue
      ry(n+f)=0
      do 105 j6=1,1
      ry(n+f)=ry(n+f)+roy(n+f,j6)
105 continue
      gr=ry(n+f)
      x1(lk,n+f)=0
      do 110 j7=1,1

```

```
x1(lk,n+f)=x1(lk,n+f)+w(lk,j7)*roy(n+f,j7)
110 continue
  tb=x1(lk,n+f)
  y(lk,n+f)=tb/gr
  yd(lk,f+n)=y(lk,f+n)
  endif
  do 69 j=1,n+f
    ydd(lk,j)=yd(lk,j)
69 continue
  write(*,*)'le point predit',f,' ', 'est',ydd(lk,n+f)
  if(f.lt.m)goto 350
  write(*,*)'
  write(*,*)'-----'
  write(*,*)'
  lk=lk+1
  if(lk.le.kk)goto 21
  write(1,*)'          t(h)          T=120°C          T=80°C '
  write(1,*)'-----'
  do 400 j=1,m+n
    write(1,*) t1(j), ydd(1,j) ,ydd(2,j)
400 continue
  write(*,*)'
  write(*,*)'----- l execution est termine-----'
  write(*,*)' aller au fichier yd.don pour voir tes resultats '
  stop
end
```

Annexe 3 : Le programme RBF2

```

integer nc,q,d,nex,b,pl,k,kk,npr,ii
real w(150,150),ed(100),sig(150),yd(150,150),w(150,150),min,max,fyhat,yhat
real w(150,150),ed(100),sig(150),yd(150,150),y(150,150),e(150,150),yv(5,5)
real c(150),x(150),extr,saut,ksi(150,150),bl(150,150),err(3,3),um(5,5)
real az(150,150),debut,pas,extr1,saut1,min,max,fyhat,yhat,sigm,u,deb
real min,max,fyhat,yhat,sigm,u,deb,fin,tau,t2,delta,tr,ts,v
open(11,file='resultat.dat',status='unknown')
d=1
write(*,*)'donner le taux de recouvrement'
read(*,*)tau
write(*,*)'donner le nombre de points desirees'
read(*,*)nex
write(*,*)' le debut des temps en heures'
read(*,*)debut
write(*,*)'donner le pas en heures'
read(*,*)pas
write(*,*)' donner le nombre de points a predire'
read(*,*)npr
write(*,*)'donner les valeurs de la propriete 1'
do i=1,nex
read(*,*)y(1,i)
enddo
write(*,*)'donner les valeurs de la propriete 2'
do i=1,nex
read(*,*)y(2,i)
enddo
write(*,*)'fixer l erreur de la propriete 1'
read(*,*)ed(1)
write(*,*)'fixer l erreur de la propriete 2'
read(*,*)ed(2)
nc=nex
z4=0
do t2=debut,debut+pas*(npr+nc+1),pas
z4=z4+1
x(z4)=t2
enddo
ii=1
23 continue
q=0
2000 continue
v=.3
min=y(ii,q+1)
max=y(ii,q+1)
call minmax(y,min,max,ii,nex,deb,fin,q)
call centre(deb,fin,tau,nc,delta,sigm)
call orga(y,nc,deb,ii,fin,delta,extr1,saut1)
p=extr1-saut1
do i=1,nc
c(i)=p
p=p+saut1
sig(i)=sigm
enddo
k=0
! programme principal
do i=1,nc**d

```

```

call random(tr)
w(ii,i)=tr
enddo
um(ii,1)=.0
1000 continue
mk=1
do i=1,nc**d
call random(ts)
ksi(ii,i)=v*ts+b1(ii,i)
enddo
do lm=1,3
if(lm.le.1)then
kk=0
else
if(lm.le.2)then
kk=1
else
kk=-1
endif
endif
do i=1,nc**d
w(ii,i)=w(ii,i)+kk*ksi(ii,i)
enddo
do a=q+1,nex+q-d
b=0
do t=d+a-1,a
b=b+1
yv(ii,b)=y(ii,t)
enddo
pl=d+a
call ne(yv,c,sig,ii,w,nc,d,yhat)
fyhat=yhat-y(ii,pl)
az(ii,pl)=yhat
e(ii,pl)=fyhat
enddo
u=0
do i=d+q+1,nex+q
u=u+(e(ii,i)**2/(nex-d))
enddo
err(ii,lm)=sqrt(u)
enddo
if(err(ii,2).lt.err(ii,1))then
do i=1,nc**d
w(ii,i)=w(ii,i)+ksi(ii,i)
b1(ii,i)=.4*ksi(ii,i)+.2*b1(ii,i)
enddo
else
if((err(ii,2).ge.err(ii,1)).and.(err(ii,3).lt.err(ii,1)))then
do i=1,nc**d
w(ii,i)=w(ii,i)-ksi(ii,i)
b1(ii,i)=-.4*ksi(ii,i)+b1(ii,i)
enddo
else
do i=1,nc**d
b1(ii,i)=.5*b1(ii,i)
enddo
endif
endif
if(err(ii,1).gt.ed(i))then
k=k+1
mk=mk+1

```

```

um(ii,mk)=err(ii,1)
if(um(ii,mk)-um(ii,mk-1).eq..0)then
if(v.ge..0)then
v=v-.1
else
v=.01
endif
endif
um(ii,mk-1)=um(ii,mk)
goto 1000
else
q=q+1
h=0
do t=pl,nex+q-d
h=h+1
yv(ii,h)=y(ii,t)
enddo
call orga(y,nc,deb,ii,fin,delta,extr,saut)
call pon(y,nc,extr,ii,saut,sigm,w)
call ne(yv,c,sig,ii,w,nc,d,yhat)
y(ii,pl+1)=yhat
! yd(ii,p+1)=y(ii,pl+1)
endif
write(*,*)'le nombre de points predits est',q,' ',y(ii,pl+1)
if(q.lt.npr)then
goto 2000
else
ii=ii+1
if(ii.le.2)goto 23
write(*,*)' 1 execution terminee '
!do i=1,nex+q
!write(11,*)x(i),y(ii,i)
!end do
endif
end
subroutine ne(yvf,cf,sigf,ii,wf,ncf,df,yhatf)
real yvf(5,5),wf(150,150),yhatf,xfl(150,150),cf(150),distf(150,150)
real sigf(150),s1,s2,rf
integer ncf,df,ii
do i=1,ncf
distf(ii,i)=yvf(ii,1)-cf(i)
xfl(ii,i)=exp(-(distf(ii,i)**2)/(2*sigf(i)**2))
enddo
s1=.0
s2=.0
do i=1,ncf**df
s1=s1+xfl(ii,i)*wf(ii,i)
s2=s2+xfl(ii,i)
enddo
rf=s2
yhatf=s1/rf
return
end
subroutine centre(fdeb,ffin,ftau,fnc,fdelta,fsig)
real fdeb,ffin,ftau,fsig,fdelta
integer fnc
fdelta=-(fdeb-ffin)/fnc
fsig=sqrt(-fdelta**2/(8*log(ftau)))
return
end
subroutine minmax(yt,mint,maxt,ii,next,debt,fint,qt)

```

```
real yt(150,150),mint,maxt,debt,fint
integer next,qt,ii
do i=1+qt,next+qt
if(mint.ge.yt(ii,i))then
mint=yt(ii,i)
endif
if(maxt.le.yt(ii,i))then
maxt=yt(ii,i)
end if
end do
debt=mint
fint=maxt
return
end
subroutine pon(yp,ncp,extrp,ii,sautp,sigp,wp)
real pop(150,150),sautp,extrp,wp(150,150),yp(150,150)
integer ncp,ii
kt=0
do pp=extrp,extrp+2*sautp,sautp
kt=kt+1
pop(ii,kt)=exp(-(pp-extrp)**2/(2*sigp**2))
enddo
if(yp(ii,ncp-1).le.yp(ii,ncp))then
do i=1,3
wp(ii,i)=wp(ii,i)/pop(ii,i)
enddo
else
do i=1,3
wp(ii,i)=wp(ii,i)*pop(ii,i)
enddo
endif
return
end
subroutine orga(yo,nco,debo,ii,fino,deltao,extro,sauto)
real yo(150,150),debo,fino,deltao,extro,sauto
integer nco,ii
if(yo(ii,nco-1).le.yo(ii,nco))then
extro=fino
sauto=-deltao
else
extro=debo
sauto=deltao
endif
return
end
```

BIBLIOGRAPHIE

- [1] A.L.Nemmour et A.Khelifi, « Utilisation Des Réseaux De Neurones Artificiels Dans Les Câbles MT », thèse de PFE, ENP, Juin 99.
- [2] A.Boubakeur, N.Mokhnache, A.Khelifi et A.L.Nemmour, « Neural Networks Application In The Field Of High Voltage Cable Insulation Thermal Ageing », Oxford , 1999.
- [3] S.Abdi et S.Zelmat, « Influence Du Vieillissement Thermique Sur Les Propriétés De L'huile De Transformateur », thèse de PFE, ENP, Juin 98.
- [4] Hamimi Amina Nadhra, « Contribution A L'étude Des Phénomènes De Préclaquage Dans Une Huile De Transformateur Sous Champs Alternatif Divergent », thèse de magister, ENP, Octobre 97.
- [5] A.Boubakeur, « Claquage Des Diélectriques Liquides », cours de Haute Tension, 4^{ème} Année électrotechnique.
- [6] M.Kostenko et L.Piotrovski, « Machines Electriques », Tome I, Deuxième édition, Edition MIR-MOSCOU 1976.
- [7] M.Mameri, « Contribution A La Caractérisation Physico-chimique Et Electrique D'une Huile Isolante De Transformateur Neuve Et Usagée », thèse de PFE, ENP, Juillet 97.
- [8] B.Chalabi et S.Ouabdesselam, « Analyse des Gaz Dissous Dans Les Huiles Isolantes », thèse de PFE, ENP, Juillet 97.
- [9] R.Fournié, « Les Isolants En Electrotechnique, Essais, Mécanismes De Dégradation, Applications Industrielles », Ed Eyrolles, Paris 90 .
- [10] D.Ouakli et H.Brouri, « Propriétés Diélectriques De L'huile De Transformateur En Présence De Gaz Dissous », thèse de PFE, ENP, Juillet 97.
- [11] W.H.Vandeschueren, « Liquides Diélectriques », techniques de l'ingénieur, D227, 1974.
- [12] A.Labbaci, « Contribution A L'étude Et A La Régénération Des Huiles Minérales Isolantes », thèse de magister, ENP, Décembre 1991.
- [13] R.J.Milischer, « Les PCB Et Leur Produits De Décomposition : Peut-on Evaluer Un Seuil De toxicité ? », RGE 1987.
- [14] M.Lebid, « Caractérisation Physico-chimique Des Huiles Lubrifiantes Et Diélectriques » (1991).

- [15] R.Tobazeon, « Liquides Diélectriques; Claquage Et Préclaquage », techniques de l'ingénieur, D226, 1986.
- [16] H.Moulai, « Contribution A L'étude Des Phénomènes De Streamers Dans Les Diélectriques Liquides En Géométrie Pointe-plan Sous Tension De Choc De Foudre », thèse de magister, ENP, 1990.
- [17] M.Nedjar, A.Boubakeur, R.Khaili, « Etude Des Propriétés Du Polychlorure De Vinyle Avant Et Après Vieillessement Thermique », journal of technologie, novembre 1994.
- [18] SM.Sekhri, K.Nouasri, « Influence Du Vieillessement Sur L'apparition Des Décharges Partielles Dans Un câble MT Isolé Au PRC », thèse de PFE, ENP, Juin 1995.
- [19] A.Hamimi, A.Boubakeur et M.Lebid, « Influence De La Température Sur Certaines Propriétés De L'huile De transformateur Utilisée Par La SONELGAZ », deuxième conférence nationale sur la haute tension, Tizi-ouzou (1997).
- [20] Jeanny Hérault, Christian Jutten, « Réseaux Neuronaux Et Traitement Du Signal », Traité des Nouvelles Technologies, HERMES, 1994.
- [21] Kohonen.T, « Associative Memories », Printice Hall (1977).
- [22] Eric Davalo, Patrick Naïm, « Des Réseaux De Neurones », Deuxième édition, Eyrolles (1993).
- [23] M.Yazid M'hamed Yeddou, « Etude De Synthèse Sur Les Réseaux De Neurones Et Leurs Applications », thèse de magister, ENP, Juin 1998.
- [24] WERBOS.P.J, « Backpropagation Trought Time, What Is Does And How To Do It ? », Proceedings of IEEE, Octobre (1995).
- [25] KOSKO.B, « Neural Networks And Fuzzy Systems. Adynamical Machine Approch To Machines Systems », Printice Hall International (1992).
- [26] Antony N.M and Farrel J.A, «Associtive Memories Via Artificial Neural Networks », IEEE Control Systems Magazine (1990).
- [27] Simpson P.K, « Artificial Neural Networks », Pergamon Press (1990).
- [28] Kung S.Y, « Digital Neural Networks », PTR Printice Hall (1993).
- [29] Freman,J.D. and. M .Skapura. Neural Networks, Algorithms, Applications and Programming Technique. Adison Wesley Company (1992).
- [30] Rumelhart D.E, G.E.Hinton and R.J.Wiliams, « Learning Internal Representation By Error Propagation », 1986.
- [31] A.Mazouz,N.Bouaouli, « Reconnaissance De Formes Par Réseaux De Neurones, Application, Signatures », thèse de PFE , septembre 1996.

ملخص :

كانت الشبكات العصبية الاصطناعية تشكل عند ظهورها، محاولة لتحسيد خصائص العقل البشري بغرض الاستفادة منها .
إن هدف هذا العمل هو استخدام الشبكات العصبية الاصطناعية ذات تمرن مشرف في دراسة التنبؤ بتدهور خصائص الزيوت المعدنية المستعملة في المحولات الكهربائية عند خضوعها لدرجة حرارة معينة.
النتائج المتحصلة عليها من خلال هذه الدراسة ، تسمح لنا بتقدير مدة صلاحية الزيت.
الكلمات المفتاحية - الشبكات العصبية، الزيوت، التنبؤات، الشيخوخة، المحولات.

Résumé :

Les réseaux de neurones artificiels constituaient, lors de leur création, une tentative visant à mimer le cerveau biologique dans son fonctionnement afin de bénéficier de ses caractéristiques très intéressantes .

Le but de ce travail est l'utilisation des réseaux de neurones artificiels dans le vieillissement thermique de l'huile de transformateur. L'étude consiste à prédire l'évolution de certaines propriétés électriques de l'huile neuve soumise à un vieillissement thermique par un réseau de neurones à apprentissage supervisé. Les résultats que nous avons obtenus vont nous renseigner sur la durée de vie de l'huile .

Mots-clés : Réseaux de neurones, Huiles, Prédications, Vieillessement , transformateurs.

Abstract:

The artificial neuron networks constituted, at first, a tentative aiming to mimic the biologic brain in order to benefit of its very interesting features.

The goal of this work is the utilization of the artificial neuron networks in the thermal ageing of the transformer oil. The survey consists in predicting the evolution of certain electric properties of the submissive widowed oil to a thermal ageing by a supervised training neur networks . Results that we got are going to inform us on the life span of oil.

Key words: Neural networks, Oil, Prediction, Ageing, transformers.