

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DE MENTOURI-CONSTANTINE-

FACULTE DES SCIENCES DE L'INGENIEUR  
DEPARTEMENT D'ELECTRONIQUE

N°d'ordre :

Série :

**MEMOIRE**

Présenté pour l'obtention du diplôme de Magister en électronique

**THEME**

*Utilisation des réseaux de neurones  
dans l'estimation des paramètres de la  
distribution Ki-2 Non Centrale Gamma*

**Option**

**Traitement du signal**

**Par**

**OTHMANI MARABOUT Farouk**

Devant le jury :

Président :	Mr.K.BELARBI	Prof	Univ.Mentouri.Constantine.
Rapporteur :	Mr.F.SOLTANI	Prof	Univ.Mentouri.Constantine.
Examineurs :	Mr.A.BENNIA	M.C	Univ.Mentouri.Constantine.
	Mr A.CHAREF	Prof	Univ.Mentouri.Constantine.

Soutenu le .....2005

## **Remerciements**

*Avant tout je voudrai remercier alah qui ma permis d'arriver la ou je suis, et de conclure avec succès ce modeste travail, et pour les parents qu'il ma offert aussi affectifs que soucieux de mon avenir.*

*Je voudrai témoigner de ma grande gratitude envers le Professeur F.Soltani, qui ma apporter toute l'aide dont j'on avait besoin, et pour les documents qu'il ma fournis, et pour toutes les connaissances que j'ai pu acquérir sous sa direction et ses précieux conseils, et je voudrai tout spécialement le remercier pour sa patience et pour avoir été aussi compréhensive.*

*Je suis très reconnaissant à Mr Chaabi, qui a ma ouvert les portes de son laboratoire, pour sa confiance on me considérant comme un membre de son laboratoire, pour son aide matérielle grâce à laquelle j'ai accompli une grande partie de ce travail.*

*Mes vifs remerciements vont au Professeur **K.BELARBI** pour l'honneur qu'il me fait de présider le jury et également au Docteur **A.BENNIA** et au Professeur **A. Charef** d'avoir accepter de faire partie du jury.*

*Que toutes les personnes, qui ont contribué de près ou de loin, directement ou indirectement à ce travail, trouvent ici le témoignage de ma profonde reconnaissance. Enfin je ne pourrais terminer ces remerciement sans une pensée à l'ensemble des mes enseignants qui sont à l'origine de tout mon savoir, et une pensée profonde pour Madame BenDana que je voudrai remercier tout particulièrement et j'encourage a continuer à faire son travail tel qu'elle l'a toujours fait.*

# Sommaire

<b>Introduction Générale</b> .....	1
<b>Chapitre 1 : Théorie de la détection et quelques types de clutter</b>	
1-1- Introduction.....	3
1-2- Théorie de la détection.....	3
1-3- Détection CFAR (Constant False Alarm Rate) .....	5
1-4- Critères de détection .....	6
1-4-1 Critère de Bayes .....	6
1-4-2 Critère de Neyman-Pearson.....	9
1-5- Le clutter.....	10
1-5-1 Clutter de volume.....	10
1-5-2 Clutter de surface.....	10
1-5-2.1 Clutter de terre.....	10
1-5-2.2 Clutter de mer.....	11
1-6- Clutter généralisé .....	12
1-7- La distribution Gamma généralisé ( $G\Gamma$ - pdf) .....	13
1-8- Le clutter composé généralisé.....	13
1-9- Les Cas spéciaux de la pdf CG et les modèles existants du clutter .....	15
1-9-1 La distribution K (K-pdf) .....	15
1-9-2 La distribution Weibull.....	16
1-10- Clutter de distribution KI-2 non centrale Gamma.....	17
1-11 Conclusion.....	21
<b>Chapitre 2 : Rappel sur Les réseaux de neurones</b>	
2-1 Introduction.....	22
2-2 Les neurones biologiques.....	22
2-2-1. Structure des neurones .....	23
2-2-1-1. Le corps cellulaire.....	23
2-2-1-2. Les dendrites .....	24
2-2-1-3. L'axone .....	24
2-3- Neurone formel .....	25
2-4- L'utilisation des réseaux de neurones .....	26
2-4-1. Classification .....	26
2-4-2. Catégorisation (Clustering) .....	27
2-4-3. Approximation de fonctions .....	27
2-4-4. Prédiction .....	27
2-4-5. Optimisation .....	27
2-4-6. Mémoire associative .....	27
2-4-7. Contrôle .....	28
2-5- Apprentissage .....	28
2-5-1. L'apprentissage supervisé .....	28
2-5-2. L'apprentissage non supervisé .....	29
2-5-3. L'apprentissage hybride .....	29
2-6- Quelques topologies des réseaux de neurones .....	29
2-6-1. Le réseau de Hopfield.....	29
2-6-2. Le réseau de neurone de Hamming .....	30
2-6-3. Le classificateur de Carpenter et Crossberg .....	31

2-6-4. Perceptron à une seule couche.....	32
2-6-5. Le perceptron multicouches .....	32
2-6-6. Carte auto-organisatrice de Kohonen .....	37
2-7 Conclusion.....	37
<b>Chapitre 3 : La théorie de l'estimation</b>	
3-1- Introduction.....	38
3-2- Méthodes d'estimation de paramètres constants.....	38
3-2-1- Méthode du maximum de vraisemblance .....	38
3-2-2- La méthode des moments .....	39
3-3- Les critères d'un bon estimateur .....	40
3-4- Méthodes d'estimation de paramètres de quelques distributions de fouillis.....	43
3-4-1- Estimation des paramètres de la distribution K .....	43
3-4-2- Estimation des paramètres de la distribution de Weibull et Log-normal .....	46
3-5- Conclusion.....	46
<b>Chapitre 4 : Introduction de l'estimateur neuronal</b>	
4-1- Introduction.....	48
4-2- Estimation des paramètres de la distribution NG par la méthode des moments.....	48
4-2-1- L'utilisation des réseaux de neurones dans l'estimation de la distribution K. ....	50
4-2-2- L'utilisation des réseaux de neurones dans l'estimation des paramètres de la distribution Ki-2 Non Centrale Gamma.....	51
4-2-2-1- L'ensemble d'apprentissage .....	52
4-2-2-2- La génération des échantillons du Clutter.....	53
4-2-2-3- La validation .....	53
4-2-3- Résultats et discussion.....	55
4-3- Les effets des estimations sur la régulation de la fausse alarme .....	59
4-3-1 Résultats et discussion .....	61
4-4- L'influence des estimées sur les performances de la détection dans le Clutter Ki-2 Non Centrale Gamma .....	64
4-4-1- Résultats et discussion.....	68
4-5- Conclusion.....	75
<b>Conclusion Générale.....</b>	<b>76</b>
<b>Références.....</b>	<b>78</b>

# *Introduction Générale*

## **Introduction Générale**

Les systèmes RADAR ont connus un grand engouement durant ce siècle et dans de nombreux domaines, aussi bien dans le domaine militaire, que dans les domaines de communications, navigations, météorologie, et spatiale, vu ce qu'il offre comme avantages, principalement la capacité de la détection automatique d'objets mobiles dans l'espace, et sa flexibilité à pouvoir s'adapter à tout environnement que sa soit la détection d'objectifs mobiles sur terre ou mer, et quelque soit les conditions dans les quelles il opère.

En envoyant un faisceau d'ondes sous formes d'impulsions qui se propage à une vitesse d'environ 300 000 Km/s, dans une ou plusieurs directions, ce dernier sera réfléchi par les obstacles qu'il rencontrera (avion, bateau, montagne, bâtiments, vagues...etc.), le signal reçu sera constitué globalement de deux signaux ; le signal de la cible et le signal du clutter (montagnes, bâtiments, pluie, nuages, vagues,...etc.), en addition au bruit thermique qui est négligeable par rapport au signal du fouillis.

Pour atteindre des performances optimums de détection, il faut déjà choisir la bonne méthode de détection, mais avant tous, il faut identifier et caractériser précisément la nature du clutter pour pouvoir l'extraire et le séparer du signal principal pour pouvoir distinguer la cible.

Dans les radars classiques, les clutter sont modélisés par plusieurs distributions suivant l'environnement auquel est confronté le radar (mer, terre, atmosphère). Les distributions les plus employées sont la distribution de Rayleigh, Weibull, Ki-2, K...etc. Mais ce n'est pas le cas pour les nouveaux radars (radar à haute résolution) où les cellules de résolution en portée sont plus petites. Dans ce cas, les distributions déjà citées résultent en un grand taux de fausses alarmes qui dégradent considérablement les performances de détection. Des modèles récemment proposés, appelés les modèles généralisés sont caractérisés par une plus longue queue et qui s'adaptent à la nature du clutter.

La bonne estimation des paramètres qui caractérisent le clutter dans le quel opère le radar est une opération cruciale pour la détection à taux de fausse alarme constant (CFAR), plusieurs méthodes peuvent être utilisées dans l'estimation des paramètres de ces distributions (qu'elles soient mono-paramétrique ou bien

multiparamétrique). Cependant, des obstacles dans l'utilisation de ces méthodes peuvent être rencontrés que se soit au niveau du modèle de la distribution lui-même qui est complexe, ou bien au niveau de la méthode elle-même qui peut nécessiter un grand nombre d'opérations la rendant non efficace.

Afin d'éviter ces obstacles, on a jugé nécessaire de chercher d'autres méthodes plus efficaces, plus flexibles et qui ne nécessitent pas un grand nombre d'opérations. Alors voyons ce que les réseaux de neurones artificiels offrent comme caractéristiques les rendant aptes à être utilisés dans ce domaine, notre but est d'introduire les réseaux de neurones dans l'estimation des paramètres d'une distribution modélisant un clutter généralisé.

Dans ce mémoire, on se propose d'étudier et de mettre en œuvre un estimateur à base de réseaux de neurones et de observer les influences qu'il pourrait avoir sur toute la chaîne de détection et quels avantages pourrait-il apporter depuis l'estimation des paramètres du clutter généralisé jusqu'à la détection adaptative CFAR, et on comparant nos résultats avec ceux obtenus à partir d'une méthode classique et qui est la méthode des moments.

### **Organisation du mémoire :**

Dans le chapitre 1 on donne un rappel sur la théorie de la détection en rappelant quelques critères de détection, pour ensuite décrire les différents types de modèles de clutter qui existent dans la littérature.

Le chapitre 2 on donne une description des réseaux de neurones artificiels avec une vue générale sur les différents domaines dans lesquels les réseaux de neurones sont applicables ainsi que quelques modèles de réseaux de neurones. On focalisera notre attention sur l'architecture MLP, et l'algorithme de rétro-propagation.

Le chapitre 3 est consacré à la théorie de l'estimation en introduisant quelques méthodes d'estimation des paramètres du clutter et en présentant quelques exemples existants dans la littérature. Un intérêt particulier sera porté sur la méthode des moments.

Dans le chapitre 4, nous appliquons les deux méthodes ; la méthode des moments et les réseaux de neurones pour l'estimation des paramètres du clutter Ki-2 Non Centrale Gamma. Ensuite, on fait une comparaison entre leurs performances à travers une étude sur leurs influences sur la régulation de fausse alarme, ainsi que sur les performances de la détection. Pour terminer, on synthétise nos résultats dans une conclusion générale.

**Chapitre1 : Théorie de la détection  
et  
quelques types de clutter**



## **1-1 Introduction :**

Le principe physique soulignant le fonctionnement d'un système radar a été démontré par Hertz dans les années 1880. Son expérience a démontrée que les ondes électromagnétiques son réfléchies par des objets métalliques tout comme le lumière est réfléchie par des miroirs, et c'est seulement en 1904 qu'appareil pratique a été mis à l'épreuve et ensuite commercialisé par un ingénieur Allemand (Christian Hulsmeyer), pour la prévention contre la collision des bateaux. Le développement d'un radar opérationnel n'a sérieusement commencé qu'en 1930. Durant cette période, l'implication militaire du radar pour la localisation des avions sous des conditions de visibilité très limitées été devenue très importante. Le travail a été entrepris en Allemagne, France, Grande Bretagne et les USA et le plus distingué de ces efforts été celui des Anglais .Durant la guerre, des efforts intensifs dans plusieurs pays ont résulté en un développement de nouvelles technologies et de nouvelles techniques de traitement de signal [1]. Dans les années 1950, la nature de la menace militaire a changé net et des systèmes radar plus sophistiqués ont été développés pour s'adapter aux nouvelles conditions .Comme exemple, le système SAGE, qui représente le première usage de l'ordinateur avec de nouvelles techniques pour remplacer l'opérateur humain. Dans les dernières années, le radar a été largement utilisé dans d'innombrables domaines aussi bien militaires que civils, comme le contrôle du trafic aérien ,navigation maritime, météorologie ,et le trafic routier .

## **1-2- Théorie de la détection :**

La détection est définie comme étant l'opération qui consiste à prendre une décision sur la présence ou l'absence de cible à partir des observations disponibles à la sortie du récepteur, et cela en comparant le signal reçu à un seuil  $T$  calculé à priori. Ce dernier dépend des paramètres du bruit dans lequel il est noyé et d'autres paramètres qu'on verra plus loin. Ceci peut se formuler en termes des tests d'hypothèses statistiques qu'on appellera hypothèse nulle  $H_0$  et hypothèse alternative  $H_1$  où, sous l'hypothèse  $H_0$  la cible est déclarée absente (c-a-d le signal reçu ne dépasse pas le seuil  $T$  ) ,et le contraire pour  $H_1$  . A cause de la nature aléatoire du bruit et les fluctuations de la cibles, la détection est étudiée comme un processus aléatoire [2].

On suppose que la décision est prise sur une observation du signal reçu  $Y$ . L'ensemble des valeurs que prend la variable aléatoire constitue l'espace d'observation  $Z$ , comme il est montré sur la Figure 2.1, qui est en fait l'union de deux espaces  $Z_0$  et  $Z_1$  :

$$Z = Z_0 \cup Z_1 \quad (1-1)$$

où  $Z_0$  : correspond à l'hypothèse  $H_0$

$Z_1$  : correspond à l'hypothèse  $H_1$

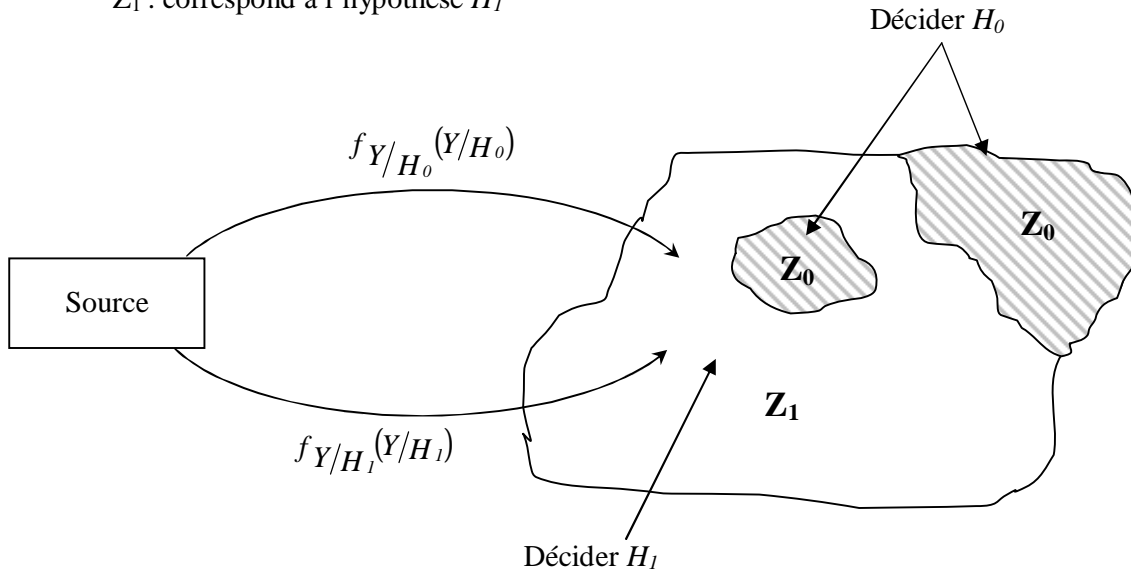


Figure 2.1 : Les régions de décision

$f_{y/H_0}(Y/H_0)$  et  $f_{y/H_1}(Y/H_1)$  sont les densités de probabilité conditionnelles connues sous le nom de fonctions de vraisemblance .

Chaque fois qu'une décision est prise, les performances du radar sont déterminées par les probabilités de quatre cas :

1.  $H_0$  est vraie et  $H_0$  est décidée.
2.  $H_0$  est vraie et  $H_1$  est décidée.
3.  $H_1$  est vraie et  $H_1$  est décidée.
4.  $H_1$  est vraie et  $H_0$  est décidée.

On remarque que le récepteur prend une bonne décision dans les cas (1) et (3) et qu'il commet une erreur dans les deux autres cas. L'événement (3) est appelé détection, la probabilité associée à ce dernier est probabilité de détection  $P_d$ . L'événement (2) est une fausse alarme car une présence de cible est décidée lorsqu'elle est absente, donc sa probabilité est  $P_{fa}$ . Quant à l'événement (4), il se réalise lorsque (3) n'est pas réalisé, par conséquent sa probabilité est  $1 - P_d$  et on l'appelle la probabilité de non détection  $P_m$  (probability of miss). Pour l'événement (1), sa probabilité est  $1 - P_{fa}$ .

Comme il a été déjà mentionné, le seuil  $T$  qui détermine la décision dépend étroitement de l'estimation des paramètres du bruit dans le quel il est noyé. Ceci est l'objet de

notre travail et aussi de la Pfa pré-spécifiée, qu'on doit garder constante, ce qui nous mène à parler du CFAR (Constant False Alarm Rate) avant de passer aux bases théoriques de la détection [2].

### 1-3- Détection CFAR (Constant False Alarm Rate) :

La détection d'une cible radar est affectée par la présence d'un retour d'écho qui est dû aux réflexions de bâtiments, forêts, reliefs, la mer ...etc. Du moment où la puissance du clutter est inconnue, on ne peut utiliser des techniques à seuil fixe car elles peuvent résulter en un nombre excessive de fausses alarmes. Le détecteur CFAR le plus simple est le CA-CFAR (Cell Averaging CFAR) montré en Figure 2.2. Il fait une comparaison entre la cellule test et un seuil adaptatif (qui varie avec les variations de l'environnement) calculé à partir des cellules avoisinantes appelées cellules de références. Le niveau total  $Q$  du bruit dans la cellule test est estimé en faisant la somme arithmétique du contenu de ces dernières comme indiqué sur la Figure 2.2, pour enfin obtenir la valeur du seuil en multipliant par un facteur  $T'$  [3].

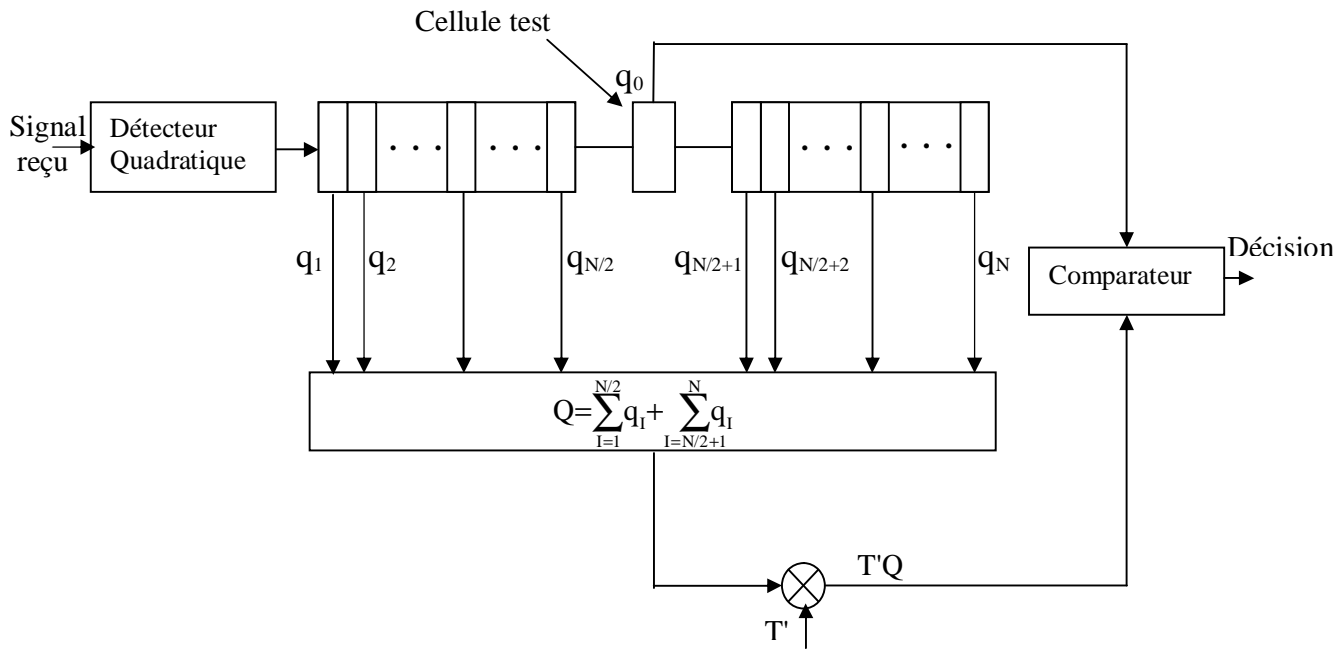


Figure 2-2 : Le détecteur CA-CFAR

La performance de ce détecteur est bonne si les cellules de références contiennent du bruit gaussien indépendant et identiquement distribué. Cependant, lorsque le bruit dans les cellules de références est corrompu par la présence d'une cible interférente ou par une

transition de la puissance du clutter, les performances se dégradent considérablement. Pour faire face à ce genre de cas, plusieurs algorithmes ont été proposés tel que le GO-CFAR (Greatest Of CFAR), le SO-CFAR (Smallest Of CFAR), OS-CFAR (Ordered Statistic CFAR), et dernièrement les ANN-CFAR [3] ...etc. Chacun de ces algorithmes est performant dans un cas de clutter ou d'environnement bien spécifique.

#### 1-4- Critères de détection :

L'objectif principal de la détection est d'arriver à établir une règle qui optimise les performances du récepteur. Pour arriver à cette fin, on se base sur plusieurs critères et qui sont [2]:

##### 1-4-1 Critère de Bayes :

Pour pouvoir introduire le critère de Bayes dans la détection, il est nécessaire de satisfaire deux conditions importantes :

- ✚ On doit connaître au préalable les probabilités des deux hypothèses

$$P(H_0) = P_0, \text{ et } P(H_1) = P_1 \quad (1-2)$$

Qu'on appelle les probabilités à priori.

- ✚ On doit associer un coût à chaque décision possible, car dans une détection radar les conséquences varient d'une décision à une autre. On notera ces coûts par  $C_{ij}$  où :

- « i » indique l'hypothèse choisie.
- « j » indique l'hypothèse vraie.

Donc on aura :

- $C_{00}$  à 1<sup>er</sup> cas
- $C_{10}$  à 2<sup>ème</sup> cas
- $C_{11}$  à 3<sup>ème</sup> cas
- $C_{01}$  à 4<sup>ème</sup> cas

On suppose aussi que les coûts d'une décision erronée est supérieur à celui d'une décision correcte c.à.d :

$$\begin{cases} C_{10} > C_{00} \\ C_{01} > C_{11} \end{cases}$$

(1.3)

Le but du critère de Bayes est de déterminer la règle de décision qui minimise le coût moyen appelé aussi le risque de Bayes  $R=E(C)$ . Pour cela on procède de la manière suivante ; sachant que la probabilité d'une décision correcte et la probabilité d'erreur, respectivement, sont comme suit :

$$\begin{aligned}P(\text{correct}) &= P(\text{correct}, H_0) + P(\text{correct}, H_1) \\ &= P(\text{correct}/H_0) * P(H_0) + P(\text{correct}/H_1) * P(H_1) \\ &= P(H_0/H_0) * P(H_0) + P(H_1/H_1) * P(H_1) \\ P(\text{correct}) &= (1-P_{fa}) * P(H_0) + P_d * P(H_1)\end{aligned}\tag{1-4}$$

$$\begin{aligned}P(\text{error}) &= P(\text{error}, H_0) + P(\text{error}, H_1) \\ &= P(\text{error}/H_0) * P(H_0) + P(\text{error}/H_1) * P(H_1) \\ &= P(H_1/H_0) * P(H_0) + P(H_0/H_1) * P(H_1) \\ P(\text{error}) &= P_{fa} * P(H_0) + (1-P_d) * P(H_1)\end{aligned}\tag{1-5}$$

Où:

- $P(H_0/H_0) = P(\text{decider } H_0, H_0 \text{ vraie})$
- $P(H_1/H_1) = P(\text{decider } H_1, H_1 \text{ vraie})$
- $P(H_1/H_0) = P(\text{decider } H_1, H_0 \text{ vraie})$
- $P(H_0/H_1) = P(\text{decider } H_0, H_1 \text{ vraie})$

Les probabilités conditionnelles  $P(H_i/H_j)$  ;  $i, j=0,1$ , en fonction des régions d'observation sont :

$$P(H_i/H_j) = \int_{Z_i} f_{Y/H_j}(Y/H_j) dy \quad ; i, j=0,1\tag{1-6}$$

Sachant que les probabilités  $P(H_0/H_1)$ ,  $P(H_1/H_0)$ ,  $P(H_1/H_1)$  représentent respectivement, la probabilité de non détection  $P_m$ , la probabilité de fausse alarme  $P_{fa}$  et la probabilité de détection  $P_d$ . On peut maintenant donner la formule du risque qui est :

$$\begin{aligned}
R &= C_{00} P(\text{correct}, H_0) + C_{01} P(\text{error}, H_1) + C_{10} P(\text{error}, H_0) + C_{11} P(\text{correct}, H_1) \\
R &= C_{00} (1 - P_{fa}) P(H_0) + C_{01} (1 - P_d) P(H_1) + C_{10} P_{fa} P(H_0) + C_{11} P_d P(H_1) \quad (1-7)
\end{aligned}$$

En remplaçant (1-6) dans (1-7) on obtient:

$$\begin{aligned}
R &= C_{00} P(H_0) \int_{Z_0} f_{Y/H_0}(Y/H_0) dy + C_{01} P(H_1) \int_{Z_1} f_{Y/H_1}(Y/H_1) dy \quad + \\
&C_{10} P(H_0) \int_{Z_1} f_{Y/H_0}(Y/H_0) dy + C_{11} P(H_1) \int_{Z_1} f_{Y/H_1}(Y/H_1) dy \quad (1-8)
\end{aligned}$$

En prenant on compte les égalités suivantes :

$$\int_Z f_{Y/H_0}(Y/H_0) dy = \int_Z f_{Y/H_1}(Y/H_1) dy = 1 \quad (1-9)$$

$$\int_{Z_0} f_{Y/H_i}(Y/H_i) dy + \int_{Z_1} f_{Y/H_i}(Y/H_i) dy = 1, \quad i=0,1 \quad (1-10)$$

Donc on peut écrire :

$$\int_{Z_1} f_{Y/H_i}(Y/H_i) dy = 1 - \int_{Z_0} f_{Y/H_i}(Y/H_i) dy, \quad i=0,1 \quad (1-11)$$

On substituant l'équation (1-11) dans l'équation (1-8)

$$R = \underbrace{C_{10}P_0 + C_{11}P_1}_{A} + \underbrace{\int_{Z_0} \left\{ P_1(C_{01} - C_{11}) f_{Y/H_1}(y/H_1) - \left[ P_0(C_{10} - C_{00}) f_{Y/H_0}(y/H_0) \right] \right\} dy}_{B} \quad (1-12)$$

On peut remarquer que les probabilités à priori ainsi que les coûts sont des constantes, alors A est une constante et on peut aussi noter que les termes B et C sont des quantités positives. Le risque est minimisé en choisissant la région de décision  $Z_0$  de façon que la valeur à intégrer (c à d le terme B) soit négative en tout point de  $Z_0$ , donc le critère de Bayes résultera sur la règle suivante :

$$\Omega(Y) = \frac{f_{y/H_1}(y/H_1)}{f_{y/H_0}(y/H_0)} \begin{matrix} > \\ < \end{matrix} \frac{H_1}{H_0} \frac{(C_{10}-C_{00})P_0}{(C_{01}-C_{11})P_1} \quad (1-13)$$

$\Omega(Y)$  est le rapport de vraisemblance, et  $\eta = \frac{(C_{10}-C_{00})P_0}{(C_{01}-C_{11})P_1}$  est le seuil de décision.

Dans la plus part des cas pratiques, les probabilités à priori,  $P_0$  et  $P_1$  sont inconnues, il est donc impossible d'utiliser le critère de Bayes dans ce cas. Il y a une autre approche pour obtenir le critère de décision qui consiste à sélectionner à priori une valeur de probabilité  $P_1$  pour la quelle le risque est maximum et donc minimisant cette fonction du risque .Ce principe de minimisation du coûts moyen pour la probabilité sélectionnée  $P_1$  est connue sous le nom du critère du Minimax [2].

#### 1-4-2 Critère de Neyman-Pearson :

En l'absence d'information sur les probabilités à priori et les coûts de décision  $C_{ij}$ ,  $i = 0,1$  et  $j = 0,1$ , on utilise le critère de Neyman-Pearson qui consiste à minimiser la probabilité de non détection  $P_m$ , sachant que la probabilité de fausse alarme  $P_{fa}$  est fixée à une certaine valeur [2]. Pour cela, nous construirons la fonction objective  $J(\lambda)$  donnée par :

$$J(\lambda) = P_m + \lambda (P_{fa} - \alpha) \quad (1-14)$$

Où  $\lambda$  est le multiplicateur de Lagrange. D'après les équations (1-6) et (1-14) nous pouvons écrire :

$$J(\lambda) = \lambda (1-\alpha) + \int_{z_0} [f_{y/H_1}(y/H_1) - I f_{y/H_0}(y/H_0)] dy \quad (1-15)$$

# ***Chapitre2 : Rappel sur Les réseaux de neurones***



## 2.1 Introduction :

Historiquement les premiers modèles neuromimétiques ont été développés dans la mouvance cybernéticienne. Ils devaient permettre de valider et d'utiliser en robotique les premiers modèles neurobiologiques de la conscience. Les cybernéticiens tendaient en effet vers une conception matérialiste réductrice de la conscience (conception béhavioriste selon laquelle la conscience est issue d'une combinaison de boucles réflexes). L'adaptation informatique des principes organisationnels du cortex devait alors leur permettre de vérifier cette hypothèse.

Les réseaux de neurones sont des modèles mathématiques et informatiques, des assemblages d'unités de calculs appelés neurones formels, et dont l'inspiration originelle était un modèle de la cellule nerveuse humaine. Cet héritage de la neurobiologie forme une composante importante de la matière, et le souci de maintenir une certaine correspondance avec le système nerveux humain a animé et continue à animer une part importante des recherches dans le domaine. Malgré cet héritage, l'essentiel des travaux d'aujourd'hui ont pour objet le réseau de neurones formels et non son corrélat neurobiologique. Vu comme des systèmes de calcul, les réseaux de neurones possèdent plusieurs propriétés qui les rendent intéressants d'un point de vue théorique, et fort utiles en pratiques [9].

Les premiers travaux sur les neurones artificiels ont débuté au début des années 1940 et ont été menés par McCulloch et Pitts. Ils décrivent les propriétés du système nerveux à partir de neurones idéalisés : ce sont des neurones logiques (0 ou 1). Dix années plus tard, on a constitué le premier modèle réel d'un réseau de neurones. En 1960, le premier perceptron est créé par Rosenblatt. Puis, durant les années 1970 il y eut une remise en cause de l'intérêt des réseaux car les ordinateurs de neurones apprenaient lentement, coûtaient très cher et leurs performances n'étaient pas si impressionnantes. La disponibilité croissante des minis et micro ordinateurs, vers la fin des années 1970, a permis aux réseaux de neurones de prendre un nouveau départ [3].

## 2-2. Les neurones biologiques :

Les cellules nerveuses, appelées neurones, sont les éléments de base du système nerveux central. Celui-ci en posséderait environ cent 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 :

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

### 2-2-1. Structure des neurones :

Un neurone est constitué de trois parties :

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

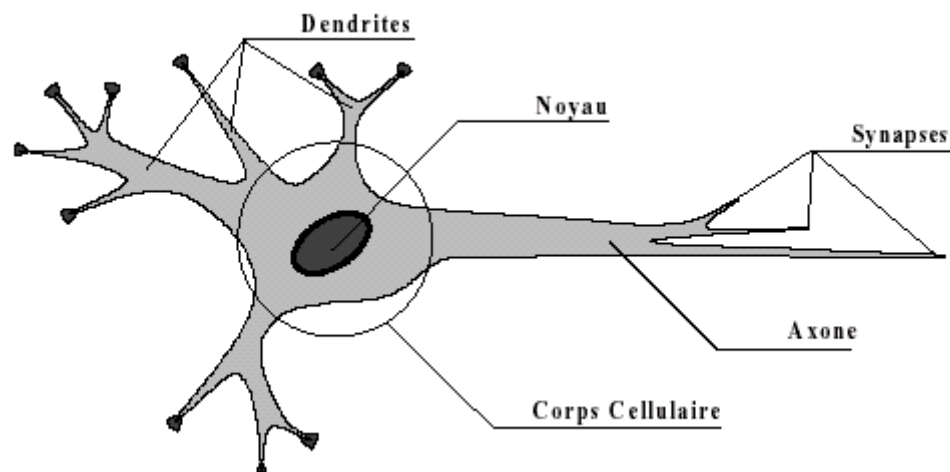


Figure 2-1. Neurone biologique

#### 2-2-1-1. Le corps cellulaire :

Il contient le noyau du neurone et effectue les transformations biochimiques nécessaires à la synthèse des enzymes et des autres molécules qui assurent la vie du neurone, il a une forme pyramidale ou sphérique dans la plus part des cas.

### 2-2-1-2. Les dendrites :

Chaque neurone possède une « chevelure » de dendrites. Celles-ci sont de fines extensions tubulaires, de quelques dixième de microns. De diamètre et d'une longueur de quelques dizaine de microns. Elles sont les récepteurs principaux du neurone pour capter les signaux qui lui parviennent.

### 2-2-1-3. L'axone :

L'axone, qui est à proprement parler la fibre nerveuse, sert de moyen de transport pour les signaux émis par le neurone. Pour former le système nerveux, les neurones sont connectés les uns aux autres suivant des répartitions spatiales complexes, les connexions entre deux neurones se font en des endroits appelés synapses où ils sont séparés par un petit espace synaptique de l'ordre d'un centième de micron. D'une façon simple, on peut dire que le soma du neurone traite les courants électriques qui lui proviennent de ses dendrites, et qu'il transmet le courant électrique (sous forme d'impulsions chaque une de durée d'environ 1ms et une amplitude d'environ 100mv) résultant de ce traitement aux neurones aux quels il est connecté par l'intermédiaire de son axone.

Le schéma classique présenté par les biologistes est celui d'un soma effectuant une sommation des influx nerveux transmis par ses dendrites. Si la sommation dépasse un certain seuil, le neurone répond par un influx nerveux au potentiel d'action qui se propage le long de son axone. Si la sommation est inférieure à ce seuil, le neurone reste inactif [10].

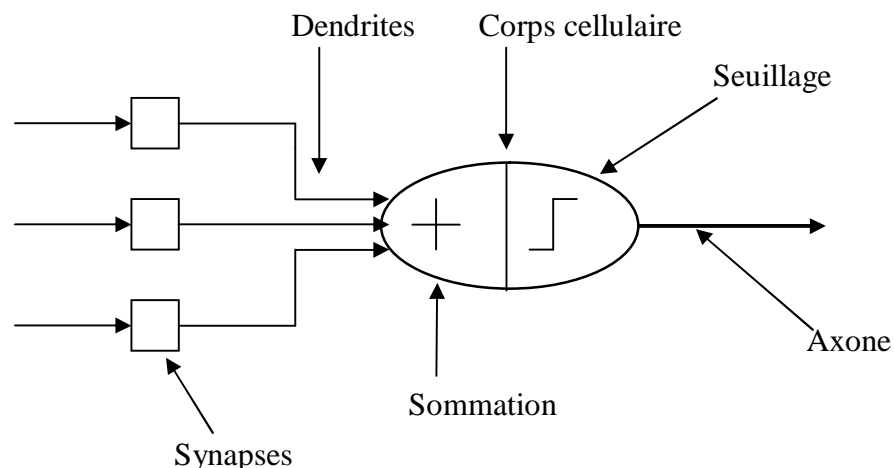


Figure 2-2. Le modèle d'un neurone biologique

### 2-3- Neurone formel :

McCulloch et Pitts ont proposé une unité à seuil binaire comme un modèle de neuronal diagramme de McCulloch-Pitts est montré dans la Figure 2-3 :

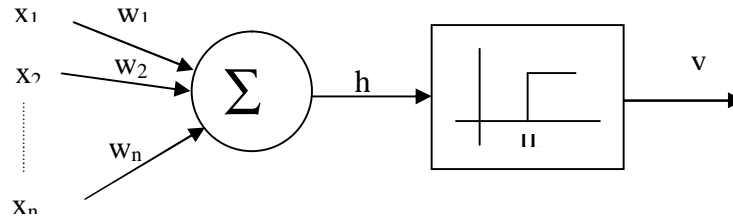


Figure 2-3. Le modèle de McCulloch-Pitts

Ce modèle mathématique calcule la somme pondérée de ses  $n$  entrées,  $x_j$ ,  $j=1,2,\dots,n$ . Le nœud est caractérisé par un seuil  $\mu$ , il génère une sortie de « 1 » si cette somme est supérieure au seuil  $\mu$ , et autrement une sortie de « 0 ». Mathématiquement :

$$v = \begin{cases} 1 & \text{si } \sum_{i=1}^n w_i x_i \geq \mu \\ 0 & \text{si } \sum_{i=1}^n w_i x_i < \mu \end{cases}$$

où  $\Phi(\cdot)$  est la fonction d'activation, et  $w_j$  est le poids synaptique associé à la  $j^{\text{ème}}$  entrée, parfois, on considère le seuil  $\mu$  comme un autre poids  $w_0 = -\mu$  et qui est connecté à un neurone avec une entrée constante,  $x_0 = 1$ . Les poids positifs correspondent à des synapses excitatrices, tandis que les poids négatifs correspondent aux synapses inhibitrices. Il y a une analogie évidente (Tableau 2-1) avec les neurones biologiques :

Neurones biologiques	Neurones artificiels
Synapses	Connections pondérées
Axons	Sorties
Dendrites	Entrées
Sommateur	Fonction d'activation

Tableau 2-1. L'analogie entre les neurones biologiques et les neurones artificiels

Le modèle de McCulloch et Pitts a été généralisé par de nombreuses façon, une généralisation évidente et celle d'utiliser une autre fonction d'activation que la fonction à seuil, e.g, la fonction logistique, tangente hyperbolique, ou bien Gaussienne, qui sont montrées dans la Figure 2-4. La fonction sigmoïde est de loin la plus utilisée dans les RNA, C'est une fonction qui est caractérisée par sa propriété asymptotique [12]. La fonction sigmoïde standard est la fonction logistique et elle est définie par :

$$G(x)=1/(1+\exp(-bx))$$

Où  $b$  est le paramètre de pente.

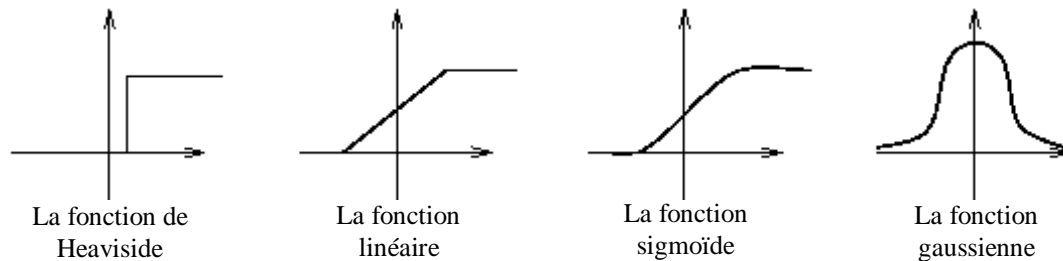


Figure 2-4. Quelques types de fonctions d'activations.

#### 2-4- L'utilisation des réseaux de neurones :

Les réseaux de neurones fournissent une technique pour obtenir la capacité traitement exigée en utilisant un grand nombre d'éléments de calcul simples fonctionnant en parallèle et c'est ce qui fait la puissance des réseaux de neurones [11], plusieurs défis et problèmes au quels ont fait face ingénieurs et scientifiques et dans les quelles les RNA ont apportées une avance considérable et des solutions fructueuses, et de ce succès les RNA sont devenus maintenant un vrais outil de résolution de problèmes ,dont on vas citer quelques uns :

##### 2-4-1- Classification :

La tâche de classification est d'assigner un modèle d'entrée(e.g un signale vocale ou des caractères manuscrits) représenté par un vecteur de caractérisant une des classes pré-spécifiée. Des applications bien connues de la classification est la reconnaissance de caractères, reconnaissance de la parole, classification des signaux EEG classification des cellules sanguines, l'inspection des circuit imprimés.

### 2-4-2- Catégorisation (Clustering) :

Dans la catégorisation aussi connue comme la classification non supervisée, il n'y a pas de modèles d'apprentissage avec bien connue. L'algorithme de catégorisation explore et teste la similarité entre les différents modèles qui lui sont présentés, et place les modèles similaires dans une même catégorie (cluster).

### 2-4-3- Approximation de fonctions :

Plusieurs problèmes scientifiques et d'ingénieries requièrent l'approximation. Ayant un ensemble de  $N$  paires d'apprentissage (entrée-sortie), générées par une fonction inconnue  $M(x)$ , la tâche d'approximation de fonction est de trouver une estimation, disons  $f$ , de la fonction inconnue  $M$ .

### 2-4-4- Prédiction :

Ayant un ensemble de  $N$  échantillons dans une séquence temporelle déterminée,  $\{y(t_1), y(t_2), \dots, y(t_N) / t_1 < t_2 < \dots < t_N\}$ , le but est de déterminer la valeur de  $y(t_{N+1})$  à l'instant future  $t_{N+1}$ . La prédiction est très utilisée dans la bourse, le contrôle des procédés, prévisions météorologiques.

### 2-4-5- Optimisation :

Un problème d'optimisation peut généralement comprendre les éléments suivants :

- ✚ Un ensemble de variables indépendantes qui font référence à l'état du processus
- ✚ Une fonction objective (fonction de coût/erreur) qui doit être optimisée.
- ✚ Un ensemble de contraintes si elles existent bien sûr.

Le but de l'optimisation est de trouver un état qui satisfait ces contraintes de telle façon que la fonction objective soit optimisée.

### 2-4-6- Mémoire associative :

La mémoire associative, peut être accédée par l'intermédiaire de sa contenance, et les données de cette mémoire peuvent être reconnues par cette dernière même si ils sont bruités ou seulement une partie d'un modèle d'entrée est disponible. E.g la récupération

bibliographique des références d'une revue de l'information partielle, de ce fait la mémoire associative est très appréciée dans la construction des bases de données multimédia.

#### 2-4-7- Contrôle :

Considérons un système définie par la paire  $\{u(t),y(t)\}$  où  $u(t)$  est le contrôle du système et  $y(t)$  est la sortie résultante à l'instant  $t$ , (dans le contrôle adaptatif) le but est de générer une entrée de contrôle  $u(t)$  de façon que le système suit la sortie désirée déterminée par le modèle de référence.[12].

#### 2-5- Apprentissage :

La capacité à apprendre un trait fondamental de l'intelligence, d'un autre côté ce que veut dire l'apprentissage est parfois difficile à décrire, le processus d'apprentissage, dans le contextes des réseaux de neurones peut être vu comme un problème ajustement et d'adaptation de l'architecture et des poids du réseaux de neurone de façon que le RNA exécute efficacement une tâche spécifique. Au lieu de spécifier un ensemble de règles, les réseaux de neurones apprennent automatiquement à partir de la collection d'exemples, et c'est ce qui les rend très attractives.

Il y a trois types de paradigmes d'apprentissage :

- ✚ Supervisé
- ✚ Non supervisé
- ✚ Hybride

#### 2-5-1- L'apprentissage supervisé :

C'est actuellement le mode d'apprentissage le plus couramment utilisé. Son principe est élémentaire : on soumet au réseau à un grand nombre d'exemples pour lesquels l'entrée et la sortie associée sont connues et les poids sont modifiés de façon à corriger l'erreur commise par le réseau. Le plus répandu des algorithmes d'apprentissage supervisé est l'algorithme de rétropropagation du gradient d'erreur qui, appliqué aux réseaux Multi Couches.

### 2-5-2- L'apprentissage non supervisé :

Contrairement aux modes supervisés, seule une base d'entrées est fournie au réseau. Celui-ci doit déterminer lui-même ses sorties en fonction des similarités détectées entre les différentes entrées, c'est-à-dire en fonction d'une règle d'auto-organisation. Il explore la structure des données, ou encore la corrélation entre les modèles d'entrée et les organise à partir de cette corrélation en des catégories.

### 2-5-3- L'apprentissage hybride :

Il combine l'apprentissage supervisé et non supervisé. Dans de nombreux problèmes on ne dispose pas de l'information nécessaire à la construction d'une base d'apprentissage complète. On ne dispose souvent que d'une information qualitative permettant l'évaluation de la réponse calculée, sans pour autant connaître la réponse la plus adaptée.

### 2-6- Quelques topologies des réseaux de neurones :

Nous allons maintenant voir brièvement quelques types de réseaux de neurones, sur la Figure 2-5 on a une vue générale sur quelques topologies d' RNA

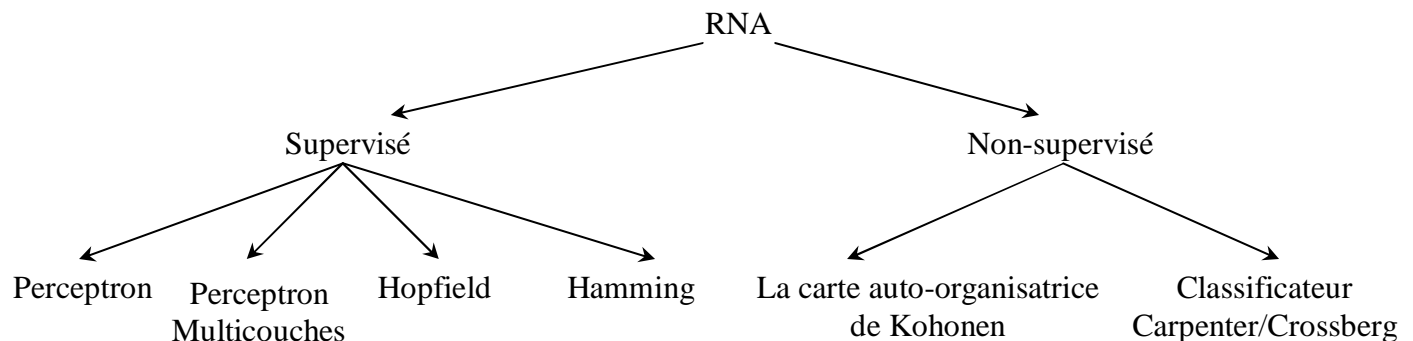


Figure 2-5. Quelques types de réseaux de neurones

#### 2-6-1- Le réseau de Hopfield :

Le réseau de Hopfield ainsi que les deux autres réseaux montrés dans la Figure 2-5 sont généralement employés avec des entrées binaires. Ces réseaux sont les plus appropriés lorsque des représentations binaires exactes sont possibles (e.g des images en noir et blanc où les éléments d'entrée sont des valeurs des pixels, ou avec un texte en code ASCII où



chaque caractère peut être représenté par des bits). Sur la Figure 2-6 on voit un réseau de Hopfield contenant  $N$  nœuds avec des fonctions seuil, des entrées binaires et des sorties pouvant prendre les valeurs  $+1$  et  $-1$ . La sortie de chaque neurone est connectée à tous les autres nœuds par l'intermédiaire de facteurs de pondération (poids).

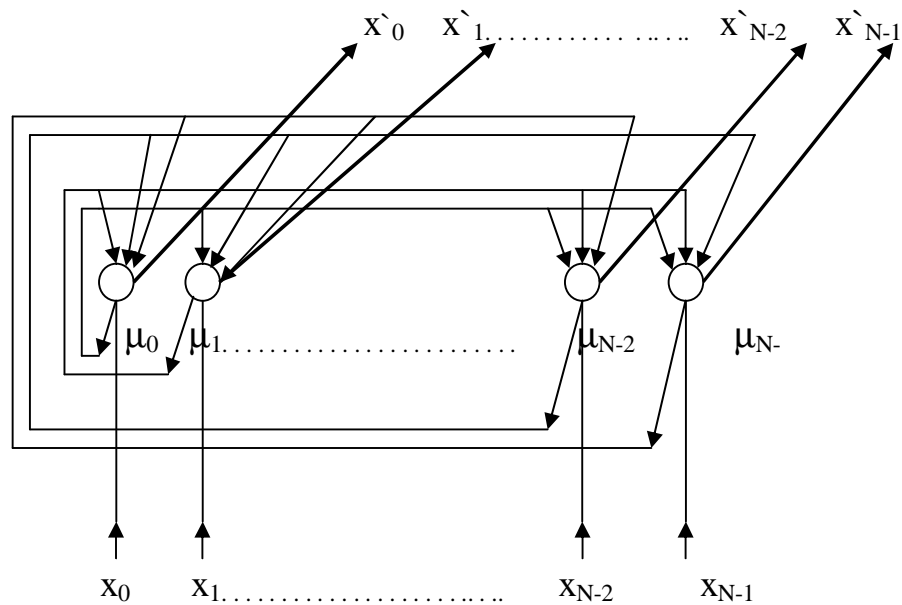


Figure 2-6. Réseau de Hopfield

### 2-6-2- Le réseau de neurone de Hamming :

Le réseau de Hopfield est souvent testé sur des problèmes où les entrées sont produites en choisissant un exemplaire en inversant les bits d'une manière aléatoire et indépendante avec une probabilité donnée. C'est un problème classique dans la théorie de communications qui se produit quand des signaux binaires à longueur constante sont envoyés par un canal binaire symétrique sans mémoire. Le classificateur optimal d'erreur minimale dans ce cas calcule la distance de Hamming de l'exemplaire pour chaque classe et choisit cette classe en fonction de la distance minimum de Hamming.

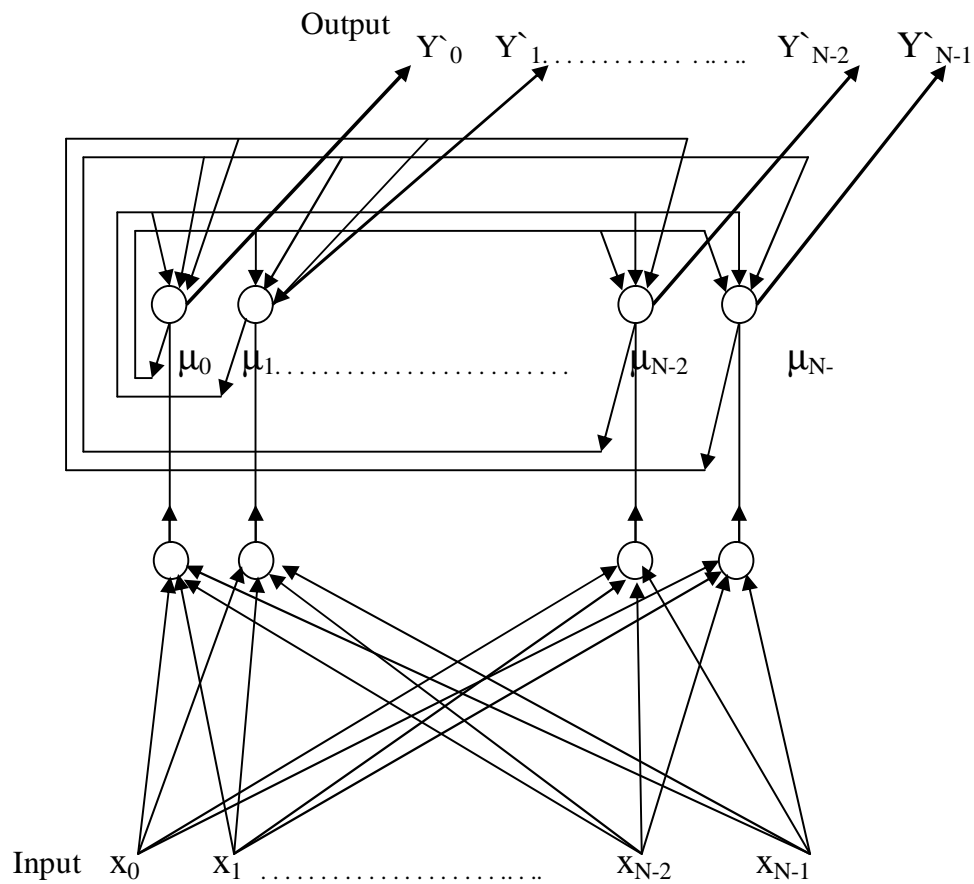


Figure 2-7. Réseau de Hamming

### 2-6-3- Le classificateur de Carpenter et Crossberg :

Carpenter et Grossberg, dans le développement de leur théorie de résonance adaptative ont conçu un réseau de neurones de classification avec un apprentissage non supervisé. Ce réseau met en application un algorithme de catégorisation qui est très semblable à l'algorithme séquentiel simple de chef. L'algorithme de chef choisit la première entrée comme exemplaire pour le premier faisceau. La prochaine entrée est comparée au premier exemplaire de faisceau. Elle suit "le chef" et est groupée avec le premier si la distance à la première est moins qu'un seuil. Autrement, c'est l'exemplaire pour un nouveau faisceau. Ce processus est répété pour toutes les entrées suivantes.

#### 2-6-4- Perceptron à une seule couche :

La perception à une seule couche est le premier de trois réseaux d'après la taxonomie de la Figure 2-3 qui peuvent être employés avec des entrées binaires ou continues. Ce réseau simple a suscité beaucoup d'intérêt, quand initialement, a été développé en raison de sa capacité d'apprendre à identifier des modèles simples.

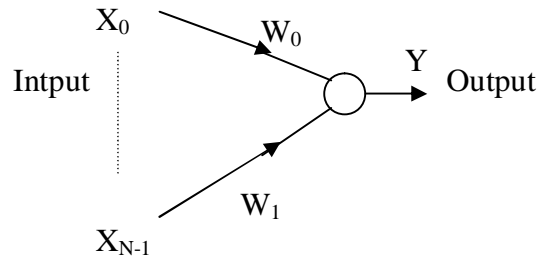


Figure 2-8. Perceptron à une seule couche

Les poids de connexion ainsi que le seuil dans un perceptron à une seule couche peuvent être fixés ou adaptés en utilisant un certain nombre d'algorithmes. La procédure originale de convergence du perceptron à une seule couche pour l'ajustement des poids a été développée par Rosenblatt.

#### 2-6-5- Le perceptron multicouches :

Les perceptrons multicouches sont des réseaux non-récurrents avec une ou plusieurs couches de neurones entre les couches d'entrée et de sortie. Ces couches additionnelles contiennent des unités cachées ou des neurones qui ne sont pas directement reliés aux neurones des couches d'entrée et de sortie. Un perceptron avec trois couches dont deux couches cachées est montré dans la Figure 2-9. Ils sont de plus en plus employés avec le développement de nouveaux algorithmes d'apprentissage.

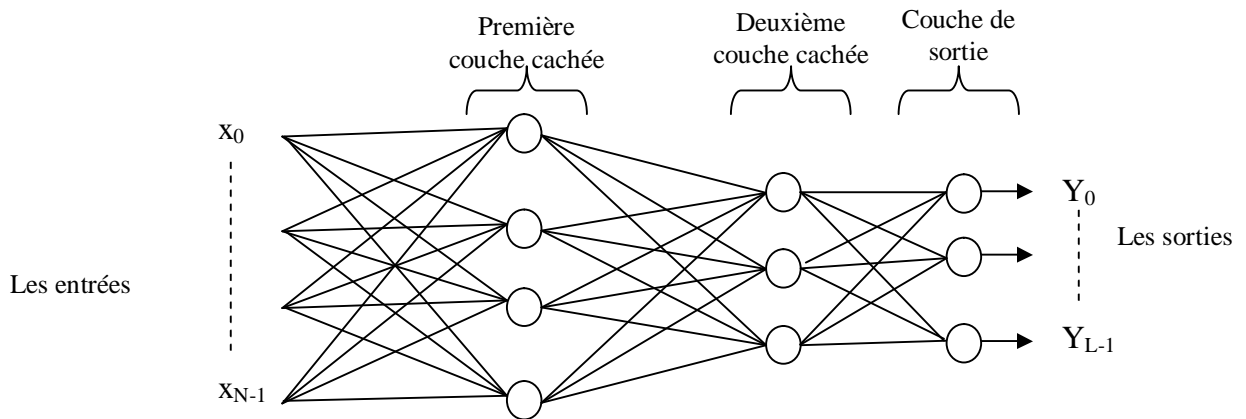


Figure 2-9. Perceptron Multi-Couche

Les capacités des perceptrons multicouches proviennent des non-linéarités utilisées dans les nœuds. Les capacités des perceptrons avec une, deux, et trois couches qui utilisent des non-linéarités illustrées dans le Tableau 2-10. La deuxième colonne dans cette Figure indique les types de régions de décision qui peuvent être formées par différents réseaux. Les deux colonnes suivantes montrent des exemples des régions de décision qui peuvent être formées pour le problème du OU exclusif et un problème avec des régions engrenées. La colonne en extrême droite donne des exemples des régions de décision les plus générales qui peuvent être formées.

Cette analyse fournit un aperçu sur le problème de choix du nombre de nœuds à employer dans un perceptron à deux couches.

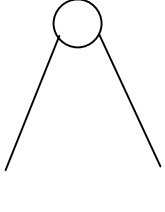
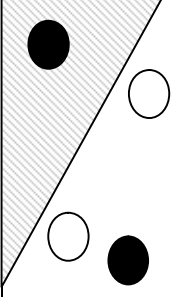
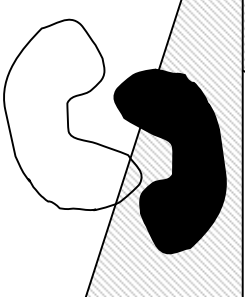
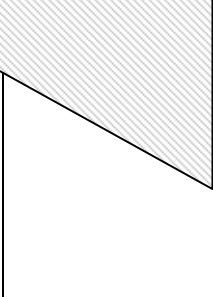
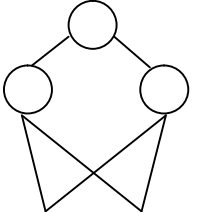
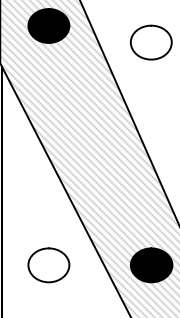
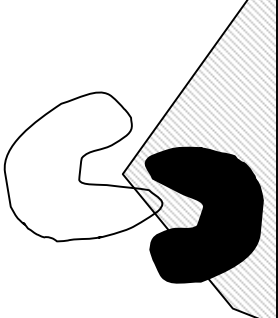
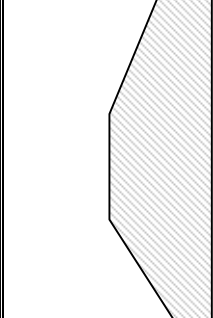
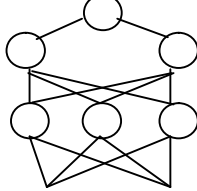
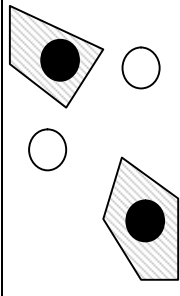
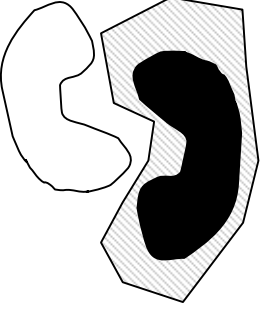
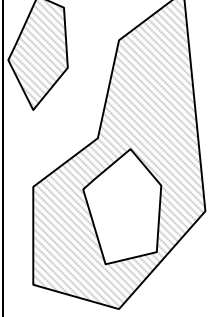
Architecture	Le problème de l' XOR	Type de la région de décision	Classes avec des régions chevauchées	Les formes de régions les plus générales
 Simple couche		Un demi plan borné par un hyperplan		
 Deux couche		Des régions convexes fermées ou ouvertes		
 Trois couches		Arbitraires		

Tableau 2-10. Interprétation géométrique du rôle des couches cachées

L'algorithme de rétro-propagation décrit dans la Figure 2-11 est une généralisation de l'algorithme de l'LMS. Il emploie une technique de recherche de gradient pour minimiser une fonction coût égal à l'erreur quadratique moyenne entre les sorties désirées et réelles. Le réseau est entraîné au début en initialisant aléatoirement les poids avec des petites valeurs, ainsi que les seuils internes et puis en présentant toutes les données d'apprentissage sur plusieurs itérations. Les poids sont ajustés après chaque itération jusqu'à ce que les poids convergent.

Figure 2-11. L'algorithme de rétropropagation

Il requière une fonction non linéaire différentiable. On utilise une fonction sigmoïde  $f(\alpha)$ , Figure 2-4, de la forme:

$$f(\alpha) = \frac{1}{1 + e^{-\alpha}}$$

**Etape1.** Initialiser les poids et les seuils, en les mettant à des petites valeurs aléatoires

**Etape 2.** Présenter la base de donnée

Présenter un vecteur d'entrée  $x_0, x_1, \dots, x_{N1-1}$  et spécifier le vecteur de sortie désirée

$$d_0, d_1, \dots, d_{N2-1}.$$

**Etape 3.** Calculer les sorties actuelles  $Y_0, Y_1, \dots, Y_{N-1}$ , en propageant l'entrée à travers les différentes couches du réseau de neurones comme suit:

$$x_j^m = f\left(\sum_{i=0}^{L-1} W_{ij} x_i^{m-1} - q_j^m\right), \quad 0 \leq j \leq N \quad \text{et} \quad 0 \leq m \leq M$$

$M$  Le nombre de couches

$W_{ij}$  Est le poids de connexion entre le  $j^{\text{eme}}$  neurone de la couche  $m$  et le  $i^{\text{eme}}$  neurone de la couche  $(m-1)$

$x_j^m$  Est la sortie de la couche  $m$

$x_i^{m-1}$  Est l'entrée de la couche  $m$  (ou la sortie de la couche  $(m-1)$ )

$q_j^m$  Est le seuil de la couche  $m$

**Etape 4.** L'ajustement des poids

On utilise un algorithme récursif, et en commençant par les nœuds de sortie, pour lesquels on utilisera la formule suivante d'ajustement des poids:

$$W_{ij}(t+1) = W_{ij}(t) + hd_j x_j^{M-1}$$

$h$  Est le terme du gain

$d_j$  Est le terme d'erreur sur le nœud  $j$ . Si le nœud  $j$  est une sortie alors:

$$d_j^M = Y_j(1 - Y_j)(d_j - Y_j)$$

Où  $d_j$  est la sortie désirée du nœud  $j$ , et où  $Y_j$  est la sortie calculée

Si  $j$  est un nœud interne alors le terme d'erreur sera:

$$d_j^{m-1} = x_j^m (1 - x_j^m) \sum_k d_k^m W_{jk}$$

Où  $k$  représente le  $k^{\text{ème}}$  nœud de la couche  $m$  (couche supérieure)

Les seuils sont ajustés de la même façon que les poids, et ça en les supposant comme étant des poids de connexion reliés à des entrées à valeurs . Parfois pour accélérer la convergence on utilise un terme  $a$  qu'on appelle momentum :

$$W_{ij}(t+1) = W_{ij}(t) + hd_j x_i + a(W_{ij}(t) - W_{ij}(t-1)) \quad \text{Où} \quad 0 \leq a \leq 1$$

**Etape 5.** Répéter à partir de "Etape 2" jusqu'à atteindre une erreur quadratique moyenne prédéfinie, ou jusqu'à atteindre un certain nombre d'itérations prédéfinies

Un théorème intéressant qui jeta une certaine lumière sur les performances des perceptrons multicouches a été prouvé par Kolmogorov. Ce théorème déclare que n'importe quelle fonction continue de  $N$  variables peut être calculée en utilisant seulement des additions linéaires et des fonctions non-linéaires continues et croissantes à seule variable. Il a déclaré qu'un perceptron de trois couches avec  $N(2N+1)$  neurones employant des non-

linéarités continues et croissantes peut calculer n'importe quelle fonction continue de N variables.

#### **2-6-6- Carte auto-organisatrice de Kohonen :**

Un important principe d'organisation des voies sensorielles dans le cerveau est que le placement des neurones soit ordonné et qu'il reflète souvent certaines caractéristiques physiques du stimulus externe ressenti. Par exemple, à chaque niveau de la voie auditive, des cellules nerveuses et des fibres sont arrangées anatomiquement en fonction de la fréquence qui obtient la plus grande réponse dans chaque neurone. Cette organisation tonotopique dans la voie auditive s'étend jusqu'au cortex auditif. Bien qu'une grande partie de l'organisation de bas niveau soit génétiquement prédéterminée, il est probable qu'une partie de l'organisation à des niveaux plus élevés soit créée pendant l'apprentissage par des algorithmes qui favorisent l'auto-organisation. Kohonen présente un algorithme semblable, qui produit ce qu'il appelle les cartes à organisation automatique semblables à celles qui se produisent dans le cerveau.

#### **2-7 Conclusion :**

Le développement qu'ont connu les réseaux de neurones artificiels a suscité un grand enthousiasme mais aussi des critiques. Beaucoup d'études comparatives comme celle effectuée dans cette thèse ont fournies une vue optimiste pour les réseaux de neurones artificiels, tandis que d'autres offrent une vue pessimiste. Pour de nombreuses tâches et problèmes, tel que la reconnaissance de formes, il n'y a pas une seule approche et qui pourrait dominer les autres. Le choix, le choix de la technique doit être en fonction de l'application en question, c'est pour ça qu'on doit comprendre et connaître les capacités, suppositions, la possibilité d'appliquer plusieurs approches dans différentes disciplines.



# **Chapitre 3 : La théorie de l'estimation**

### 3-1- Introduction :

Dans le chapitre précédent, on a vu que le problème de la détection est de déterminer quelle hypothèse est correcte; la présence de la cible  $H_1$ , ou bien son absence  $H_0$ . L'hypothèse correcte est supposée connue, mais un ou des éléments essentiels pour ce processus restent inconnues. La théorie de l'estimation prend part dans le traitement des observations reçues d'une façon optimale pour donner une valeur (faire une estimation) aux paramètres inconnus, qui sont en fait, les paramètres qui représentent la distribution du clutter dans le quel est noyé notre signal, pour ainsi avoir le taux de fausses alarmes le plus constant possible.

Il existe dans la littérature, plusieurs méthodes d'estimation des paramètres des distributions du clutter. Le choix de ces méthodes se repose principalement sur deux critères :

- ✚ La nature des paramètres à estimer (c à d, soit une variable aléatoire ou bien une constante), par exemple pour les paramètres de nature aléatoire, on utilise différentes méthodes tel que l'estimateur de Bayes et l'estimateur des probabilités a priori maximales, tandis que pour les paramètres constants on utilise plutôt, la méthode du maximum de vraisemblance, la méthode des moindres carrées, la méthode des moment ... etc.
- ✚ La complexité du modèle mathématique de la distribution, ainsi que le nombre de paramètres à estimer.

Comme la distribution qui représente notre modèle de clutter est en fonction de paramètres supposés constants, on ne va s'intéresser qu'aux méthodes d'estimation de paramètres constants.

### 3-2- Méthodes d'estimation de paramètres constants :

On suppose qu'en pratique le paramètre inconnue est considéré comme une constante, cette approche a comme but de faire une estimation raisonnable et de tester la qualité de cette estimation à travers quelques critères qui seront présenté plus loin, mais tout d'abord on va voir un aperçu sur les méthodes d'estimation de paramètres constants [13].

#### 3-2-1- Méthode du maximum de vraisemblance :

La procédure d'estimation et qui est la plus couramment utilisé dans le cas d'un paramètre constant est appelée la méthode du maximum de vraisemblance (ML) [2].

L'estimation ML d'un paramètre constant  $\theta$  est tout simplement la valeur de  $\hat{q}$  qui maximise la probabilité conditionnelle  $f_{y/q}(Y/q)$ , où  $Y$  est le vecteur d'observation de variables aléatoires indépendantes et identiquement distribuées, d'une autre façon l'estimée ML est la valeur de  $\hat{q}$ , l'estimation de  $\theta$  pour laquelle  $Y$  est le plus vraisemblable à avoir lieu.  $f_{y/q}(y/q)$  est appelée la fonction de vraisemblance [13], donc on aura :

$$\hat{q}_{ML} = \arg \max \left( f_{y/q}(y/q) \right) \quad (3-1)$$

Mais généralement on utilise la valeur logarithmique, comme suit :

$$\hat{q}_{ML} = \arg \max \left( \ln \left( f_{y/q}(y/q) \right) \right) \quad (3-2)$$

Dans ce cas les conditions requises pour trouver le maximum (3-2) sont :

$$\frac{d \left( f_{y/q}(y/q) \right)}{dq} = 0 \quad \text{et} \quad \frac{d^2 \left( f_{y/q}(y/q) \right)}{d^2 q} < 0 \quad (3-3)$$

Comme la fonction logarithme est une fonction monotone de  $y$ , alors la maximisation de  $f_{y/q}(y/q)$  est équivalente à la maximisation de  $\ln \left( f_{y/q}(y/q) \right)$ . Ainsi on arrive à la condition nécessaire pour obtenir l'estimation à maximum de vraisemblance qui est tout simplement comme suit:

$$\frac{d \ln \left( f_{y/q}(y/q) \right)}{dq} = 0 \quad (3-4)$$

### 3-2-2- La méthode des moments :

Etant donné une séquence de  $M$  variables aléatoires statistiquement indépendants, la méthode des moments revient à résoudre un système d'équations, et cela en égalisant les  $M$  premiers moments obtenus à partir de des échantillons observés

avec les moments de la distribution calculés théoriquement et qui sont des fonctions des  $M$  paramètres inconnus. Par conséquent, on aura ce genre de système d'équation :

$$\begin{cases} \hat{m}_1(\hat{q}_1, \hat{q}_2, \dots, \hat{q}_M) = m_1 \\ \hat{m}_2(\hat{q}_1, \hat{q}_2, \dots, \hat{q}_M) = m_2 \\ \vdots \\ \hat{m}_M(\hat{q}_1, \hat{q}_2, \dots, \hat{q}_M) = m_M \end{cases} \quad (3-5)$$

Après la résolution de ce système d'équation; soit directement (analytiquement) ou numériquement, on obtient une estimation des paramètres  $\hat{q}_m$ ,  $m = 1, 2, \dots, M$ . Cette méthode est la plus simple et la plus intuitive et elle ne nécessite que la connaissance a priori des moments théoriques des distributions, sauf que les performances de cet estimateur commence à se dégradées et avoir une plus grande variance au dessous d'un certain nombre d'échantillons.

### 3-3- Les critères d'un bon estimateur :

Dans l'estimation d'un vecteur de paramètres  $\theta$  à partir d'un ensemble de  $N$  échantillons, un ensemble de mesure de performances est utilisé pour quantifier et comparer les caractéristiques des différents estimateurs, en général l'estimée d'un vecteur est une fonction du vecteur d'observations  $y$ , de la longueur des observation et du modèle  $M$  du processus lui même, cette dépendance peut être exprimé comme suit :

$$\hat{q} = f(y, N, M) \quad (3-6)$$

Différents paramètres produisent différents résultats en fonction de la méthode d'estimation et des observations utilisées ainsi que l'influence des informations à priori. Comme les observations sont aléatoires, un même estimateur va produire différente estimation pour différentes observations pour un même processus. A partir de là l'estimée elle même est une variable aléatoire, qui a une moyenne et une variance, et elle peut être décrite par une fonction de densité de probabilité. Tandis que, pour la plus part des cas il est difficile de caractériser un estimateur en terme de la moyenne et la variance de l'erreur de l'estimateur [13].

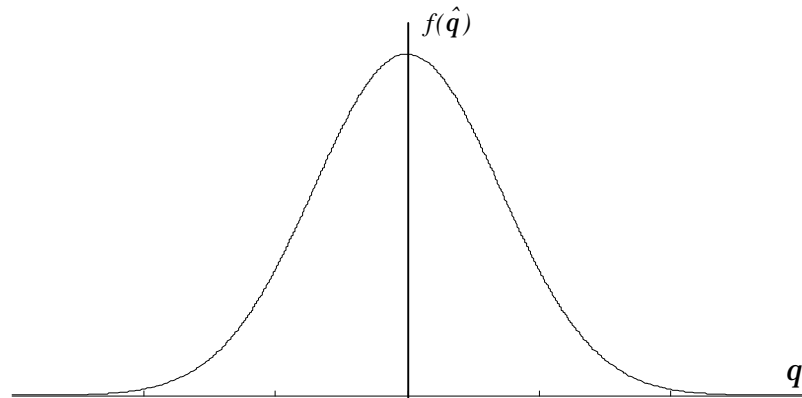


Figure 3-1. Variance d'un estimateur

Les mesures de performance les plus communément utilisées pour un estimateur sont les suivantes :

- (a) L'espérance mathématique de l'estimée:  $E[\hat{q}]$
- (b) Le Biais de l'estimée :  $E[\hat{q} - q] = E[\hat{q}] - q$
- (c) La covariance de l'estimée :  $Cov[\hat{q}] = E[(\hat{q} - E[\hat{q}])(\hat{q} - E[\hat{q}])^T]$

L'estimateur optimal a pour but d'avoir un biais nulle et une covariance minimum pour l'erreur d'estimation. Les caractéristiques désirables d'un bon estimateur peuvent être données comme suit :

- (a) **Estimateur non biaisé** : un estimateur du paramètre  $\theta$  est dit non biaisé si l'espérance mathématique de l'estimée est égale à la vraie valeur du paramètre c'est à dire :

$$E[\hat{q}] = q$$

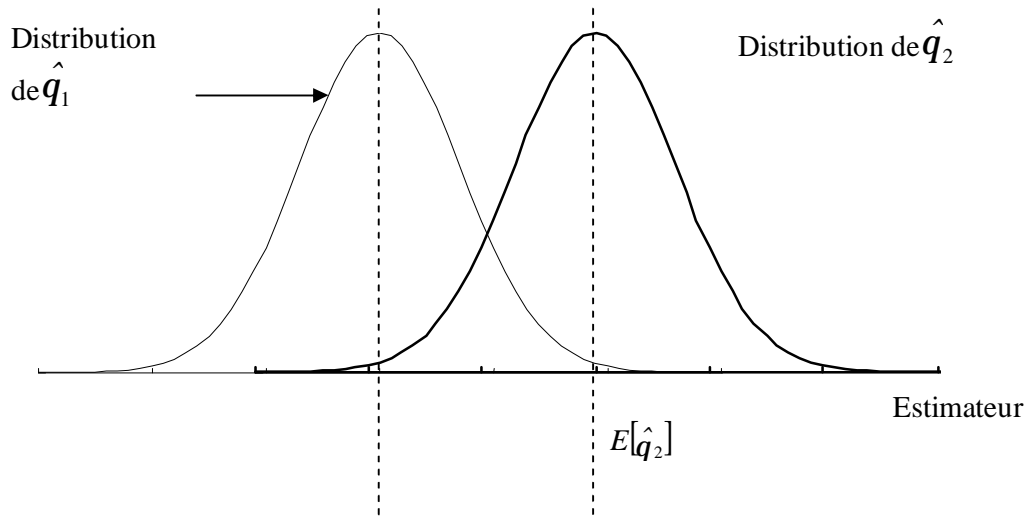


Figure 3-2. Le biais d'un estimateur

Un estimateur est dit asymptotiquement non biaisé si en augmentant le nombre d'observations  $N$  on a :

$$\lim_{N \rightarrow \infty} E[\hat{q}] = q \tag{3-7}$$

(b) **Estimateur efficace** : un estimateur non biaisé de  $\theta$  est dit efficace si il possède la plus petite matrice de covariance comparé avec les autre estimées de  $\theta$  non biaisées.

$$Cov[\hat{q}_{Efficace}] \leq Cov[\hat{q}] \tag{3-8}$$

où  $\hat{q}$  est n'importe quelle autre estimation de  $q$

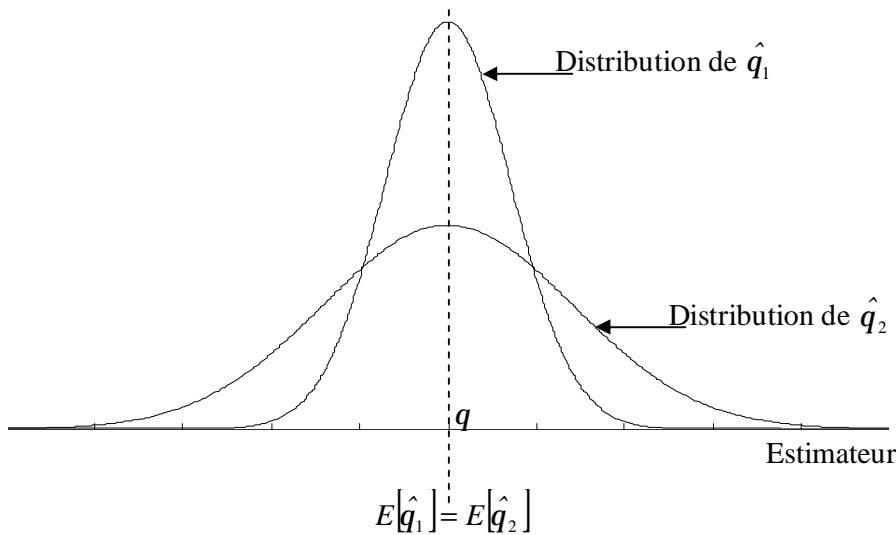


Figure 3-3. Estimateur efficace

- (c) **Estimateur consistant** : un estimateur est consistant si l'estimation s'améliore avec l'augmentation du nombre d'observation  $N$ , de façon que la valeur de l'estimation  $\hat{q}$  converge vers la vraie valeur  $q$  sachant que la valeur  $N$  qui deviendra infiniment grande [13].

$$\lim_{N \rightarrow \infty} P\left[|\hat{q} - q| > e\right] = 0 \quad (3-9)$$

Où  $e$  est une valeur arbitraire très petite, ou en d'autres termes :

$$\lim_{N \rightarrow \infty} E[\hat{q}_1] = q \quad (3-10)$$

$$\lim_{N \rightarrow \infty} E[\hat{q}_1^2] = 0 \quad (3-11)$$

### 3-4- Méthodes d'estimation de paramètres de quelques distributions de clutter :

Plusieurs approches ont été étudiées dans le domaine de l'estimation des paramètres de distribution du clutter tel que la méthode des moments d'ordre supérieur, moments fractionnaires, l'estimateur log normalisé ... etc. On va essayer de voir quelques applications des méthodes déjà évoquées :

#### 3-4-1- Estimation des paramètres de la distribution K :

Différentes approches ont été adoptées pour l'estimation des paramètres du clutter qui suit une distribution K, les estimateurs les plus utilisées dans la littérature sont basées sur la méthode des moments [14] et la méthode de maximum de vraisemblance et dernièrement des travaux ont été effectués sur les estimateurs à base de réseaux de neurones et [15].

En se basant sur le fait que la distribution K est en fonction de paramètres,  $n$  qui est le paramètre de forme, et  $a$  le paramètre d'échelle. Comme on a déjà vu dans la définition de la méthode des moments, et comme va faire l'estimation de deux paramètres  $n$  et  $a$ . Jahangir et al [16] ont suivis une approche très simple, on utilisant les moments du deuxième et quatrième ordre dont les expressions sont comme suit :

$$m_2 = \frac{\Gamma(2)\Gamma(n+2)}{\Gamma(n+1)}(2a)^2 \quad (3-17)$$

$$m_4 = \frac{\Gamma(3)\Gamma(n+3)}{\Gamma(n+1)}(2a)^4 \quad (3-18)$$

On résolvant le système d'équations suivant, on pourra trouver les estimées des paramètres, comme suit :

$$\frac{m_4}{m_2^2} = \frac{2\Gamma(n+3)\Gamma(n+1)}{\Gamma^2(n+2)} \quad (3-19)$$

Sachant que :  $\Gamma(n+1) = n\Gamma(n)$

$$n = \frac{4 - \frac{m_4}{m_2^2}}{\frac{m_4}{m_2^2} - 2} \quad (3-20)$$

Quant au paramètre  $a$ , il est calculé à partir de l'une des deux équations (3-18) ou (3-19). En utilisant l'équation (3-20):

$$a = \frac{m_1 \Gamma(n+1)}{\sqrt{p} \Gamma(n+1.5)} \quad (3-21)$$

Iskander et al [17] ont proposés une approche semblable à celle des moments d'ordre supérieur mais qui utilise les moments fractionnels, avec une formule plus généralisée :

$$r_{p,q} = \frac{m_{p+2q}}{m_p m_{2q}} \quad (3-22)$$

Cette méthode n'a besoin que d'un petit nombre d'opérations pour calculer les estimées, et ne requièrent pas la résolution d'équations non linéaires, par conséquent, l'estimateur sera plus performant que celui utilisant les moments d'ordre deux et d'ordre quatre. Ian et al [18] ont fait une étude comparative entre la méthode des moments et celle du maximum de vraisemblance. En faisant varier la longueur des



échantillons d'observation, ils ont déduis que les performances obtenus par la méthode de maximum de vraisemblance sont proches de ceux atteints par la méthode des moments, mais la méthode de maximum de vraisemblance est plus performante pour des tailles d'échantillons inférieures à 1000. Une approche un peut plus récente proposée par Iskander et *al* [17], qui ont proposés un estimateur basé sur les moments et le maximum de vraisemblance, on utilise la formule des moments de la distribution K :

$$E[x^k] = a^k \frac{\Gamma\left(\frac{k}{2} + \nu\right) \Gamma\left(\frac{k}{2} + 1\right)}{\Gamma(\nu)} \quad (3-23)$$

Pour trouver l'estimation du paramètre d'échelle  $a$  comme suit :

$$\hat{a} = \frac{1}{2} \left[ \frac{\Gamma(\hat{n}+1)}{\Gamma(0.5k+1) \Gamma(\hat{n}+1+0.5k)} \frac{1}{N} \sum_{i=1}^N x_i^k \right]^{\frac{1}{k}} \quad (3-24)$$

Ainsi que l'estimation du paramètre d'échelle trouvé par l'estimateur de maximum de vraisemblance [17] comme suit :

$$\hat{a} = \frac{1}{2} \exp\left(\frac{g - \Psi(\hat{n}+1)}{2} + \frac{1}{N} \sum_{i=1}^N \log(x_i)\right) \quad (3-25)$$

Où  $g=0.5772\dots$  est la constante d'Euler

$\Psi(\cdot)$  Est la fonction Digamma [19]

$$\Psi(z) = \frac{d[\ln\Gamma(z)]}{dz} \quad (3-26)$$

On combinant les deux équations précédentes on obtient :

$$\log\left[\frac{\Gamma(\hat{n}+1)}{\Gamma\left(\hat{n}+1+\frac{k}{2}\right)}\right] + \frac{k\Psi(\hat{n}+1)}{2} = \frac{1}{N} \sum_{i=1}^k \log(x_i) - \log\left[\frac{1}{N} \sum_{i=1}^k x_i^k\right] + \frac{kg}{2} + \log\left[\Gamma\left(1+\frac{k}{2}\right)\right]$$

On met :

$$g_k(\hat{n}) = \log\left[\frac{\Gamma(\hat{n}+1)}{\Gamma\left(\hat{n}+1+\frac{k}{2}\right)}\right] + \frac{k\Psi(\hat{n}+1)}{2} \quad (3-27)$$

En remplaçant (3-26) dans (3-27) on trouve :

$$g_k(\hat{n}) = \frac{1}{N} \sum_{i=1}^k \log(x_i^k) - \log \left[ \frac{1}{N} \sum_{i=1}^k \log(x_i^k) \right] + \frac{kg}{2} + \log \left[ \Gamma \left( 1 + \frac{k}{2} \right) \right] \quad (3-28)$$

On remarque que la partie droite de l'équation (3-28) est indépendante de  $\hat{n}$  :

$$\hat{n} = g_k^{-1} \left( \frac{1}{N} \sum_{i=1}^k \log(x_i^k) - \log \left[ \frac{1}{N} \sum_{i=1}^k x_i^k \right] + \frac{kg}{2} + \log \left[ \Gamma \left( 1 + \frac{k}{2} \right) \right] \right) \quad (3-29)$$

Où  $g_k^{-1}(\cdot)$  est la fonction inverse de  $g_k(\cdot)$

### 3-4-2 Estimation des paramètres de la distribution de Weibull et Lognormal :

Les deux paramètres (paramètres de forme et d'échelle) du clutter caractérisant les distributions de Weibull et Rayleigh sont estimés à l'aide de l'estimateur du maximum de vraisemblance. Le calcul du seuil de détection en utilisant cette méthode démontre une variance moindre et ainsi une perte moindre de performances du CFAR par rapport aux autres estimateurs tel que les moments, ainsi que la méthode bi-paramétrique qui est basée sur le maximum de vraisemblance [20], [21].

### 3-5 Conclusion:

Dans ce chapitre nous avons donné un aperçu de quelques méthodes d'estimation qui permettent d'avoir de bonnes approximations des valeurs de paramètres caractérisant quelques types de clutter, et qui offrent une variance acceptable pour ne pas dire minimale, pour ainsi produire un taux de fausse alarme possédant une petite variance. La généralisation du modèle du clutter a rendu l'application de ces méthodes très difficile, car elles nécessitent d'énormes calculs qu'on ne peut simplifier que par l'utilisation des méthodes numériques qui influent sur l'efficacité de ces estimateurs. Ces distributions généralisées dépendent d'un plus grand nombre de paramètres que les distributions classiques. Donc, il est indispensable de trouver d'autres méthodes d'estimations de leurs paramètres, permettant ainsi, de réduire les calculs par rapport aux méthodes classiques d'une part et d'avoir une estimation

efficace avec une variance minimale d'autre part. De nouvelles méthodes basées sur l'intelligence artificielle ont été proposées pour l'estimation des paramètres de distribution représentant le clutter du radar; comme les réseaux de neurones. Dans le chapitre suivant nous allons appliquer les réseaux de neurones dans l'estimation des paramètres du clutter de distribution  $Ki-2$  non centrale gamma, et comparer leurs performances avec ceux de l'estimateur à base de la méthode des moments.

# ***Chapitre4 : Introduction de l'estimateur neuronal***

#### 4-1- Introduction

Plusieurs modèles appelés modèles distribution généralisées ont été proposés pour la représentation des Clutter du radar. Ces modèles ont été proposés pour pallier aux faiblesses de la distribution K à modéliser un certains types de Clutter tel que le Clutter de terre SAR, Clutter de mer...etc. Le modèle composé généralisé, bien que validé par des mesures faites par des radars à hautes résolution, il possède une forme analytiques complexe, ce qui rend l'utilisation des méthodes conventionnelles (pour l'estimation des paramètres le caractérisant) déjà citées, difficiles et parfois totalement inefficace. Le modèle proposé par Shnidman [6], est représenté par le distribution Ki-2 Non Centrale Gamma. Cette dernière nous permet de produire des distributions avec une queue plus longue. Durant les dernières années, l'intelligence artificielle a connu un grand engouement dans tous les domaines et la recherche scientifique n'a pas était épargnée. Ceci est dû principalement à l'efficacité, flexibilité et facilité d'implémentation de ces méthodes. Dans ce chapitre nous allons utiliser les réseaux de neurones artificiels dans l'estimation des paramètres de la distribution Ki-2 Non Centrale Gamma qui est une distribution bi-paramétrique. Vu la difficulté d'utilisation de la méthode de vraisemblance [17], on a préféré retenir la méthode des moments en vue de faire une comparaison avec les réseaux de neurones artificiels.

Tout d'abord on va comparer la précision des deux estimateurs. Pour ensuite étudier l'effet des estimations sur la régulation de la fausse alarme. Enfin, on analysera les performances de la probabilité de détection. Pour commencer, on va introduire la méthode des moments dans l'estimation des paramètres de la distribution Ki-2 Non Centrale Gamma.

#### 4-2- Estimation des paramètres de la distribution NG par la méthode des moments:

Avant tout, dans ce qui suit on suppose que  $N_c=2$ . Rappelons que la distribution Ki-2 Non Centrale Gamma, dépend de deux autres paramètres ;  $X_c, L$ .

Pour l'estimation des paramètres  $X_c$  et  $L$ , on aura besoin préalablement de connaître deux des moments statistiques de la distribution Ki-2 Non Centrale Gamma. Vu la complexité de la formule des moments de la distribution Ki-2 Non Centrale Gamma, formule (1-55), et pour simplifier les calculs, on a choisi les moments du premier et du

second ordre. A partir de la formule (1-55) on obtient les formules mathématiques suivantes :

$$m_1 = N_c + X_c \quad (4-1)$$

$$m_2 = \frac{L+1}{L}(X_c+L)^2 + 2(N_c-L)(X_c+L) + (N_c-L)(N_c-L+1) \quad (4-2)$$

On peut obtenir une forme plus compacte du moment du second ordre, qui nous permettra d'avoir une estimation plus précise en évitant d'introduire des méthodes numériques pour la résolution du système d'équations (4-1) et (4-2), connaissant la forme de la variance qui est :

$$s^2 = \frac{X_c^2}{L} + N_c + 2X_c \quad (4-3)$$

On sait que :

$$s^2 = m_2 - m_1^2 \quad (4-4)$$

$$m_2 = s^2 + m_1^2 \quad (4-5)$$

$$m_2 = \frac{X_c^2}{L} + 2X_c + N_c + (N_c + X_c)^2 \quad (4-6)$$

A partir des deux formules (4-6) et (4-1), on obtient un simple système d'équation à deux inconnus  $X_c, L$  dont les solutions sont comme suit :

$$\left\{ \begin{array}{l} X_c = m_1 - N_c \end{array} \right. \quad (4-7)$$

$$\left\{ \begin{array}{l} L = \frac{X_c^2}{m_2 - 2X_c - N_c - (N_c + X_c)^2} \end{array} \right. \quad (4-8)$$

Sachant que les moments statistiques se calculent à partir des échantillons d'observations, comme suit:

$$\hat{m}_k = \frac{1}{N} \sum x_i^k, \quad k = 1, 2 \quad (4-9)$$

Où  $\{x_i, i = 1, 2, \dots, N\}$  sont les observations des différentes réalisations des événements  $X$ . Dans notre étude, et pour les deux méthodes, la méthode des moments ainsi que celle réseaux de neurones artificielles, on va prendre les deux cas extrêmes, c à d :

- ✚  $L = 1$ , dans ce cas tous les moyennes des cellules du Clutter ont la même réponse, la pdf du niveau moyen de la puissance devient :

$$f(c/Xc) = \frac{1}{Xc} \exp(-c/Xc) \quad (4-10)$$

L'expression (4-10) représente la pdf de la distribution exponentielle. Elle est obtenue à partir de la somme des deux sorties (en phase, en quadrature de phase) du détecteur quadratique, qui sont des variables aléatoires gaussiennes à moyenne nulle et de variance  $\frac{Xc}{Nc} \sigma^2$ . Dans ce cas, les moyennes des cellules sont complètement corrélées (toutes les sorties des  $Nc$  cellules sont égales).

- ✚  $L = Nc = 2$ , dans ce cas toutes les moyennes des cellules sont indépendantes, donc la pdf du niveau moyen de la puissance devient :

$$f(c/Xc, L) = \frac{c^{L-1} Nc^L}{\Gamma(Nc) Xc^L} \cdot \exp(-c Nc/Xc) \quad (4-11)$$

Cette dernière expression représente la pdf de la distribution Gamma, qui est la somme des moyennes  $m_k$  des sorties (en phase, et en quadrature de phase) de  $Nc$  cellules. Ces moyennes sont des variables aléatoires gaussiennes de moyennes nulles et de variance  $\frac{Xc}{Nc} \sigma^2$  chacune. Les fluctuations entre les cellules dépendront du nombre de cellules  $Nc$ .

#### 4-2-1- L'utilisation des réseaux de neurones dans l'estimation de la distribution K :

Plusieurs approches à base de réseaux de neurones ont été proposées dans l'estimation des paramètres des distributions qui modélisent les Clutter de radar. La plus

part des approches se penchent sur l'estimation des paramètres de la distribution K. En 1996, Jahangir *et al* [16], ont proposés l'utilisation en premier lieu d'un perceptron multicouches comme un modèle d'estimateur, avec une sortie, 16 neurones dans la couche cachée, et 225 entrées, aux quelles on présente directement 225 échantillons qui suivent une distribution K, ce modèle donne de bonnes estimations mais il présente une variance plutôt importante. Alors il a été décidé d'entraîner le réseau de neurones à l'aide des moments statistiques (moyenne, variance, variance du log de la puissance, ou encore en combinant entre les différents paramètres déjà cités) à partir de 2000 observations toutes de longueurs de 225 échantillons. Le réseau de neurones présente une bonne approximation, en le comparant avec d'autres estimateurs, spécialement, quand la variance est utilisée pour l'apprentissage du réseau de neurones.

Wachowiak *et al* en 2002 [15], ont proposés un réseau de neurones pour l'estimation du paramètre de forme de la distribution K. Ils ont effectué l'apprentissage d'un réseau de neurones multicouches (deux couches cachées) à deux entrées et une sortie (qui sera la sortie du réseau de neurones) en utilisant l'algorithme d'apprentissage de rétropropagation, et en utilisant l'inverse des carrés des moments statistiques du troisième et quatrième ordre comme entrées. L'apprentissage à été effectuer pour plusieurs longueurs d'échantillons  $N= 250, 2500$  et  $10000$  échantillons.

En comparant avec d'autres méthodes telles que la méthode des moments, la méthode ML/MOM (chapitre 2), ils ont pu conclure que l'estimateur à base de réseaux de neurones est plus précis et plus consistant.

#### **4-2-2- L'utilisation des réseaux de neurones dans l'estimation des paramètres de la distribution Ki-2 Non Centrale Gamma :**

Vu la complexité et l'inefficacité de l'utilisation des méthodes classiques (à part la méthode des moments) dans l'estimation des paramètres de la distribution Ki-2 Non Centrale Gamma, on a proposé une approche à base de réseaux de neurones. Comme montré sur la Figure 4-1. Le réseau proposé est un réseau à deux couches cachées, une couche d'entrées et une couche de sortie, la couche d'entrée est constituée de deux entrées, où sera injectée l'ensemble d'apprentissage. 25 neurones dans la première couche cachée, et 19 neurones dans la deuxième couche cachée, et comme on aura besoin d'estimer les deux paramètres ( $X_c, L$ ) la couche de sortie contient deux neurones.



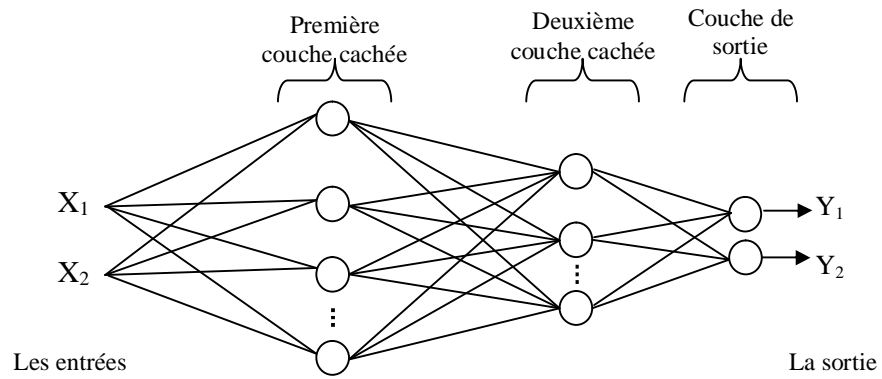


Figure 4-1 : L'estimateur à base de RNA

**4-2-2-1- L'ensemble d'apprentissage :**

C'est une base de données, qui contient un vecteur de paires d'entrées/sorties, à partir des quels on va établir l'ajustement des paramètres (poids et biais) du réseau de neurones. Notre base de données est constituée tout naturellement des valeurs réelles des paramètres  $Xc$  et  $L$ . Les valeurs pour les quelles le réseau de neurones a été entraîné s'étendent comme suit:

$$\begin{cases} 1 \leq Xc \leq 15 \\ L = \begin{cases} 1 \\ 2 \end{cases} \end{cases}$$

Ces derniers sont les sorties du réseau de neurones artificielle, et a partir des quels ont va calculer les moments du premier et du second ordre qui seront présentés comme entrées du réseau, donc la base de données est constituée comme suit:

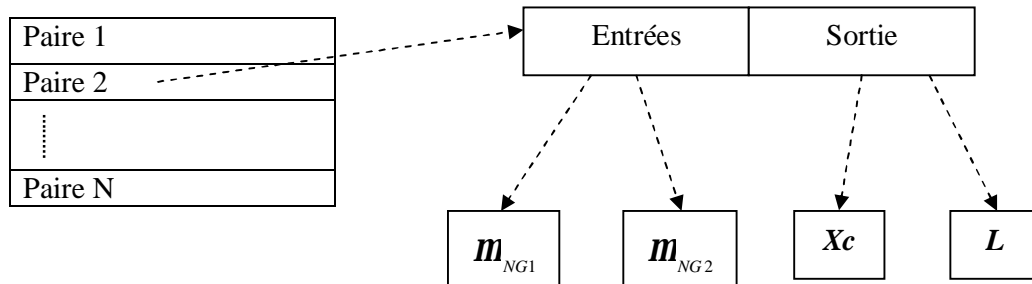


Figure 4-2 : la composition de la base de données

Sachant que les moments du premier et du deuxième ordre sont calculés comme on a vu précédemment dans la formule (4-9) à partir des échantillons des observations du Clutter, l'apprentissage du réseau de neurones a été effectué avec différentes tailles d'échantillons,  $K=50,100,200,500,1000,2000$ .

#### 4-2-2-2- La génération des échantillons du Clutter:

Après la génération de  $2Nc$  signaux gaussiens de variance  $\sigma^2$ , et de moyennes :  $m_{kq}$ ,  $k=1,2,\dots,Nc$ , les moyennes des entrées du détecteur en quadrature de phase.  $m_{ke}$ ,  $k=1,2,\dots,Nc$ , les moyennes des entrées du détecteur en phase.

Sachant que la somme des carrés de ces moyennes  $I$ , suit une distribution Gamma, ou  $I$  est appelé le paramètre de non centralité, où:

$$I = \sum_{k=1}^{Nc} m_k^2$$

La distribution du niveau moyen de la puissance du Clutter est modélisée par l'expression (1-48) (chapitre1). Sachant que  $m_k^2 = m_{kq}^2 + m_{ke}^2$ , cela revient à présenter à l'entrée du détecteur quadratique des signaux gaussiens à variance  $\sigma^2$  et moyennes  $m_k$ ,  $k=1,2,\dots,Nc$ . Pour chaque paire de paramètres ( $Xc, L$ ) on génère 30 vecteurs contenant  $N$  variables aléatoires suivant une distribution Ki-2 Non Centrale Gamma. Comme le paramètre  $Xc$  peut prendre les valeurs suivantes  $Xc=0.5, 1, 1.5, 2, 2.5, \dots, 14.5, 15$ , et  $L=1,2$  donc pour chaque tailles d'échantillon  $N$  ( $N=50,100,200,500,1000,2000$ ), on dispose d'un vecteur de longueur de 1800 paires (entrées,sorties). L'apprentissage du réseau de neurones, est basé sur la descente du gradient ou la rétropropagation, qui a déjà été présenté dans le chapitre 2.

Il faut noter qu'on a utilisé la racine carrée des moments du second ordre, afin de minimiser sa variance, et avoir un changement linéaire de ce dernier en fonction de  $Xc$

#### 4-2-2-3- La validation

La validation du réseau de neurones a été vérifiée en utilisant deux autres bases de données qui contiennent d'autres valeurs de moments calculées à partir des

valeurs des paramètres  $X_c$  et  $L$  qui sont comprises dans le même intervalle  $[0.5, 15]$ , mais avec des valeurs différentes de celles choisies pour l'apprentissage.

A partir de ce point, on peut juger si le réseau de neurones peut être considéré comme un bon estimateur, et c'est le cas, seulement si le comportement du réseau est le même pour toutes les bases de données utilisées dans le test de généralisation, c a d, si le réseau de neurones donne une bonne approximation pour toutes les valeurs de paramètres  $X_c$  et  $L$ .

Les tableaux suivants montrent l'erreur quadratique moyenne obtenue pour les différents ensembles ; apprentissage et validation :

	Erreur quadratique moyenne					
	2000	1000	500	200	100	50
Ensemble d'apprentissage	0.00455	0.00701	0.00937	0.0201	0.0413	0.137
Ensemble de validation 1	0.00457	0.00701	0.00922	0.0204	0.0413	0.137
Ensemble de validation 2	0.00461	0.00716	0.01055	0.0225	0.0410	0.138

Tableau (4-1) L'erreur quadratique moyenne des différents réseaux de neurones

Quant au tableau suivant, il montre les résultats obtenus après utilisation des réseaux de neurones autant qu'estimateurs. Nous avons calculé l'erreur quadratique moyenne et nous l'avons comparé avec celle trouvée en utilisant la méthode des moments.

	Erreur quadratique moyenne											
	RNA						Moments					
	N=2000	N=1000	N=500	N=200	N=100	N=50	N=2000	N=1000	N=500	N=200	N=100	N=50
Ensemble d'apprentissage	0.0432	0.0777	0.247	0.309	0.613	1.692	0.109	0.144	0.220	0.440	0.785	1.92
Ensemble de validation 1	0.0426	0.0791	0.220	0.341	0.615	1.704	0.107	0.140	0.215	0.436	0.835	2.17
Ensemble de validation 2	0.0473	0.0857	0.181	0.326	0.595	1.671	0.107	0.137	0.212	0.423	0.791	1.65

Tableau (4-2) Comparaison entre l'erreur quadratique moyenne des deux estimateur

Ainsi pour pouvoir juger que notre estimateur est un bon estimateur, on va se baser sur deux critères, sur lesquelles on se basera pour établir une comparaison avec l'estimateur à base de moments.

#### ✚ Biais de l'estimateur :

Les expressions des biais sont données comme suit :

$$\begin{aligned} \text{Biais}_{X_c} &= E[\hat{X}_c] - X_c \\ \text{Biais}_L &= E[\hat{L}] - L \end{aligned} \quad (4-13)$$

Où  $\hat{X}_c$  et  $\hat{L}$  sont les valeurs estimées des paramètres  $X_c$  et  $L$  respectivement, plus le biais tend vers zéro, plus notre estimateur donne de bonne approximation. Mais cela est insuffisant pour juger l'efficacité de notre estimateur. Il faut en plus du biais, calculer la variance ou l'écart type qui nous aidera dans notre comparaison entre les deux estimateurs.

#### ✚ Variance de l'estimateur :

Ayant comme formule les expressions suivantes :

$$\begin{aligned} \text{Var}_{X_c} &= E[\hat{X}_c^2] - E^2[\hat{X}_c] \\ \text{Var}_L &= E[\hat{L}^2] - E^2[\hat{L}] \end{aligned} \quad (4-12)$$

#### 4-2-3- Résultats et discussion :

Les tableaux suivants démontrent clairement les performances des deux estimateurs en se basant principalement sur les deux critères, biais et variances (calculés à partir de différentes valeurs de  $X_c$  et  $L$ , et pour les longueurs d'échantillons  $N=50,100,200,500,1000,2000$ , à partir de 1000 essais effectués.). On peut observer que dans la majorité des cas, Tableaux 4-3 à 4-8 ainsi que pour toutes les valeurs de  $X_c$  et  $N$  que l'estimateur neuronal montre une nette avance sur celui basé sur la méthode des moments et ça dans l'estimation du paramètre de fluctuation  $L$ . Ceci est clair par la comparaison des biais et variances, il est aussi à noter que dans la plus grande partie des cas, les valeurs de la puissance moyenne  $X_c$ , les biais de l'estimée  $L$  atteint par l'estimateur neuronale sont inférieurs à ceux calculés à partir de la méthode des moments. Dans ce cas précis, on peut observer que l'estimation du paramètre de fluctuation  $L$  garde de très bon biais et variances alors que l'estimateur classique perd

considérablement de ses performances, ceci est clairement illustré dans le Tableau 4-8 pour les valeur de  $N=50, X_c=3.5$  et  $L=2$  la variance est supérieur à 16.

D'autre part, on peut noter que les résultat obtenus pour  $N=2000$  et  $N=1000$ , sont meilleurs que ceux obtenus pour  $N=500, 200, 100, 50$ , pour les deux méthodes utilisées et ce résultat est complètement prévisible, sachant que la diminution du nombre d'échantillons augmente le bruit dans les moments du premier et second ordre. Ce bruit peut se traduire par une grande variation entre les valeurs des moments, donc pour  $N=2000, 1000$  les moments du premier est second ordre sont moins bruités que les moments calculés pour des valeurs de  $N$  inférieur. Par contre, on peut relever que les estimations du paramètre de puissance  $X_c$  sont moins bonnes pour l'estimateur neuronal, mais qui restent toutefois acceptables et ce que nous pouvons voir à partir des tableaux, c'est que même si le réseau de neurone présente un biais de  $\hat{X}_c$  assez élevé par rapport à celui atteint par la méthode des moments, la variance de cette dernière reste dans le même ordre de grandeur et parfois inférieure que celle de l'estimateur a base de la méthode des moments.

	Estimateur à base de RNA			Estimateur à base des Moments		
	Moyennes $\begin{cases} X_c \\ L \end{cases}$	Biais	Variances	Moyennes	Biais	Variances
$X_c=1$	1.0874	0.0874	0.000835	1.0006	0.00062	0.0022
$L=2$	1.6408	0.3591	0.01831	0.403	1.596	0.0016
$X_c=5$	5.0127	0.0127	0.188	4.9965	0.00343	0.01816
$L=1$	1.00016	0.000167	0.000000325	0.932	0.0678	0.00321
$X_c=7$	6.9624	0.0376	0.0197	6.9917	0.0083	0.0199
$L=2$	1.9999	0.0001	0.000000169	1.8611	0.138	0.0129
$X_c=9$	9.0424	0.0424	0.00552	8.996	0.00402	0.053
$L=1$	1.00008	0.000081	0.000000015	0.979	0.0204	0.0029
$X_c=12$	12.076	0.0767	0.0476	11.9934	0.0066	0.0486
$L=2$	1.998	0.00183	0.00000306	1.956	0.0436	0.0105
$X_c=15$	14.955	0.0442	0.0676	14.999	0.00082	0.0719
$L=2$	1.997	0.0025	0.00000167	1.969	0.03024	0.00957

Tableau 4-3 Valeur des biais et variance des estimateurs de  $X_c$  et  $L$  pour  $N=2000, N_c=2$

	Estimateur à base de RNA			Estimateur à base des Moments		
	Moyennes $\begin{cases} X_c \\ L \end{cases}$	Biais	Variances	Moyennes	Biais	Variances
$X_c=2$	1.6942	-0.3058	0.0132	2.0047	0.0047	0.0101
$L=1$	1.0146	0.0146	0.0028	0.6732	-0.3268	0.0054
$X_c=3$	3.1084	0.01084	0.0101	2.995	-0.005	0.013
$L=2$	1.9719	-0.0280	0.00797	1.4022	-0.5977	0.02
$X_c=5$	4.9705	-0.0294	0.04299	4.9933	-0.0066	0.0377
$L=1$	1.0004	0.00045	0.000079	0.93706	-0.0629	0.00653
$X_c=7$	6.9609	0.03911	0.06115	7.0083	0.0083	0.0631
$L=1$	1.00027	0.00027	0.000013	0.96718	-0.0328	0.0056
$X_c=9$	9.0514	0.0514	0.0634	9.0029	0.0029	0.0614
$L=2$	1.999	0.00098	0.0000000086	1.9203	0.07967	0.0214
$X_c=15$	14.9338	-0.0662	0.1328	14.9959	0.00401	0.1389
$L=2$	1.9989	-0.001	0.0000000025	1.9641	-0.0358	0.0188

Tableau 4-4 Valeur des biais et variance des estimateurs de  $X_c$  et  $L$  pour  $N=1000, N_c=2$ 

	Estimateur à base de RNA			Estimateur à base des Moments		
	Moyennes $\begin{cases} X_c \\ L \end{cases}$	Biais	Variances	Moyennes	Biais	Variances
$X_c=2.5$	2.1524	-0.3476	0.0109	2.5014	0.0014	0.2139
$L=2$	1.9324	-0.0675	0.0356	1.244	-0.756	0.03952
$X_c=4$	3.8942	-0.1058	0.0498	8.0093	0.0093	0.0505
$L=1$	1.1455	0.1455	0.00911	1.9117	-0.0882	0.0139
$X_c=8$	8.1214	0.1214	0.1231	8.0093	-0.0154	0.102
$L=2$	1.9979	-0.002	0.0012	1.9117	0.0034	0.0506
$X_c=13.5$	13.7545	0.2545	0.5138	13.4846	-0.0154	0.4119
$L=1$	1.00089	0.00089	0.00012	1.0034	0.0034	0.01

Tableau 4-5 Valeur des biais et variance des estimateurs de  $X_c$  et  $L$  pour  $N=500, N_c=2$

	Estimateur à base de RNA			Estimateur à base des Moments		
	Moyennes $\begin{cases} X_c \\ L \end{cases}$	Biais	Variances	Moyennes	Biais	Variances
$X_c=2$	2.00042	0.00042	0.0383	2.0063	0.0063	0.04023
$L=2$	1.6437	-0.3562	0.0549	1.072	-0.9281	0.1001
$X_c=5.5$	5.461	-0.0389	0.2089	5.5045	0.00456	0.2106
$L=1$	1.0537	0.0537	0.0209	0.9864	-0.0136	0.0314
$X_c=8$	7.9144	-0.08556	0.2310	8.0012	-0.0012	0.2452
$L=2$	1.9436	-0.0563	0.02114	1.9641	-0.0358	0.1275
$X_c=12$	11.8741	-0.1258	0.8682	11.969	-0.0309	0.8122
$L=1$	1.0314	0.03147	0.01392	1.0248	0.0248	0.0285
$X_c=14.5$	14.4482	-0.0517	1.2436	14.501	0.001	1.149
$L=1$	1.0278	0.0278	0.0136	1.0358	0.0358	0.0278

Tableau 4-6 Valeur des biais et variance des estimateurs de  $X_c$  et  $L$  pour  $N=200, N_c=2$ 

	Estimateur à base de RNA			Estimateur à base des Moments		
	Moyennes $\begin{cases} X_c \\ L \end{cases}$	Biais	Variances	Moyennes	Biais	Variances
$X_c=1$	1.0538	0.0538	0.0382	0.9975	-0.0025	0.2482
$L=1$	1.5557	0.5557	0.00779	0.376	-0.6239	0.353
$X_c=5$	5.1683	0.1683	0.1945	5.0067	0.0067	0.2482
$L=2$	1.847	-0.1529	0.0379	1.8937	-0.1063	0.3530
$X_c=7$	6.8261	-0.1738	0.4313	7.0323	0.0323	0.6961
$L=1$	1.0154	0.0154	0.0038	1.0461	0.0460	0.0642
$X_c=9$	8.9114	-0.0885	0.7244	8.9876	-0.0124	0.6094
$L=2$	1.9014	-0.0985	0.0388	2.0525	0.0525	0.2817
$X_c=15$	15.2783	0.2783	1.6551	15.0102	0.0102	1.5078
$L=2$	1.9256	0.0743	0.0352	2.0944	0.0944	0.2249

Tableau4-7 Valeur des biais et variance des estimateurs de  $X_c$  et  $L$  pour  $N=100, N_c=100$

	Estimateur à base de RNA			Estimateur à base des Moments		
	Moyennes $\begin{cases} X_c \\ L \end{cases}$	Biais	Variances	Moyennes	Biais	Variances
$X_c=1$	0.9624	0.0375	0.0302	0.9797	0.0203	0.1022
$L=1$	1.6218	0.6218	0.0002	0.44	0.56	0.4812
$X_c=3.5$	3.3175	0.1824	0.1605	3.4836	0.0163	0.3031
$L=2$	1.7354	0.2645	0.0212	1.9134	0.0865	16.022
$X_c=8$	7.3153	0.6847	1.6104	7.9431	0.0568	1.5574
$L=1$	1.6906	0.6905	0.0712	1.1288	0.1288	0.1417
$X_c=9$	8.398	0.6019	1.9653	8.96806	0.03194	1.1694
$L=2$	1.9517	0.0482	0.0064	2.2761	0.27615	1.0141
$X_c=12$	11.4408	0.5591	4.4495	11.9266	0.0733	3.3293
$L=2$	1.7297	0.7297	0.0862	1.1516	0.1515	0.137

Tableau 4-8 Valeur des biais et variance des estimateurs de  $X_c$  et  $L$  pour  $N=50$ ,  $N_c=2$ 

#### 4-3- Les effets des estimations sur la régulation de la fausse alarme :

Dans ce qui suit, nous allons tester l'efficacité de notre estimateur à base de réseau de neurones artificiels, et cela en étudiant son influence sur la variation de la fausse alarme, tout en la comparant avec celle de l'estimateur à base de la méthode des moments. Comme mentionné dans le premier chapitre, la fausse alarme est la déclaration de la présence d'une cible lorsque elle est absente, c à d :

$$Pfa = P(q > T / H_0) \quad (4-14)$$

Où  $T$  est le seuil de détection.

Sachant que la probabilité  $P(q > T / H_0)$  peut être remplacée par l'expression mathématique suivante :

$$Pfa = \int_T^{\infty} f_{Y/H_0}(Y/H_0) dy \quad (4-15)$$



Où  $f_{Y/H_0}$  est la densité de probabilité conditionnelle correspondant à l'hypothèse  $H_0$ . On sait que :

$$\int_0^{\infty} f_{Y/H_0}(Y/H_0)dy = 1 \quad (4-16)$$

A partir de là on peut obtenir la probabilité de fausse alarme comme suit:

$$Pfa = 1 - \underbrace{\int_T^{\infty} f_{Y/H_0}(Y/H_0)dy}_a \quad (4-17)$$

Où le terme  $a$  définit l'expression de la fonction de répartition de  $Y$  donnée par [22]

$$F_Y(Y) = P_{NG}(Y, N_c, X_c, L) = \sum_{k=0}^{\infty} \frac{\Gamma(L+k)}{\Gamma(L)k!} \left( \frac{1}{1+X_c/L} \right)^L \left( \frac{X_c/L}{1+X_c/L} \right)^k \left[ 1 - \sum_{m=0}^{N_c-1+k} \exp(-T) \frac{T^m}{m!} \right] \quad (4-18)$$

Alors la probabilité de fausse alarme devient :

$$Pfa = 1 - F_Y(Y) \quad (4-19)$$

En remplaçant (4-18) dans (4-19) on obtient :

$$Pfa = 1 - \sum_{k=0}^{\infty} \frac{\Gamma(L+k)}{\Gamma(L)k!} \left( \frac{1}{1+X_c/L} \right)^L \left( \frac{X_c/L}{1+X_c/L} \right)^k \left[ 1 - \sum_{m=0}^{N_c-1+k} \exp(-T) \frac{T^m}{m!} \right] \quad (4-20)$$

Evidemment, comme on peut le remarquer, l'expression de la probabilité de la fausse alarme est en fonction des paramètres  $X_c$ ,  $L$ ,  $N_c$ . A partir de là, on va étudier l'influence de ces paramètres sur la variation de la probabilité de fausse alarme (notons que  $N_c$ , est constant et égale à 2) en passant par les étapes suivantes :

- ✚ L'estimation des paramètres  $X_c$ ,  $L$ , en premier lieu en utilisant la méthode des moments. Ensuite, en utilisant le réseau de neurones.

- ✚ Utiliser ces derniers paramètres pour calculer le seuil  $T$  qui correspond à la probabilité de fausse alarme nominale  $Pfan$ , et cela en résolvant l'équation non linéaire suivante :

$$Pfan = 1 - \sum_{k=0}^{\infty} \frac{\Gamma(\hat{L} + k)}{\Gamma(\hat{L})k!} \left( \frac{1}{1 + \hat{X}_c/\hat{L}} \right)^L \left( \frac{\hat{X}_c/\hat{L}}{1 + \hat{X}_c/\hat{L}} \right)^K \left[ 1 - \sum_{m=0}^{N_c-1+k} \exp(-T) \frac{T^m}{m!} \right]$$

- ✚ Remplacer la valeur trouvée de  $T$  dans l'équation (4-20), mais cette fois ci en utilisant les vraies valeurs de  $X_c$  et  $L$ , pour enfin déterminer la vraie valeur de la probabilité de fausse alarme.

Enfin, répéter cette opération 1000 fois pour avoir une bonne représentation de la variation de la probabilité de fausse alarme.

Ces opérations sont exécutées en utilisant les deux estimateurs ; l'estimateur à base de réseaux de neurones et celui des moments et pour plusieurs valeurs des paramètres de la distribution Ki-2 Non Centrale Gamma, ainsi que pour différentes longueurs d'observations  $N=50, 100, 200, 500, 1000, 2000$ , et pour une probabilité nominale de fausse alarme  $Pfan=0.001$ .

#### 4-3-1- Résultats et discussion :

Les Figures suivantes montrent l'effet de l'introduction de l'estimateur neuronal sur la variation de la probabilité de fausse alarme et sa comparaison avec la variation qu'introduit l'estimateur à base des moments. En variant  $X_c$ ,  $L$  et pour différentes longueurs d'observation. On remarque à partir des Figures 4-2 et 4-3, qu'on obtient une meilleure régulation de la probabilité de fausse alarme en utilisant l'estimateur neuronal. On peut aussi remarquer, sur les Figures 4-4 jusqu'à 4-6 que la diminution de la longueur d'échantillon affecte considérablement la régulation de la Pfa qui devient proche de celle obtenus avec les estimateur à base des moments. On peut aussi relever une grande déviation par rapport à la probabilité de fausse alarme nominale sur la Figure 4-6-a, ce qui engendre un biais assez important et donc une grande perte des performance de détection par rapport à l'estimateur à base de la méthode des moments.

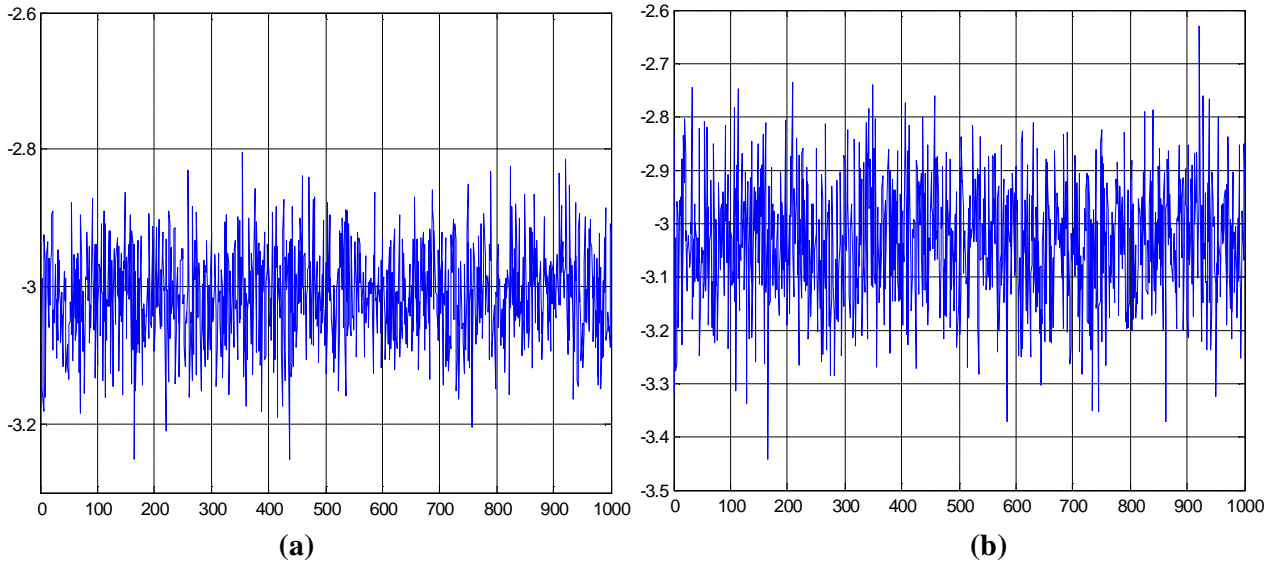


Figure 4-2 la variation de la probabilité de fausse alarme en fonction du nombre d'essai, pour  $N=2000$ ,  $X_c=7$ ,  $L=2$ . (a) l'estimateur neuronal, (b) l'estimateur à base de la méthode des moments.

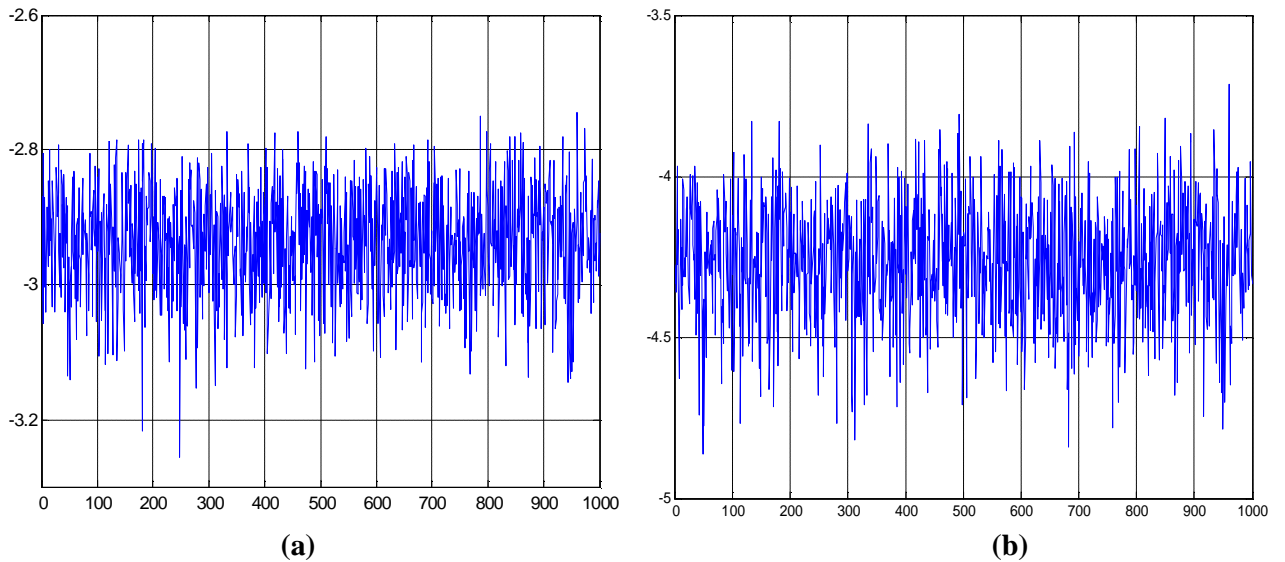


Figure 4-3 la variation de la probabilité de fausse alarme en fonction du nombre d'essai pour  $N=1000$ ,  $X_c=1$ ,  $L=1$ . (a) l'estimateur neuronal, (b) l'estimateur à base de la méthode des moments.

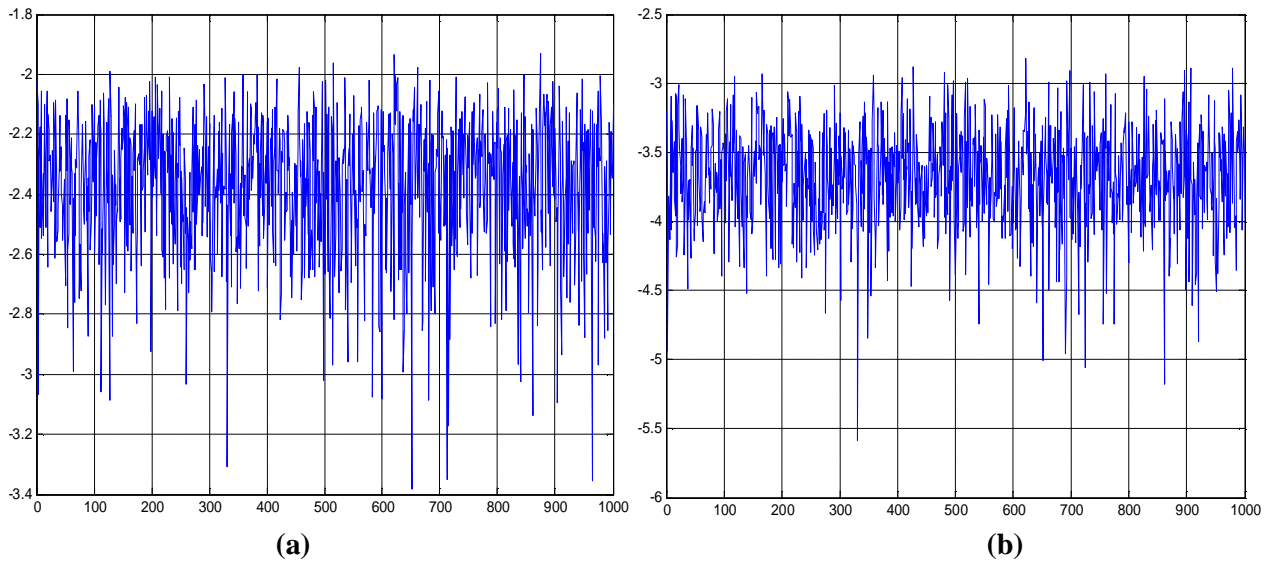


Figure 4-4 la variation de la probabilité de fausse alarme en fonction du nombre d'essai pour  $N=200$ ,  $X_c=2$ ,  $L=2$ . (a) L'estimateur neuronal, (b) l'estimateur à base de la méthode des moments.

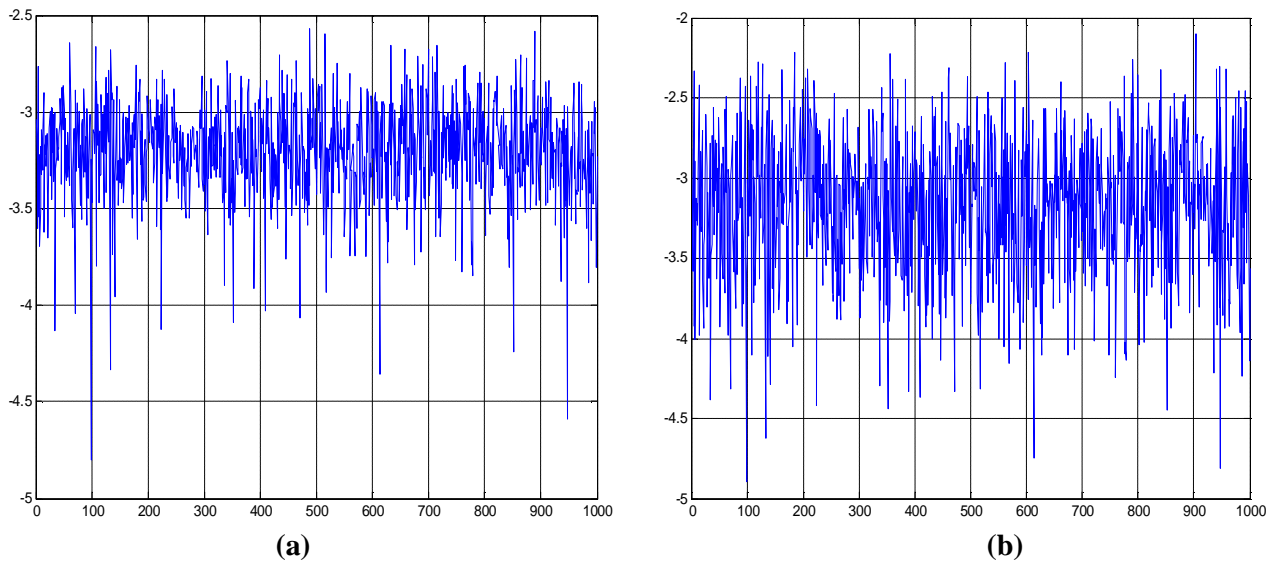


Figure 4-5 la variation de la probabilité de fausse alarme en fonction du nombre d'essai pour  $N=100$ ,  $X_c=7$ ,  $L=1$ . (a) L'estimateur neuronal, (b) l'estimateur à base de la méthode des moments.

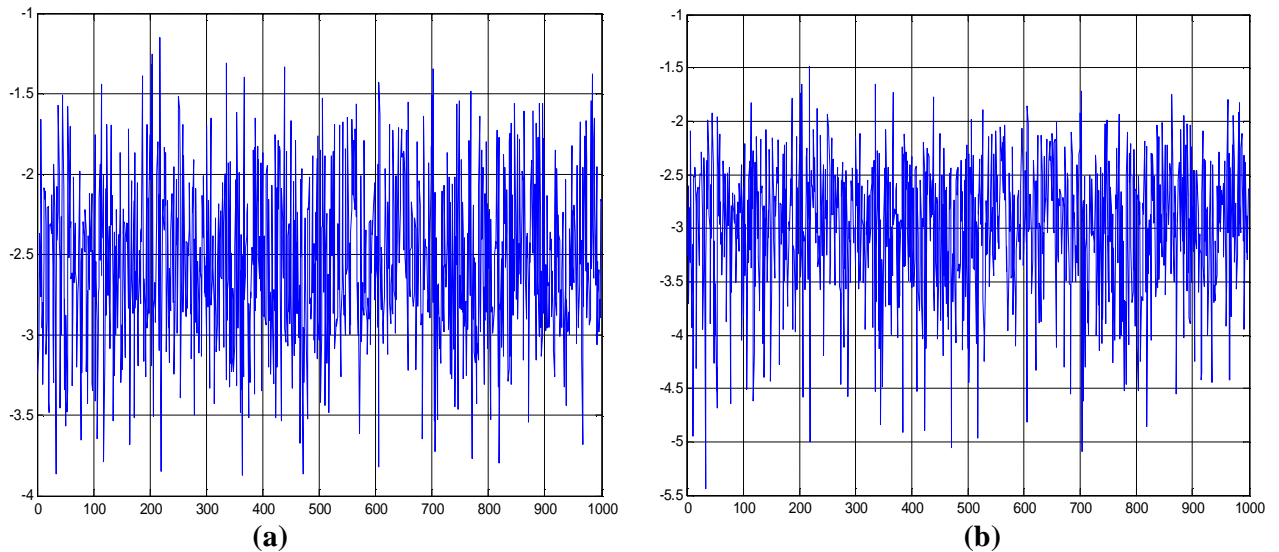


Figure 4-6 la variation de la probabilité de fausse alarme en fonction du nombre d'essai pour  $N=50$ ,  $X_c=9$ ,  $L=2$ . L'estimateur neuronal, (b) l'estimateur à base de la méthode des moments.

#### 4-4- L'influence des estimées sur les performances de la détection dans le Clutter Ki-2 Non Centrale Gamma :

Dans la section précédente, nous avons vu comment les estimateurs étudiés influent sur la régulation de la Pfa. Dans cette section, nous allons étudier les effets qu'aura l'estimateur neuronal sur les performances de la détection, tout en faisant une comparaison avec ceux dus à l'estimateur à base de la méthode des moments.

Comme déjà évoqué dans le premier chapitre, la détection est la déclaration de la présence d'une cible quand elle l'est réellement. La probabilité de détection est donnée par l'expression :

$$P_D = \int_T^{\infty} f_{y/H_1}(y/H_1) dy \quad (4-21)$$

où  $f_{y/H_1}(y/H_1)$  est la densité de probabilité conditionnelle sachant  $H_1$ , et  $T$  est le seuil de adaptatif de détection. Comme on ne peut obtenir une forme analytique compacte et

simple à manipuler de la probabilité de détection, on a procédé au calcul de la probabilité de détection par simulation. Pour ce calcul il est nécessaire de connaître au préalable le rapport signal-sur-Clutter (signal to clutter ratio), qui est tout simplement comme suit :

$$SCR_{dB} = 10 \log(P_S/P_F) \quad (4-22)$$

Ce qui nous amène à :

$$P_S = P_F 10^{SCR/10} \quad (4-23)$$

$P_F$  la puissance du Clutter.

$P_S$  la puissance du signal (Cible)

Comme nous nous basons sur le modèle de Swerling I, la cible suivra une loi Rayleigh définie par la forme :

$$P(q) = \frac{q}{S^2} \exp\left(-\frac{q^2}{2S^2}\right) \quad (4-24)$$

La puissance de la cible est en fonction du rapport SCR, que nous pouvons varier, et aussi en fonction de la puissance du clutter qu'il faudra calculer. Dans le chapitre 1, il a été évoqué que le signal du clutter suit une distribution gaussienne à la sortie de chaque cellule. Cette gaussienne a une variance  $S^2$ , est une moyenne qui suit une loi de distribution gaussienne à moyenne nulle et une variance  $S_M^2$  de façon que la somme des carrés de toutes les moyennes suit une distribution Gamma, qui est une condition. Donc on peut représenter le signal du clutter par une variable aléatoire [5], gaussienne de moyenne  $M$  et de variance  $S^2$ , où  $M$  est elle-même une variable aléatoire gaussienne de moyenne nulle et de variance  $S_M^2$ . On aura la formule de la puissance du clutter suivante :

$$P_F = E[X^2/M] = E[E[x^2/M]] \quad (4-24)$$

$$\begin{aligned} \text{Où } E[x^2/M] &= \text{Var}(x) + E^2[x] \\ &= s^2 + M^2 \end{aligned} \quad (4-25)$$

En remplaçant dans (4-24) :

$$P_F = s^2 + E[M^2] \quad (4-26)$$

Sachant que :

$$E[M^2] = s_M^2 + \underbrace{E^2[M]}_{=0} \quad (4-27)$$

Par conséquent, la puissance du Clutter ne sera qu'en fonction des variances du signal du Clutter et de la variance de la moyenne du Clutter. Elle est donnée comme suit :

$$P_F = s^2 + s_M^2 \quad (4-28)$$

Comme la variance du signal du Clutter  $s^2$  est constante ( $s^2 = 1/2$ ), il nous reste à déterminer la valeur de la variance de la moyenne du Clutter  $s_M^2$ , qui sera tirée à partir du modèle mathématique de la distribution Gamma, pour les deux cas sur lesquelles est basée notre étude, c-à-d,  $L=1, L=2$

#### **L=Nc :**

On sait que la somme de  $2Nc$  signale gaussien de moyennes  $m_k$  et de variance  $s^2$ , suit une distribution Ki-2 Non Centrale de densité de probabilité :

$$P_{NCS}(y/2Nc, c) = (y/c)^{(Nc-1)/2} \exp(-y-c) I_{Nc-1}(2\sqrt{yc}) \quad (4-29)$$

$I$  est le paramètre de non centralité donné par  $I = \sum_{K=1}^{Nc} m_K$ .

Cette dernière suit une distribution Gamma dont l'expression est :

$$P(I) = \frac{(IL/Xc)^{L-1}}{\Gamma(L)2S^2} \cdot \exp(-IL/2S^2 Xc) \quad (4-30)$$

Nous savons que la densité de la distribution Gamma standard (avec un paramètre d'échelle =1) est :

$$P(q) = \frac{q^{L-1}}{2S^2 \Gamma(L)} \exp(-q/2S^2) \quad (4-31)$$

Donc, par analogie entre les équations (4-30) et (4-31), on obtient directement la formule de la variance des moyennes des gaussiennes :

$$S_M^2 = S^2 \frac{Xc}{L} \quad (4-32)$$

En remplaçant la formule (4-32) dans l'équation (4-28) :

$$P_F = S^2 \left(1 + \frac{Xc}{L}\right) \quad (4-33)$$

Par conséquent on aura comme puissance du signal de la cible ce qui suit :

$$P_S = S^2 \left(1 + \frac{Xc}{L}\right) 10^{SCR/10} \quad (4-34)$$

#### ✚ L=1:

Dans ce cas les moyennes des cellules contenant le signal du Clutter ont toutes la même amplitude, alors le paramètre de non centralité devient:

$$I = \sum_K^{Nc} m_k^2, m_k = m, k = 1, 2, \dots, Nc$$

$$I = Nc m^2$$

$$\text{On obtient : } m^2 = \frac{I}{Nc} \quad (4-35)$$



De la même façon que pour le premier cas, la variance  $S_M^2$  aura l'expression suivante :

$$S_M^2 = S^2 \frac{Xc}{L.Nc} \quad (4-36)$$

En remplaçant l'équation (4-36) dans (4-28) :

$$P_F = S^2 \left( 1 + \frac{Xc}{L.Nc} \right) \quad (4-37)$$

Alors la puissance du signal de la cibles aura la forme suivante:

$$P_S = S^2 \left( 1 + \frac{Xc}{L.Nc} \right) 10^{SCR/10} \quad (4-38)$$

#### 4-4-1- Résultats et discussion :

Les Figures suivantes montrent l'influence de  $Xc$  et  $L$  sur la probabilité de détection pour  $Nc=2$  et pour différentes valeurs de longueurs d'échantillons  $N=50, 100, 200, 500, 1000, 2000$ . L'allure générale des courbes démontre clairement l'amélioration qu'apporte l'estimateur neuronal à la probabilité de détection par rapport à celui utilisant la méthode des moments. On peut aussi noter, que pour toutes les valeurs de  $Xc$  et  $L$ , les probabilités de détection obtenues à l'aide des valeurs estimées par l'estimateur neuronal sont soit meilleures ou égales a celles obtenues à l'aide de l'estimation par la méthode des moments. On remarque que la probabilité de détection obtenue par les réseaux de neurones est meilleure que celle de la méthode classique, et ça pour de petites valeurs de  $Xc$  comme on peut le voir sur toutes les Figures 4-7-a jusqu'à 4-12-a. Cela peut s'expliquer par le fait que les estimations du paramètre de fluctuation  $L$  par la méthode des moments devient meilleure au fur et à mesure que  $Xc$  augmente. Par conséquent, une augmentation de la probabilité de détection, alors que les performances obtenues par le réseau de neurone ne change pas d'une façon notable.

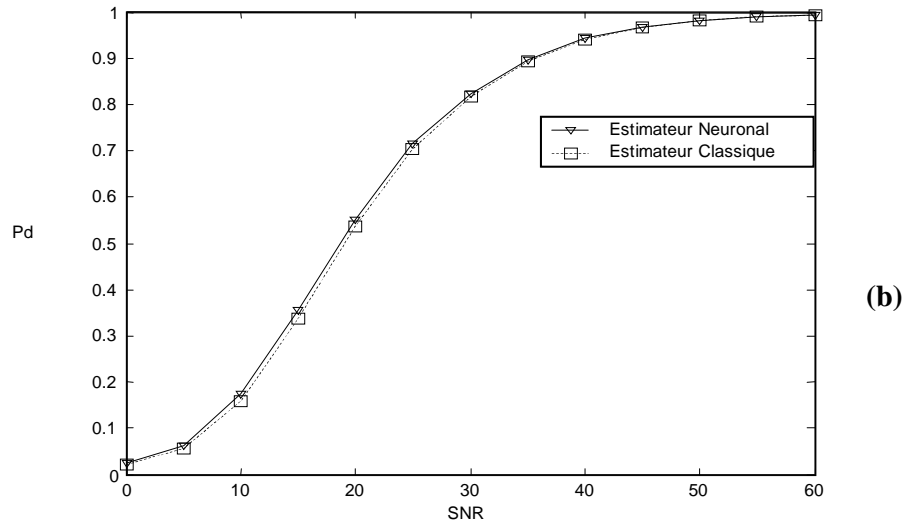
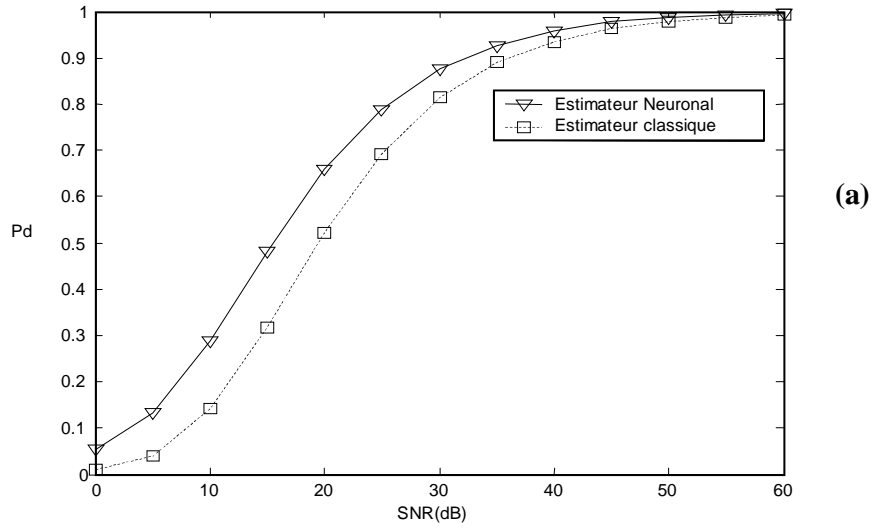


Figure 4-7 la détection de probabilité en fonction du SCR pour  $N=2000$  :  
 (a)  $X_c=1, L=1$ .  
 (b)  $X_c=5, L=1$ .

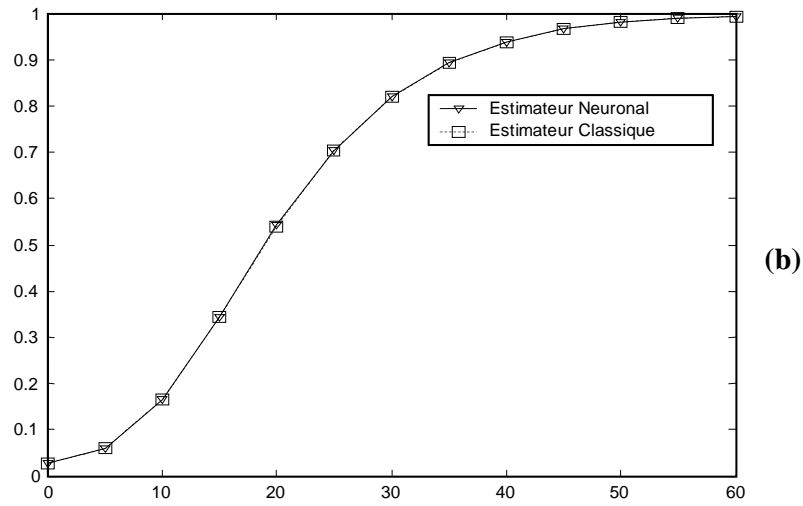
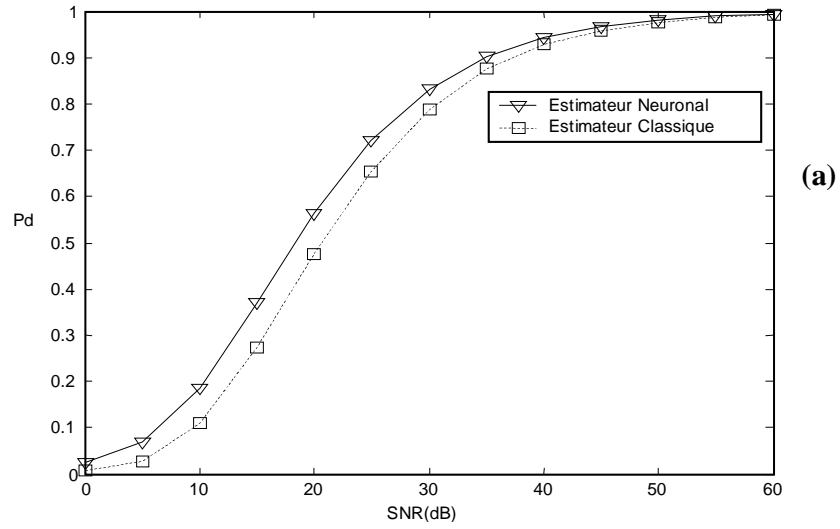


Figure 4-8 la détection de probabilité en fonction du SCR pour  $N=1000$  :  
 (a)  $X_c=1, L=2$ .  
 (b)  $X_c=12, L=2$ .

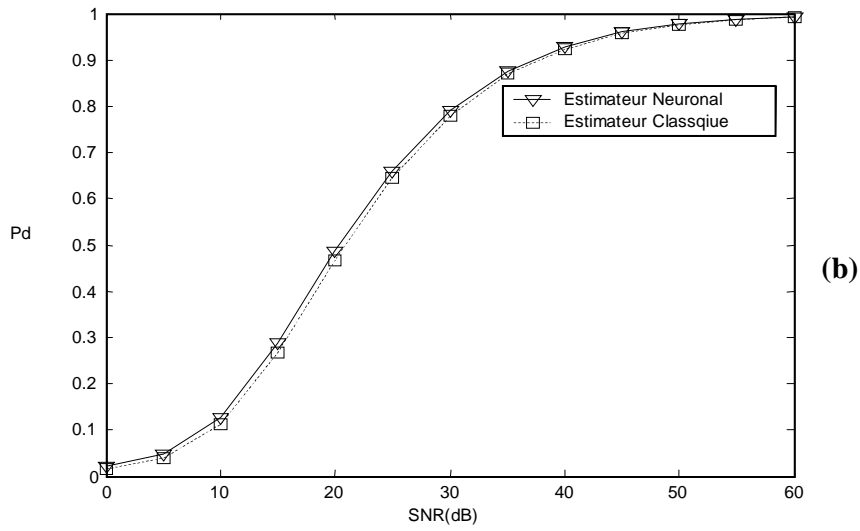
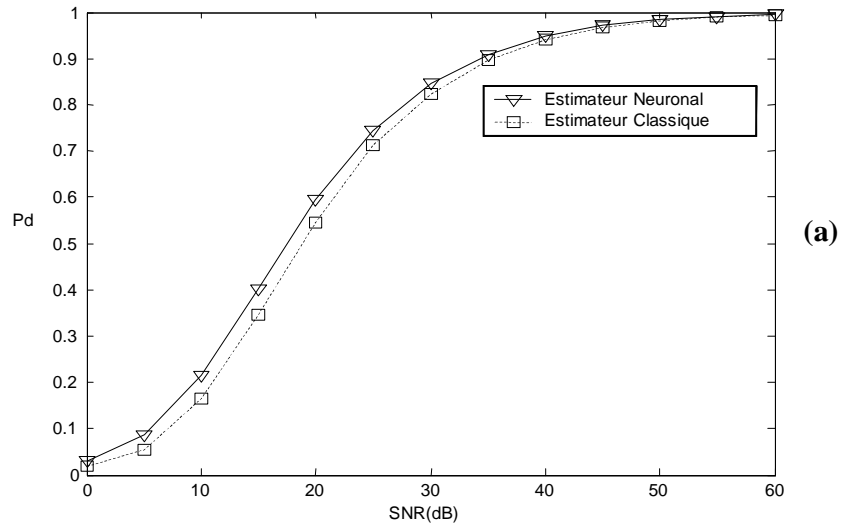


Figure 4-9 la détection de probabilité en fonction du SCR pour  $N=500$  :  
 (a)  $X_c=2, L=1$ .  
 (b)  $X_c=7, L=2$ .

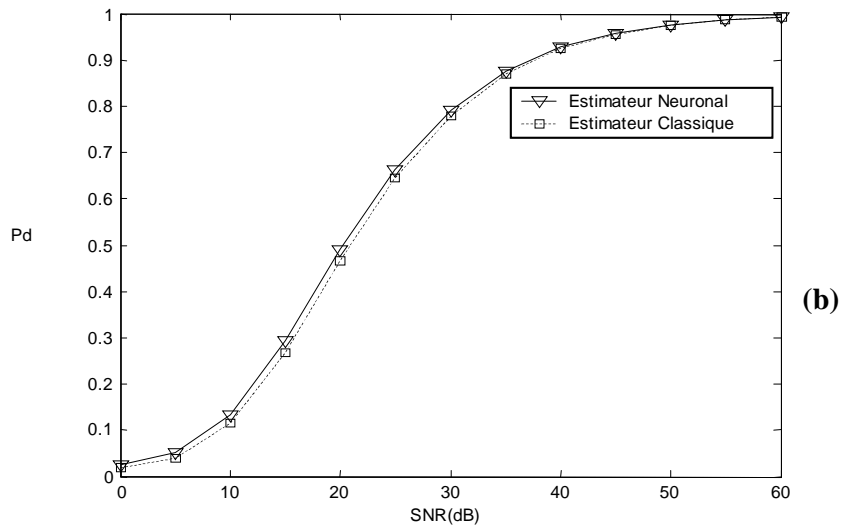
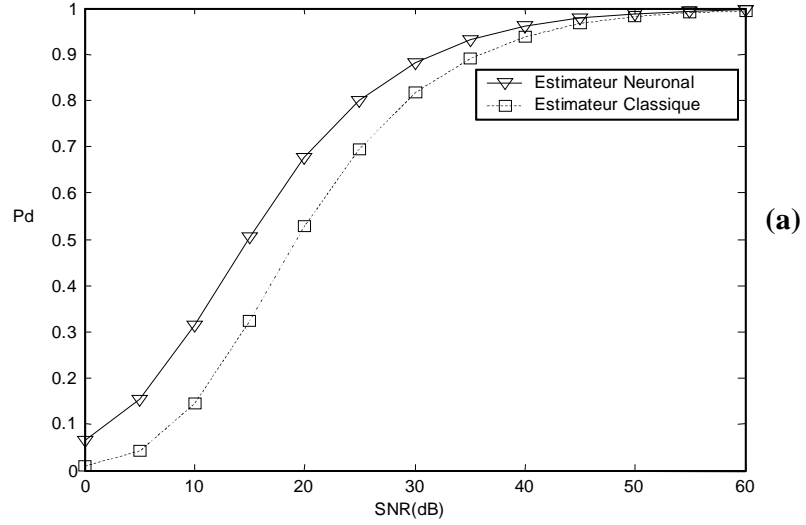


Figure 4-10 la détection de probabilité en fonction du SCR pour  $N=200$  :  
 (a)  $X_c=1, L=1$ .  
 (b)  $X_c=9, L=2$ .

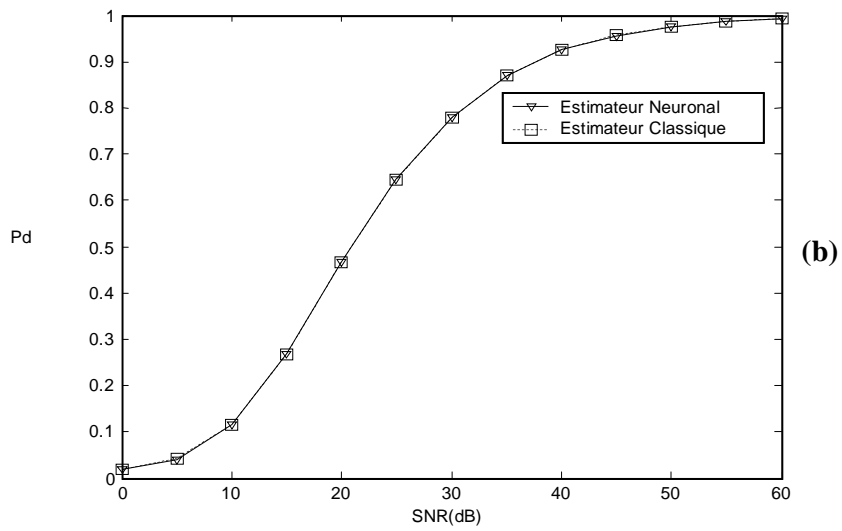
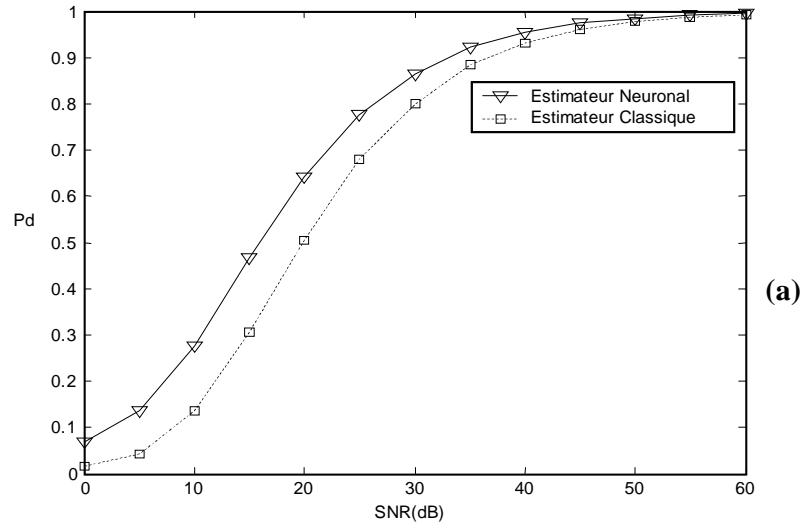


Figure 4-11 la détection de probabilité en fonction du SCR pour  $N=100$  :

- (a)  $X_c=2, L=2$ .
- (b)  $X_c=12, L=1$ .

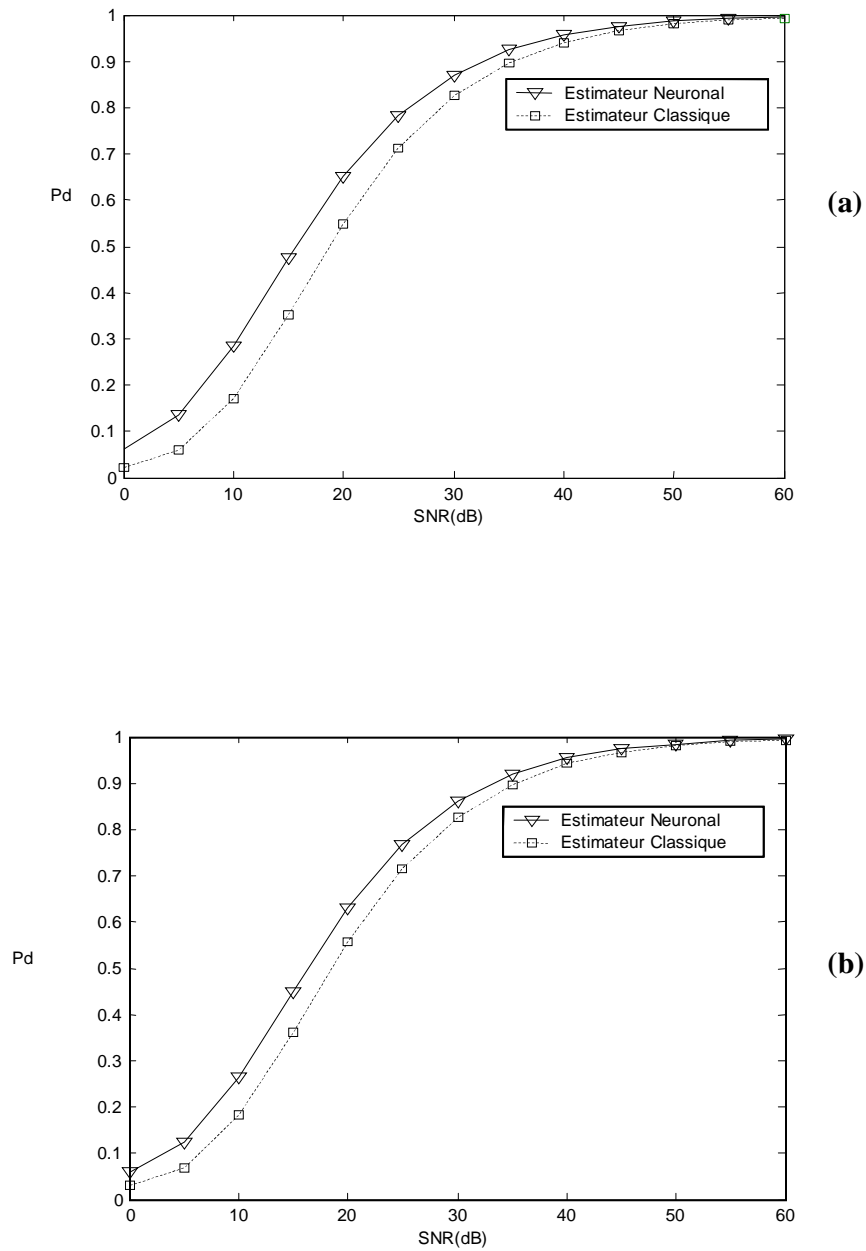


Figure (4-12) : la détection de probabilité en fonction du SCR  
pour  $N=50$  :  
(a)  $X_c=3$ ,  $L=1$ .  
(b)  $X_c=9$ ,  $L=2$ .

On peut toutefois remarquer à partir des Figures 4-7-b jusqu'à 4-12-b, que les valeurs de  $X_c$  pour lesquelles les deux probabilités de détection commencent à se rapprocher augmentent avec la diminution de la longueur d'échantillons et cela est dû au fait que les performances de l'estimateur neuronal ne se dégradent pas aussi rapidement que celles de l'estimateur classique, sachant que celui-ci connaît une grande dégradation pour de petites longueurs d'observations.

#### 4-5- Conclusion :

Dans ce chapitre, nous avons introduit les réseaux de neurones dans l'estimation des paramètres de la distribution Ki-2 Non Centrale Gamma. Les résultats obtenus à la suite de la simulation démontrent clairement que le biais et la variance des estimées, bien que, seulement le paramètre de fluctuation  $L$  présente une réelle amélioration en comparant avec l'estimateur à base des moments, le biais et la variance de la puissance moyenne  $X_c$  sont toutefois considérés comme étant acceptable, et voyant même que leurs différences avec ceux de l'estimateur à base des moments qui sont meilleurs n'empêche pas d'avoir de bon résultats. Ces résultats peuvent se faire appréciés lorsque nous avons analysé les effets que peuvent avoir les estimations sur la régulation de la fausse alarme et les performances de détection.



# *Conclusion Générale*

## Conclusion Générale

Le but de ce travail est de répondre à une problématique qui est :

- ✚ Quels avantages apporteraient les réseaux de neurones dans les systèmes de détection radar ?
- ✚ Quels influences et effets auront-ils sur toute la chaîne de détection ?

On s'est proposé de prendre en considération l'étude de l'estimation des paramètres de la distribution Ki-2 Non Centrale Gamma, qui est un des modèles généralisés qui caractérisent le clutter d'un radar à haute résolution de nature multiparamétrique ( $X_c, L, N_c$ ). Puisqu'il est difficile d'utiliser les méthodes classiques à cause de leur inadaptabilité, alors il y avait une imminente nécessité de trouver d'autres approches qui ne sont pas affectés par la complexité du modèle du clutter.

Une méthode classique et qui est la méthode des moments a été utilisée pour l'estimation des paramètres de cette distribution. Malgré sa simplicité, elle nécessite un grand nombre d'échantillons pour donner de bons résultats. Une autre méthode à base de réseaux de neurones a été introduite. En particulier, on a utilisé l'architecture MLP qui est connue pour sa simplicité d'implémentation et pour ces propriétés d'approximateur universelle, c-à-d avec un bon choix des paramètres internes de ce réseau on peut obtenir l'approximation de n'importe quelle fonction complexe.

La mise en avant des avantages de cette méthode fut par l'intermédiaire d'une étude comparative avec la méthode des moments aussi bien au niveau de l'estimation, en observant les biais et variances des deux estimateurs, qu'au niveau de la régulation de la probabilité de fausse alarme, en étudiant l'influence des valeurs estimées sur la fausse alarme et ainsi avoir une idée sur ces fluctuations par rapport à une valeur nominale. En dernier lieu, on a évalué les performances de la détection d'une cible de type Swerling I et les résultats ont montré les bonnes performances que nous pouvions obtenir à l'aide des valeurs estimées résultantes du réseaux de neurones par rapport à celles obtenues à partir de la méthode des moments.

Dans le cadre de ce travail, nous avons pris en considération une seule architecture de réseaux de neurones et qui est l'architecture MLP (Multi Layers Perceptron), qui jusque là a apporté une amélioration considérable, mais que nous estimons insuffisante dû au fait qu'elle connaît une dégradation lorsqu'on travaille avec un petit nombre d'échantillons. Donc, nous jugeons nécessaire d'étudier l'apport que pourraient amener d'autres architectures avec différents types d'algorithmes d'apprentissage.

*Références*

- [1] Alan V. Oppenheim, Applications of Digital Signal Processing., Editor. Massachusetts institute of technology Cambridge, Mass, 02139. 1978.
- [2] M. Barkat. Signal Detection and Estimation. State University of New York, 1991 Artech House.
- [3] CHEIKH Khaireddine. La détection CFAR utilisant les réseaux de neurones artificiels dans un clutter K-distribution .Thèse de Magister. Université de Mentouri, Constantine. Département d'électronique, Janvier 2004.
- [4] Drumheller, D.M , General expressions for Ricean density and targets , IEEE Transaction on Aerospace and Electronic systems, Vol.29. N°2, Pg580-588, 1993
- [5] Shnidman, D.A, Generalized radar clutter model. IEEE Transactions on Aerospace and Electronic Systems, Vol.35, pp857-865, July 1999.
- [6] Jackeman, E, and Pursey, P.N. A model for non Rayleigh sea echo, IEEE Transactions on Antennas and propagation, Vol.24, N°6, PP.806-814, 1976.
- [7] Anntassopouloss, V, Lampropoulos, G. A, Drosopoulos, A. and Rey, M, High resolution radar clutter statistics, IEEE Transactions on Aerospace and Electronic Systems, Vol.35, N°1, PP 43-60, January 1999.
- [8] Abramowitz, M, Stegun, I. Handbook of mathematical functions, New York 1972. pg 558, formule 15.2.14.
- [9] Jean-François Jodouin. Les réseaux neuromimétiques, modèles et applications. © Hermès, Paris, 1994.
- [10] Eric DAVALO, Patrick NAIM, Des réseaux de neurones. Deuxième édition, deuxième tirage 1993.

- [11] R.P Lipmann, An introduction to computing with neural network, IEEE ASSP Magazine, April 1987.
- [12] Anil K, Jain. Jianchang Mao. K, Mohiuddin. Artificial Neural Network. IEEE computer Special Issue On Neural Computing, 1996.
- [13] Vaseghi, Saeed V .Advanced Digital Signal Processing and Noise Reduction, .Published By: John Wiley & Sons, Ltd.Edition: 2<sup>nd</sup>.
- [14] D. Robert Iskander, Abdelhak M. Zoubir. Estimation of the Parameters of the K-Distribution Using Higher Order and Fractional Moments. IEEE Transactions on Aerospace and Electronic Systems, Vol.35, N°4, October 1999.
- [15] Mark P.Wachowiak, Renata Smolikova, Jacek M Zurada, Adel S. Almaghraby. Estimation of K Distribution Parameters Using Neural Networks, IEEE Transactions On Biomedical Engineering, Vol. 49, no. 6, June 2002 617.
- [16] M.Jahangir, D. Blacknell ,R.G.White. Accurate approximation to the optimum parameter estimate for K-distributed fouillis, , IEE Proc-Radar, Sonar Navig., Vol. 143, No. 6, December 1996.
- [17] D. Robert Iskander, Abdelhak M Zoubir, et Boualam Bousahsah .A Method for estimating the parameters of K distribution. IEEE Transaction on Signal Processing, Vol 47, No. 4, April 1999.
- [18] Ian.r, Joughin, Donald B, Percival, Dale, P, Winberner. Maximum likelihood estimation for k-distribution SAR data, IEEE Transactions Geosci Remote Sensing, Vol 29, pp.489-499, 1993.
- [19] Abramowitz, M, Stegun, I. Handbook of mathematical functions, New york 1972, pg 258, formule ,eq 6.3.1.

- [20] R. Ravid N. Levanon, Maximum likelihood cfar for weibull background, IEE Proceeding ,Pt, F, Communication, Radar and signal processing, Vol.139,N°3,pp.256-264,1992.
- [21] Maurizio guida, M Lango and M Lops .Biparametric Cfar Procedure for Log-Normal clutter. IEEE Transaction on Aerospace and Electronic Systems, Vol.29, N°3, pp 798-808,1993
- [22] BENAZZOUZ Nadjat, Detection Et Estimation Dans Un Clutter De Distribution Ki-2 Non Centrale Gamma, These de magister ,Université Mentouri, Constantine, Département d'électronique, 2002.

# *Références*

## *Abstract*

The goal of a radar system is the automatic control of a target in space. This is a delicate task which must be done under various conditions and environments, which makes it difficult and complex. The great diversity of the environments in which a radar operates is the principal source of the various noises and undesirable echoes, that are called Clutter which come to be added to the signal of the target and thus to scramble it .

Our work consists in introducing the neural networks into the estimation of the parameters of the distribution *Ki-2 Non Central Gamma*, focusing on MLP ( Multi Layers perceptron) architecture , which is an architecture as simple as effective especially in the applications of functions approximations. We have implemented an MLP neural network with two hidden layers, an output layer which presents two outputs and an input layer with two inputs.

According to the results, one can conclude that the artificial neural networks can considerably improve the radar detection



## ملخص

إن الهدف الرئيسي من نظام الرادار هو الكشف الآلي عن هدف ما في الفضاء. هذه العملية هي عملية حساسة لكونها تحدث تحت تأثير ظروف و بيئات مختلفة, الشيء الذي يجعلها عملية صعبة و معقدة. التنوع الكبير للبيئات التي يعمل فيها الرادار هي المنبع الأساسي للضوضاء و الصدى أو ما يسمى بالتشويش غير المرغوب فيه, هذا الأخير الذي يضاف إلى إشارة الهدف و يعمل على تشويشها.

العمل الذي قمنا به يتمحور حول إستعمال الشبكات العصبونية في تقدير المقاييس الخاصة بالتوزيع  $Ki - 2 Non Centrale Gamma$ , و بالخاصة شبكة MLP, التي تعد ذات هيكل مبسط و فعال في آن واحد, خاصة في التطبيقات التقريبية للدوال. لقد قمنا بإنشاء شبكة عصبونية من نوع MLP ذات طبقتين مخفيتين, طبقة خروج تحمل مخرجين و طبقة دخول ذات مدخلين.

من خلال النتائج المحصل عليها , نستنتج أن الشبكات العصبونية الإصطناعية بإمكانها إدخال تحسينات معتبرة للكشف في جهاز الرادار .

## *Résumé*

Le but d'un système radar est la détection automatique d'une cible dans l'espace. Ceci est une tâche délicate qui doit se faire sous diverses conditions et environnements, ce qui la rend difficile et complexe. La grande diversité des environnements dans lesquels opère un radar est la source principale des différents bruits et échos indésirables appelés "Clutter" qui viennent s'ajouter au signal de la cible et ainsi le brouiller

Notre travail consiste à introduire les réseaux de neurones dans l'estimation des paramètres de la distribution *Ki-2 Non Centrale Gamma* en se focalisant sur une architecture MLP (Perceptron Multi Couches), qui est une architecture aussi simple qu'efficace surtout dans les applications d'approximations de fonctions. Nous avons mis en œuvre un réseau de neurones MLP à deux couches cachées, une couche de sortie qui présente deux sorties qui sont en fait les paramètres à estimer, une couche d'entrée avec deux entrées.

D'après les résultats trouvés, on peut conclure que les réseaux de neurones artificiels peuvent améliorer considérablement la détection radar.