

**Université SAAD DAHLEB de Blida 1**

**Institut d'Aéronautique et des études Spatiales**



Mémoire de fin d'études Master2



Pour l'obtention du diplôme de MASTER

**Option : CNS/ATM**

Thème :

**Estimation de la fonction de densité de probabilité  
et le différentes Techniques de détection Radar**

Par : MEDJIR Aouatef

Promoteur : Dr. RAHMOUNI Mohammed.

Blida: 2020/2021

## ملخص:

منذ تطور الرادار، استخدم على نطاق واسع وأحرز تقدماً كبيراً. وتستخدم هذه الأجهزة أساساً في الكشف والقياس عن بعد. وهذا العمل يبرره هذا الواقع وهو جزء من البحث عن أفضل كاشف لهذه الحالة. الهدف هو اقتراح كاشف يجمع بين معدل الإنذار الكاذب الثابت (CFAR) تكنولوجيا الكشف مع طريقة التقدير غير البارامترية لوظيفة كثافة الاحتمالات الكومومية كوسيلة لتمييزها بشكل ثابت، وقد أدخلنا بعض مفاهيم مبدأ تشغيل الرادار وأدخلنا أيضاً أنواعاً مختلفة من الضوضاء وبعض قوانين التوزيع التي يمكن استخدامها لوصف نظم الكشف التلقائي، ولا سيما تلك المستخدمة في عملنا. تم إدخال طريقة غير بارامترية لتقدير دالة كثافة الاحتمالات، نقوم بتقديمها عن طريق إدخال معايير الخطأ بين الوظائف المستخدمة عادة في الكتابات، ثم نقوم بإدخال تقنيات التقدير من الأبسط إلى الأكثر تقدماً. بعد ذلك، نطور النتائج المتعلقة بتقنيات التقدير المختلفة لدالة الكثافة الاحتمالية. وقد قدمنا بإيجاز نظرية كشف الرادار، ونصائح بشأن تقنيات الكشف التقليدية والتكيف، وعلقنا على نتائج المحاكاة لأجهزة الكشف CA وGO وSO وOS-CFAR في مختلف أنواع الحالات.

الكلمات الرئيسية: الرادار، الكشف، تقدير دالة الكثافة، CFAR.

## Résumé:

Depuis le développement du radar, il a été largement utilisé et a fait de grands progrès. Ces appareils sont principalement utilisés pour la détection et la télémétrie. Ce travail se justifie par cette réalité et s'inscrit dans la recherche du meilleur détecteur pour cette situation. L'objectif est de proposer un détecteur qui combine la technologie de détection du taux de fausses alarmes constant (CFAR) avec la méthode d'estimation non paramétrique de la fonction de densité de probabilité de fouillis comme moyen de la caractériser statiquement, et permet enfin l'application d'un traitement CFAR approprié et Le choix optimal de la valeur du facteur d'échelle T. Nous avons introduit quelques concepts du principe de fonctionnement du radar et également introduit différents types de bruit et quelques lois de distribution qui peuvent être utilisées pour caractériser les systèmes de détection automatique, en particulier ceux utilisés dans notre travail. La méthode non paramétrique d'estimation de la fonction de densité de probabilité a été introduite, nous l'introduisons en introduisant des normes d'erreur entre les fonctions couramment utilisées dans la littérature, puis nous introduisons des techniques d'estimation des plus simples aux plus avancées. Ensuite, nous développons les résultats liés aux différentes techniques d'estimation de la fonction de densité de probabilité. Nous avons brièvement présenté la théorie de la détection radar, des conseils sur les techniques de détection classiques et adaptatives, et commenté les résultats de simulation des détecteurs CA, GO, SO et OS-CFAR dans différents types de situations.

Mots clés : radar, détection, estimation de la fonction de densité, CFAR.

## **Abstract:**

Since the development of the radar, it has been widely used and has made great progress. These devices are mainly used for detection and telemetry. This work is justified by this reality and is part of the search for the best detector for this situation. The objective is to propose a detector that combines the Constant False Alarm Rate (CFAR) detection technology with the non-parametric estimation method of the clutter probability density function as a means of statically characterizing it, and finally allows the application of a suitable CFAR treatment and the optimal choice of the value of the scale factor  $T$ . We have introduced some concepts of the operating principle of the radar and also introduced different types of noise and some distribution laws that can be used to characterize automatic detection systems, especially those used in our work. The non-parametric method of estimating the probability density function has been introduced, we introduce it by introducing error standards between the functions commonly used in the literature, then we introduce estimation techniques from the simplest to the most advanced. Next, we develop the results related to the different estimation techniques of the probability density function. We briefly presented radar detection theory, advice on conventional and adaptive detection techniques, and commented on the simulation results of CA, GO, SO and OS-CFAR detectors in different types of situations.

Keywords: radar, detection, density function estimation, CFAR.

## **Dédicace:**

Je dédie ce travail :

À mes chers parents.

À mes sœurs.

À mes frères et toute ma famille.

À toutes les personnes qui me sont chères.

## Remerciements:

Je tiens tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui m'a donné la force et la patience d'accomplir ce Modeste travail.

En second lieu, je tiens à remercier mon promoteur **Dr. RAHMOUNI Mohamed**, docteur à l'université de **SAAD Dahlab Blida1**, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à faire ce travail.

Mes vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à ma recherche en acceptant d'examiner notre travail Et de l'enrichir par leurs propositions.

Je tiens aussi à exprimer ma reconnaissance aux enseignants du département d'aéronautique et des études spatiales qui ont contribué à ma formation et mon apprentissage tout au long de ces cinq dernières années et sans eux qui je ne serais là où j'en suis aujourd'hui.

Mes derniers remerciements, vont à tous ceux qui ont contribué de près ou de loin pour l'aboutissement de ce travail.

## Liste des acronymes :

<b>AMISE</b>	Asymptotic Mean Integrated Squared Error.
<b>ASH</b>	Averaged Shifted Histograms.
<b>CST</b>	Cellule Sous Test.
<b>CA-CFAR</b>	Cell Averaging-Constant False Alarm Rate.
<b>CFAR</b>	Constant False Alarm Rate.
<b>GO-CFAR</b>	Greatest Of-Constant False Alarm Rate.
<b>HRR</b>	High Resolution Radar.
<b>IID</b>	Independant Identically Distributed.
<b>INR</b>	Interference to Noise Ratio.
<b>ISE</b>	Integrated Squared Error.
<b>MSE</b>	Mean Squared Error.
<b>MISE</b>	Mean Integrated Squared Error.
<b>OS-CFAR</b>	Ordered Statistics- Constant False Alarm Rate.
<b>PDF</b>	Fonction de Densité de Probabilité.
<b>RADAR</b>	RADio Detection and Ranging.
<b>RCS</b>	Radar Cross Section.
<b>SER</b>	Surface Equivalent Radar.
<b>SCR</b>	Signal to Clutter Ratio.
<b>SIR</b>	Signal to Interferente Ratio.
<b>SO-CFAR</b>	Smallest Of -Constant False Alarm Rate.
<b>SNR</b>	Signal to Noise Ratio.

## Liste des symboles :

$F_x$	Fonction de densité de probabilité.
$F_h$	Estimée de la fonction de densité de probabilité.
$F$	Fonction de répartition.
$E$	Espérance.
$\mu$	Moyenne.
$\sigma^2$	Variance.
$cv$	Coefficient de variation
$\gamma_1$	Premier facteur de forme Skewness.
$\gamma_2$	Deuxième facteur de forme Kurtosis.
$d_{kl}$	Distance de Kullback-Leibler.
$d_H$	Distance de Hellinger.
$p_{fa}$	Probabilité de fausse alarme.
$p_{fa\ nom}$	Probabilité de fausse alarme nominale.
$T$	Facteur d'échelle.
$N$	Nombre de cellules.
$p_d$	Probabilité de détection.

# Table des matières :

Introduction générale :.....	1
Chapitre 1 : généralités sur les systèmes Radars :.....	4
1.1 Définition d'un radar :.....	4
1.2 Classification des radars :.....	5
1.3 Différents types de bruit et signaux parasites :.....	5
1.3.1 Bruit interne (thermique) :.....	5
1.3.2 Écho parasite :.....	6
1.3.3 Fouillis ou Clutter :.....	6
1.3.4 Bord de Clutter :.....	7
1.3.5 Cibles interférentes :.....	8
1.4 Fluctuation des cibles :.....	9
1.4.1 Fluctuation des cibles :.....	9
1.4.2 Fluctuation de la SER :.....	10
1.4.3 Modèles de fluctuation des cibles :.....	10
1.5 Modèles de distributions de Clutter :.....	12
1.5.1 Modèle Gaussien :.....	12
1.5.2 Modèle de Log-normal :.....	13
1.5.3 Modèle de Weibull :.....	15
1.5.4 Modèle de Rayleigh :.....	19
1.5.5 Modèle de K-distribué :.....	20
Chapitre 2 : estimation de la fonction de densité de probabilité :.....	23
2.1 Introduction à l'estimation de la fonction de densité de probabilité :.....	23
2.1.1 Estimation de la fonction de densité de probabilité :.....	23
2.1.2 Propriétés de la fonction de densité de probabilité :.....	24
2.1.3 Applications de l'estimation de fonction de densité de probabilité :.....	24
2.2 Critère d'erreur :.....	24
2.2.1 Erreur quadratique intégrée ISE (Integrated Squar Error) :.....	25
2.2.2 Erreur quadratique moyenne MSE (Mean Squar Error) :.....	25
2.2.3 Erreur quadratique moyenne intégrée MISE (Mean Integrated Square Error) :.....	25
2.2.4 Distance de Kullback- Leibler :.....	26

2.2.5	Distance de Hellinger : .....	26
2.3	Estimation par histogramme : .....	26
2.3.1	Histogramme : .....	27
2.3.2	Choix de paramètre de lissage : .....	29
2.4	Polygone de fréquence : .....	29
2.4.1	Choix du paramètre de lissage : .....	30
2.5	Histogrammes décalés ASH (Averaged Shifted Histograms): .....	32
2.6	Estimation non paramétrique par la méthode du noyau : .....	33
2.6.1	Introduction : .....	33
2.6.2	Définition du noyau de Parzen-Rosenblatt : .....	34
2.6.3	Propriétés de l'estimateur à noyau : .....	34
2.6.4	La convergence de l'estimateur à noyau : .....	35
2.6.5	La consistance d'un estimateur : .....	36
2.6.6	Choix du noyau (exemples de noyaux) : .....	37
2.6.7	Choix du paramètre de lissage : .....	42
2.7	Estimateur à noyau adaptatif : .....	43
2.7.1	Définition d'un estimateur à noyau adaptatif : .....	44
2.8	Simulation et comparaison des résultats : .....	45
2.8.1	Introduction : .....	45
2.8.2	Le choix de langage de programmation : .....	45
2.8.3	Estimateur à noyau, que choisir le noyau ou sa largeur : .....	46
2.8.4	Exemples de densités estimées par les différents estimateurs : .....	49
2.8.5	Comparaison des performances des estimateurs : .....	50
Chapitre 3 : Technique de détection Radar : .....		56
3.1	Définition de détection : .....	56
3.2	Rappels sur la théorie de la détection radar : .....	56
3.2.1	Critères de décision : .....	57
3.3	Détection à seuil fixe : .....	59
3.4	Détection à seuil adaptatif CFAR : .....	60
3.4.1	Analyse du détecteur CA-CFAR : .....	61
3.4.2	Analyse du détecteur GO-CFAR : .....	63
3.4.3	Analyse du détecteur SO-CFAR : .....	65
3.5	Comparaison des performances des détecteurs CA, GO et SO-CFAR : .....	66
3.5.1	Cas d'absence de cibles interférentes : .....	66

3.5.2	Cas d'un bord de clutter :	67
3.5.3	Cas de présence de cibles interférentes :	68
3.6	Détection radar en présence d'interférences :	70
3.6.1	Détecteur OS-CFAR :	70
3.7	Proposition d'une architecture d'exploitation des techniques d'estimations de la PDF à la détection CFAR :	73
3.7.1	Explication de l'approche proposée :	74
	Conclusion générale :	76
	Annexes :	77
	Bibliographie :	82

# Liste des figures :

<b>Figure 1:</b> Principe de détection radar. ....	4
<b>Figure 2:</b> Cellule sous test dans le bruit thermique en présence d'un bord de clutter.....	7
<b>Figure 3:</b> Cellule sous test dans le bruit thermique et le clutter. ....	8
<b>Figure 4:</b> Cellule sous test dans le bruit thermique en présence des cibles interférentes. ....	9
<b>Figure 5:</b> Densité de probabilité normale standard. ....	13
<b>Figure 6:</b> Densité de probabilité log-normale.....	14
<b>Figure 7:</b> Densité de probabilité Weibull. ....	16
<b>Figure 8:</b> Evolution du coefficient de variation (cv) de la loi Weibull. ....	17
<b>Figure 9:</b> Variation du premier facteur de forme (Skewness) de la loi Weibull. ....	18
<b>Figure 10:</b> Variation du second facteur de forme (Kurtosis) de la loi Weibull. ....	19
<b>Figure 11:</b> Densité de probabilité de Rayleigh. ....	20
<b>Figure 12:</b> Densité de probabilité K-distribué. ....	21
<b>Figure 13:</b> Histogramme de deux valeurs de h ( $h_1=0.05$ et $h_2=0.025$ ).....	28
<b>Figure 14:</b> Densité de probabilité normale estimée par polygone. ....	30
<b>Figure 15 :</b> Densité de probabilité normale estimée par ASH.....	33
<b>Figure 16:</b> exemple de noyau d'épanechnikov utilisé pour estimer les fonctions de densité de probabilité...	39
<b>Figure 17:</b> exemple de noyau triangulaire utilisé pour estimer les fonctions de densité de probabilité.....	40
<b>Figure 18:</b> exemple de noyau tripoids utilisé pour estimer les fonctions de densité de probabilité. ....	40
<b>Figure 19:</b> exemple de noyau normal utilisé pour estimer les fonctions de densité de probabilité. ....	41
<b>Figure 20:</b> exemple de noyau bipoids utilisé pour estimer les fonctions de densité de probabilité. ....	41
<b>Figure 21:</b> Exemple de noyau silverman utilisé pour estimer les fonctions de densité de probabilité.....	42
<b>Figure 22:</b> Densité de probabilité estimée avec différents noyaux.....	46
<b>Figure 23:</b> Erreur quadratique moyenne pour les différents noyaux. ....	47
<b>Figure 24:</b> Densité de probabilité estimée avec différentes largeurs de bandes. ....	48
<b>Figure 25:</b> Erreur Quadratique moyenne pour les différents paramètres de lissage.....	48
<b>Figure 26:</b> Densité de probabilité Normale ( $m = 0, \sigma^2 = 1$ ) stimée avec un noyau à largeur fixe. ....	49
<b>Figure 27:</b> Densité de probabilité Weibull ( $c = 4.5, b = 1$ )estimé avec un noyau à largeur adaptative.....	50
<b>Figure 28:</b> Variations de la MISE en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Normale ( $N(5,1)$ ). ....	51
<b>Figure 29 :</b> Variations de la MISE en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité de Weibull ( $c=5, b=5$ ). ....	52
<b>Figure 30:</b> Variations de la MISE en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Log-normale ( $m=1.2, c=0.3$ ).....	52
<b>Figure 31:</b> Variations de la distance de Hellinger en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Normale ( $N(5,1)$ ). ....	53
<b>Figure 32:</b> variations de la distance de Hellinger en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Weibull ( $c=4.5, b=5$ ).....	54
<b>Figure 33:</b> Variations de la distance de Hellinger en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Log-normale ( $m=1.2, c=0.3$ ).....	54
<b>Figure 34:</b> Principe de détection avec un seuil fixe. ....	60
<b>Figure 35:</b> Schéma fonctionnel du détecteur CA-CFAR.....	62
<b>Figure 36:</b> Schéma fonctionnel du détecteur GO-CFAR. ....	64
<b>Figure 37:</b> Schéma fonctionnel du détecteur SO-CFAR. ....	65

**Figure 38:** Seuils CA, GO et SO-CFAR en absence de cibles interférentes. .... 66

**Figure 39:**  $P_d$  en fonction du SNR pour CA-CFAR, SO-CFAR et GO-CFAR en absence de cibles  
interférentes. .... 67

**Figure 40:** Seuils CA, GO et SO-CFAR en présence d'un bord de clutter. .... 68

**Figure 41:** Seuils CA, GO et SO-CFAR en présence des cibles interférente. .... 69

**Figure 42:** Schéma fonctionnel du détecteur OS-CFAR. .... 70

**Figure 43:** Seuil CA et OS-CFAR en présence des cibles interférentes..... 72

**Figure 44:** Organigramme basée sur l'estimation de la fonction de densité de probabilité. .... 73

## Liste des tableaux :

<b>Tableau 1:</b> Modèle de fluctuation des cibles de Swerling.....	12
<b>Tableau 2:</b> Exemples de modélisation du clutter. ....	22
<b>Tableau 3:</b> Exemples de noyaux utilisés pour estimer la fonction de densité de probabilité.....	39
<b>Tableau 4:</b> Détection et test de deux hypothèses $H_0$ et $H_1$ . ....	60

## Introduction générale :

Un des objectifs importants dans la conception d'un système radar est la détection automatique des signaux issus de cibles volants corrompus par des signaux indésirables (bruit plus fouillis). Ces signaux indésirables perturbent la qualité de la détection et provient généralement de l'électronique du radar (bruit thermique) et de l'environnement (nuage, montagne, etc.) : modélise par une certaine loi appelée fonction de densité de probabilité (PDF).

Dans la plupart des situations, le radar doit situer autour duquel se trouve la cible à détecter: c'est le cas des radars de surveillance qui scrutent l'horizon du sol ou de la mer. Dans une situation pareille, les signaux réfléchis par une cible sont corrompus par des échos parasites appelés Clutter tels que ces provenant des sols, de la végétation, de la mer...etc. Ce Clutter a été considéré, pendant longtemps, dans les radars conventionnels ayant une loi de distribution gaussiennes permettent alors une détection optimale.

Malheureusement dans nombreuses situations, la statistique du Clutter ne s'apparente que par une loi non-gaussienne de types k-distribution, Weibull, log-normal, paréto...etc.

Dans les systèmes de détection radar, la caractérisation statistique du Clutter a suscité un intérêt particulier dû essentiellement aux effets pervers de celui-ci à travers l'évolution des systèmes radar et l'extension de leurs applications. En effet, les scientifiques du domaine du radar ont pendant longtemps considéré que la statistique du Clutter est modélisée par une loi normale (gaussienne), et d'autres chercheurs ont remarqué qu'il y a une influence qui va du masquage des cibles qui passent à travers une région qu'il occupe à la génération de fausses alarmes ces derniers mobilisent inutilement des ressources ou déconcentrent les opérateurs et génèrent des dégradations de performances observées sur les radars conventionnels à savoir ceux conçus pour la détection de cibles noyées dans un Clutter gaussien, et de Clutter éclairé par des radars à haute résolution (HRR) ou à faibles angles d'incidences  $\theta \leq 5^\circ$ . Pour se rapprocher de la réalité, il fallut penser à l'estimation de la loi de distribution des signaux reçus au niveau du récepteur radar pour palier au problème des fausses alarmes. Aussi, la détection classique utilisant un seuil fixe crée des fausses alarmes du fait de la non-stationnarité de la puissance du Clutter. Pour remédier à cette situation, des détecteurs à taux de fausses alarmes constants (CFAR) ont été proposés dans la littérature et l'implémentés dans des radars réels.

Ce travail est justifié par cette réalité et s'inscrit dans le cadre de la recherche des détecteurs optimaux pour pareille situation. La finalité est de proposer un détecteur qui associe les techniques de détection à taux de fausses alarmes constant (CFAR) aux méthodes d'estimation non paramétriques de la fonction de densité de probabilité du Clutter comme moyen de le caractériser statiquement, ce qui permet de terme d'appliquer le traitement CFAR approprié et le choix optimal de la valeur du facteur d'échelle T.

Le mémoire est organisé en trois chapitres comme suit :

Dans le premier chapitre, nous introduisons quelques concepts sur le principe de fonctionnement d'un radar, nous introduisons aussi les différents types de bruits et quelques lois de distributions qui peuvent servir dans la caractérisation des systèmes de détection automatique et particulièrement celles qui sont utilisées dans notre travail.

Le deuxième chapitre, est consacré aux méthodes non paramétriques d'estimation de fonction de densités de probabilité. Nous l'introduisons par la présentation des critères d'erreurs entre fonction utilisées généralement en littérature, ensuite nous présentons les techniques d'estimations des plus simples aux plus évoluées. Puis nous déroulons les résultats relatifs aux différentes techniques d'estimations de la fonction de densité de probabilité.

Dans le troisième chapitre, nous décrivons brièvement des rappels sur la théorie de la détection radar, les techniques de détection classique et adaptative, ainsi que les commentaires des résultats de simulations des détecteurs CA,GO,SO et OS- CFAR dans les différents types de situations.

Enfin nous concluons ce travail en essayant de dégager les points importants et les perspectives à envisager.



## Chapitre 1 : généralités sur les systèmes Radars :

### Introduction :

Dans ce chapitre, nous introduisons des définitions sur les concepts radar, composition et le principe de fonctionnement, puis nous décrivons les différents types des bruits qui sont les problèmes majeurs de la détection. Enfin nous citons quelques modèles de distributions du Clutter.

### 1.1 Définition d'un radar :

Un radar est un système qui utilise la propriété des ondes électromagnétiques de se réfléchir (en totalité ou partiellement) sur tout obstacle, permettant ainsi de détecter des objets volants (cibles) qui sont situés à l'intérieur de son volume de couverture pour en extraire des informations comme la position, la vitesse...etc.

**RADAR:** est un acronyme de **RA**dio **D**etection **A**nd **R**anging qui signifie «détection et télémétrie radioélectriques ».

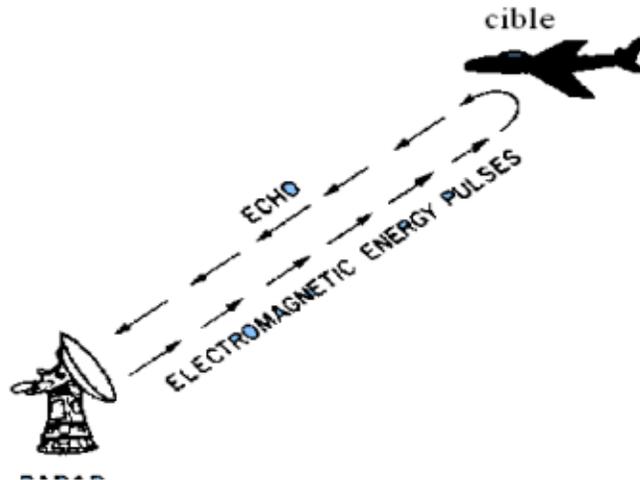


Figure 1: Principe de détection radar.

➤ **Définition d'une Cible:** Au sens large du terme, une cible (Target en Anglais) est tout objet volant qui interfère avec l'onde émise et réfléchit une partie de l'énergie vers le radar.

On fait la distinction entre une cible qui est l'objet qu'on veut détecter et le « Clutter » qui représente les objets non désirées (réflexions de la mer, de la terre, pluie, oiseaux, insectes, météorites,...) qui interceptent aussi l'énergie et la renvoient.

### 1.2 Classification des radars :

Selon l'information recherchée, les radars possèdent différentes technologies. Différentes classifications sont utilisées en fonction de certains critères : type de cible (primaire, secondaire), objectif (surveillance, poursuite, guidage...), position relative de l'émetteur et le récepteur (mono-statique, bi-statique), type de signal (radar à impulsions, radar à onde continue), la résolution (radars conventionnels, radars haute résolution).

Pour atteindre des performances optimales de la détection, il faut déjà choisir la bonne méthode de détection, mais avant tout, il faut identifier et caractériser précisément la nature du Clutter pour pouvoir l'extraire et le séparer du signal utile, pour détecter ensuite la cible.

Il existe de nombreuses sources de signaux malvenus, que les radars doivent pouvoir ignorer plus ou moins, afin de se focaliser uniquement sur les cibles intéressantes. Ces signaux malvenus peuvent avoir des origines internes ou externes, passives ou actives. La capacité d'un radar à surmonter ces nuisances définit son rapport signal sur Clutter (SCR): plus le SCR est grand, plus le radar peut séparer efficacement une cible des signaux parasites des alentours.

### 1.3 Différents types de bruit et signaux parasites :

Tout signal autre que le signal utile est un parasite qui gêne la détection radar. De plus, ce qui est pour notre cas un signal indésirable peut être le signal utile pour une autre application; comme par exemple les échos atmosphériques qui représentent des bruits pour les radars de détection de cibles aériennes, mais qui sont des signaux utiles des radars météorologiques [1]. Les signaux parasites dans les radars de surveillances aériennes sont constitués de :

#### 1.3.1 Bruit interne (thermique) :

Ce type de bruit est créé par l'agitation thermique des électrons dans les différentes composantes qui constituent le récepteur radar. C'est une source interne de variations aléatoires du signal, que tous les composants électroniques génèrent de façon inhérente à différents degrés. Il apparaît typiquement comme constitué de variations aléatoires superposées au signal d'écho reçu par le radar, lequel est celui qu'on recherche. Si la puissance du signal désiré est faible, il est donc difficile de le discerner du bruit thermique, le but alors est de minimiser le bruit thermique et ceci en minimisant le facteur de bruit (Le facteur de bruit est une mesure du bruit produit par un récepteur comparé à celui produit par un récepteur idéal), ce rapport doit être minimal. Dans le cas des radars modernes, grâce aux hautes

performances de leurs récepteurs, le bruit interne est inférieur ou égal au bruit de l'environnement extérieur, sauf si le radar est pointé vers un ciel dégagé, auquel cas l'environnement est si froid qu'il génère très peu de bruit thermique.

### 1.3.2 Écho parasite :

Des échos parasites qui sont par définition inintéressantes pour l'opérateur radar. Les causes de ces échos sont :

- Des objets naturels tels que le sol, la mer, les forêts, les précipitations (telles que la pluie, la neige ou la grêle), les tempêtes de sable, les animaux (particulièrement les oiseaux en vol de groupe), les turbulences atmosphériques, et d'autres effets atmosphériques (par exemple les chutes de météores);
- Des objets fabriqués par l'homme tels que les immeubles ou des paillettes métalliques lâchées intentionnellement comme contre-mesures dans la guerre électronique;
- Des réflexions venant de trajets par réflexions multiples sur une cible. Ainsi, le faisceau radar frappe une cible et comme l'onde émise se réfléchit dans toutes les directions, une partie peut être réfléchi sur une autre cible et retourner au radar. Comme le temps met pour cette seconde réflexion pour atteindre le radar est plus long que le retour direct, elle sera placée au mauvais endroit. On peut ainsi obtenir deux cibles au lieu d'une;
- Des objets très réfléchissants visibles à travers d'un lobe secondaire de l'antenne, alors que l'antenne pointe vers une zone moins réfléchissante. On verra alors un fantôme dans la direction de pointe le lobe principal.

Ces échos parasites de tout objet autre qu'une cible, et en l'absence de brouillage accidentel (interférence avec un autre radar par exemple) ou intentionnel (brouilleur ennemi), sont appelés fouillis, auxquels les anglo-saxons leur consacrent le mot équivalent «Clutter».

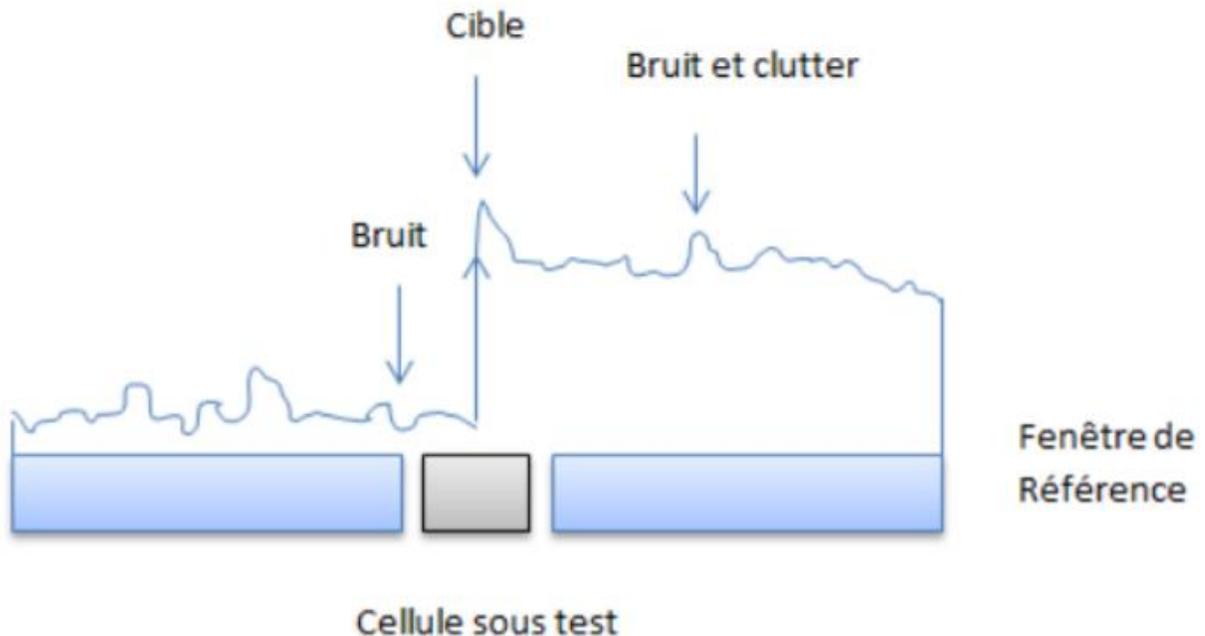
### 1.3.3 Fouillis ou Clutter :

Nous pouvons classier les Clutters en deux grandes familles. Les Clutters de surface qui peuvent représenter les échos de mer et de la terre et les Clutters de volume qui sont généralement liés aux phénomènes météorologiques. Ces deux types de Clutter sont tels que le rapport du signal sur Clutter (Signal to Clutter Ratio, SCR) est très grand par rapport au rapport signal sur bruit (Signal to Noise Ratio, SNR). La présence des interférences induit, quant à elle, un rapport signal à interférence (Signal

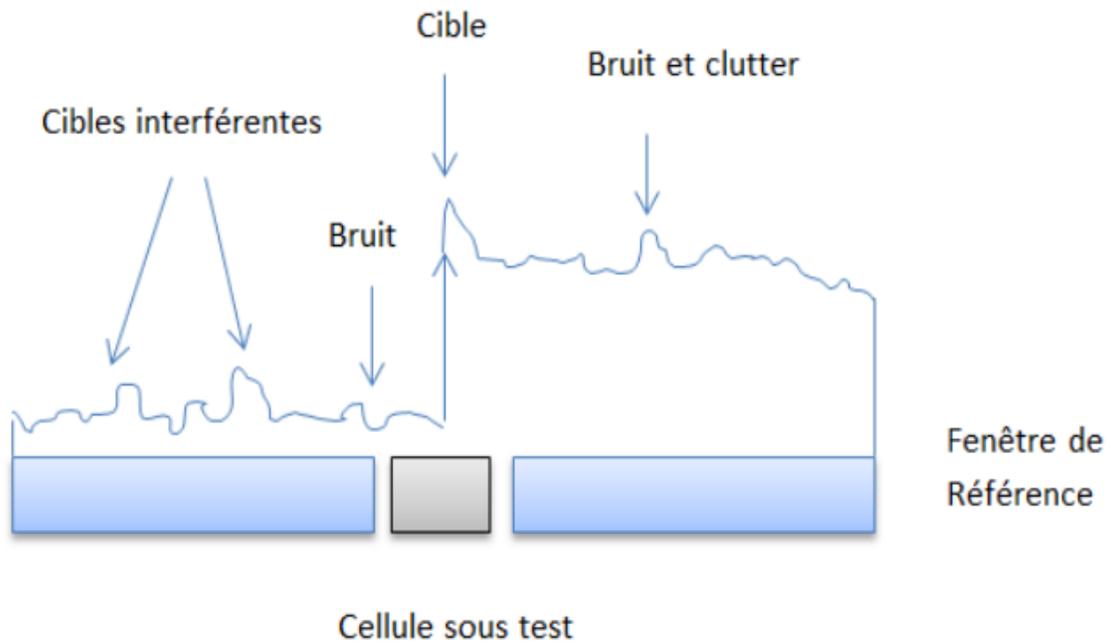
to Interference Ratio, SIR) beaucoup plus grand que le SNR. Le Clutter peut être considéré comme homogène ou non homogène. La non-homogénéité est causée essentiellement par l'effet de bord du Clutter et/ou par la présence des cibles interférentes. Durant ces dernières années, beaucoup de progrès ont été effectués pour modéliser les Clutter selon l'environnement dans lequel le radar opère (terre, mer, atmosphère, ...etc.).

### 1.3.4 Bord de Clutter :

C'est une transition entre deux milieux de natures différentes telle que la transition d'un milieu terrestre vers un milieu maritime ou d'une zone claire vers une zone nuageuse et inversement produit un changement brusque dans la puissance du Clutter. Elle pose deux types de problèmes dans la détection. Le premier se produit lorsque la cellule sous test (CST) baigne dans le bruit thermique alors qu'une partie des cellules de référence contient du Clutter et du bruit thermique. Dans ce cas, le niveau du seuil de détection augmente, ce qui a pour effet de dégrader la détection. Ce phénomène est appelé effet de masquage (masking effect), comme le montre la Figure 3. Le deuxième problème, quant à lui, se manifeste lorsque la cellule sous test baigne dans le bruit thermique et le clutter, alors qu'une partie des cellules de référence contient du bruit thermique uniquement, comme le montre la Figure 2.



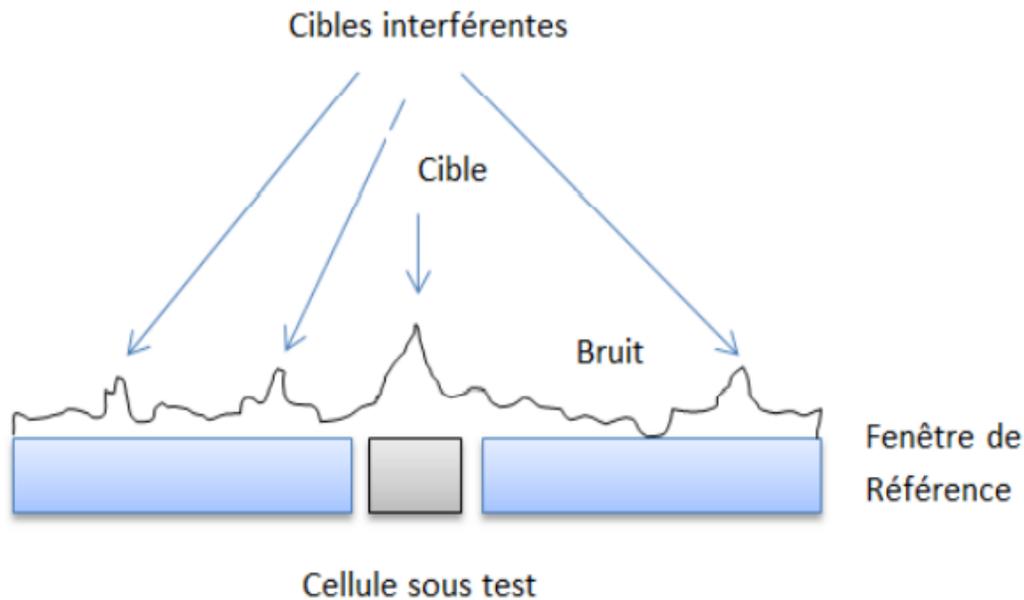
**Figure 2:** Cellule sous test dans le bruit thermique en présence d'un bord de clutter.



**Figure 3:** Cellule sous test dans le bruit thermique et le clutter.

### 1.3.5 Cibles interférentes :

Egalement appelées cibles secondaires ou interférentes, elles peuvent apparaître dans une ou plusieurs cellules de référence sous forme de pics. Elles peuvent se situer en aval et/ou en amont de la CST (Cellule Sous Test). Dans ce cas, le seuil de détection augmente et la probabilité de détection se dégrade énormément et aussi connu sous le nom d'effet capture. La présence des interférences induit un rapport de la puissance du signal sur la puissance de l'interférence (Signal-to-Interference Ratio, SIR) beaucoup plus grand que le SNR.



**Figure 4:** Cellule sous test dans le bruit thermique en présence des cibles interférentes.

### 1.4 Fluctuation des cibles :

#### 1.4.1 Fluctuation des cibles :

Si un radar de veille envoie des impulsions dans l'espace à surveiller avec une puissance, une cible se trouvant dans cet espace, frappée par cette onde électromagnétique, réfléchit une certaine puissance est par définition sa Surface Equivalente Radar (SER). Par conséquent, le signal provenant de la cible est lié au pouvoir réfléchissant de cette dernière [1].

Une cible se comporte comme une antenne de forme complexe. Elle intercepte une part de l'énergie dans laquelle elle baigne; elle absorbe une certaine quantité et réémet le reste dans toutes les directions. L'énergie émise dans la direction du radar est fortement fluctuante et dépend énormément de la forme de la cible, de l'angle d'observation, de la fréquence de l'onde et de la polarisation, ...etc. C'est ainsi qu'on attribue à chaque cible une surface de réflexion effective appelée « Surface Équivalente Radar » (SER) ou « Radar Cross Section » (RCS) en anglais.

La SER est liée à la puissance reçue  $P_r$  par la cible lorsqu'elle est immergée dans un espace où règne une densité surfacique de puissance  $W$

$$P_r = A_e W \quad (1-1)$$

Où  $A_e$  désigne la surface équivalente de la cible. L'équation (1-1) montre que plus la SER est grande, plus l'amplitude du signal réfléchi n'est élevée [2].

### 1.4.2 Fluctuation de la SER :

La SER d'une cible réelle n'est pas constante dans le temps lorsque celle-ci est en mouvement. Il en résulte que quand la cible entre dans le faisceau émise par le radar, elle reçoit un groupe d'impulsion pendant la durée de l'illumination, et réfléchit en conséquence un train d'impulsion dont l'amplitude et la phase fluctuent plus ou moins lentement d'une impulsion à une autre.

Selon leur rapidité, on distingue deux types de fluctuations de la SER :

- Dans le premier cas: la SER fluctue à chaque balayage.
- Dans le deuxième cas: la SER fluctue à chaque impulsion [3].
- Les fluctuations de balayage à balayage: si l'amplitude des impulsions réfléchis situés dans le même groupe est constante.
- Les fluctuations à chaque impulsion: elle est d'impulsion à impulsion si cette amplitude varie d'une impulsion à une autre dans le même groupe.

### 1.4.3 Modèles de fluctuation des cibles :

La SER d'une cible est considérée comme une variable aléatoire associée à une densité de probabilité que nous devons établir à partir d'une observation donnée. La SER doit être considéré comme un processus aléatoire défini par sa fonction de densité de probabilité.

Swerling a proposé des modèles de fluctuations des cibles, qui diffèrent par le taux de fluctuation et la fonction de distribution de la SER. Ces modèles sont présentés ci-dessous :

#### ▪ Swerling I :

Dans ce cas, nous supposons que les impulsions reçues de la cible ont une amplitude constante pendant toute la durée d'illumination et statistiquement indépendantes d'un passage de l'antenne à une autre «Scan-to-Scan fluctuation». Cette supposition ignore l'effet de l'antenne sur l'amplitude de l'écho, l'enveloppe du signal réfléchi à la sortie du détecteur quadratique suit une loi exponentielle de la forme :

$$f(s) = \frac{1}{\sigma^2} e^{-\frac{s}{\sigma^2}} \quad (1-2)$$

Où  $\sigma^2$  représente la puissance moyenne du signal reçu.

### ▪ **Swerling II :**

Dans ce cas, l'enveloppe du signal suit la même loi que celle du type Swerling I donnée par l'équation (1-2), mais les fluctuations sont plus rapides que dans le premier cas de telle façon que les amplitudes sont indépendantes d'une impulsion à une autre « Pulse-to-Pulse fluctuation » pour chaque balayage.

### ▪ **Swerling III :**

Pour ce modèle, les fluctuations sont considérées lentes comme dans le cas du type Swerling I; c'est à dire «Scan-to-Scan fluctuation», mais la densité de probabilité de l'enveloppe du signal à la sortie du détecteur quadratique est donnée par l'équation :

$$f(s) = \frac{4s}{\sigma^2} e^{-\frac{2s}{\sigma^2}} \quad (1-3)$$

### ▪ **Swerling IV :**

Pour ce type de modèle, les cibles possèdent des fluctuations rapides et les amplitudes sont indépendantes d'une impulsion à une autre « Pulse-to-Pulse fluctuation », la densité de probabilité de l'enveloppe du signal à la sortie du détecteur quadratique est la même que celle donnée par l'équation (1-3).

La densité de probabilité utilisée pour les cas du type Swerling I et II est relative aux cibles complexes constituées de plusieurs réflecteurs élémentaires (scatterers) indépendants et aucun réflecteur n'est prépondérant. Ce modèle est utilisé pour représenter les fluctuations des échos d'avions et la réflexion sur la plupart des terrains. Par contre, la densité de probabilité des cas de type **Swerling III** et **IV** est utilisée pour modéliser des cibles composées d'un réflecteur prédominant constant et d'un ensemble de petits réflecteurs indépendants. Ce modèle peut être appliqué pour les missiles et les satellites. Dans le cas où une seule impulsion est transmise pendant la durée d'illumination, le modèle **Swerling I** est équivalent au modèle de **Swerling II**, et le modèle de Swerling III est équivalent au modèle de **Swerling IV** [2].

cas	application	Types de fluctuation	P D F
SW0	/	Pas de fluctuation	/
SW1	Plusieurs réflecteurs comparables (avion)	«Scan-to-Scan fluctuation» Lente : l'amplitude des n impulsions reçues est constante pendant le balayage (scan) mais varie de façon indépendante de balayage à balayage	$f(s) = \frac{1}{\sigma^2} e^{-\frac{s}{\sigma^2}}$ $\sigma > 0$
SW2	Plusieurs réflecteurs comparables (avion)	«Pulse-to-Pulse fluctuation » Rapide : fluctuations indépendantes d'une impulsion à l'autre (Radars à agilité de fréquence)	$f(s) = \frac{1}{\sigma^2} e^{-\frac{s}{\sigma^2}}$ $\sigma > 0$
SW3	Cibles avec un réflecteur principal (missile)	«Scan-to-Scan fluctuation» Lente : l'amplitude des n impulsions reçues est constante pendant le balayage (scan) mais varie de façon indépendante de balayage à balayage	$f(s) = \frac{4s}{\sigma^2} e^{-\frac{2s}{\sigma^2}}$ $\sigma > 0$
SW4	Cibles avec un réflecteur principal (missile)	«Pulse-to-Pulse fluctuation » Rapide : fluctuations indépendantes d'une impulsion à l'autre Radars à agilité de fréquence	$f(s) = \frac{4s}{\sigma^2} e^{-\frac{2s}{\sigma^2}}$ $\sigma > 0$

Tableau 1: Modèle de fluctuation des cibles de Swerling.

## 1.5 Modèles de distributions de Clutter :

A cause de sa variabilité, le clutter est décrit par sa fonction de densité de probabilité, cependant plusieurs modèles statistiques (distributions) ont été proposés pour caractériser la fluctuation du clutter. Parmi ces modèles on trouve le clutter de distribution Rayleigh, Weibull, K-distribuée, Log-Normal,....etc.

### 1.5.1 Modèle Gaussien :

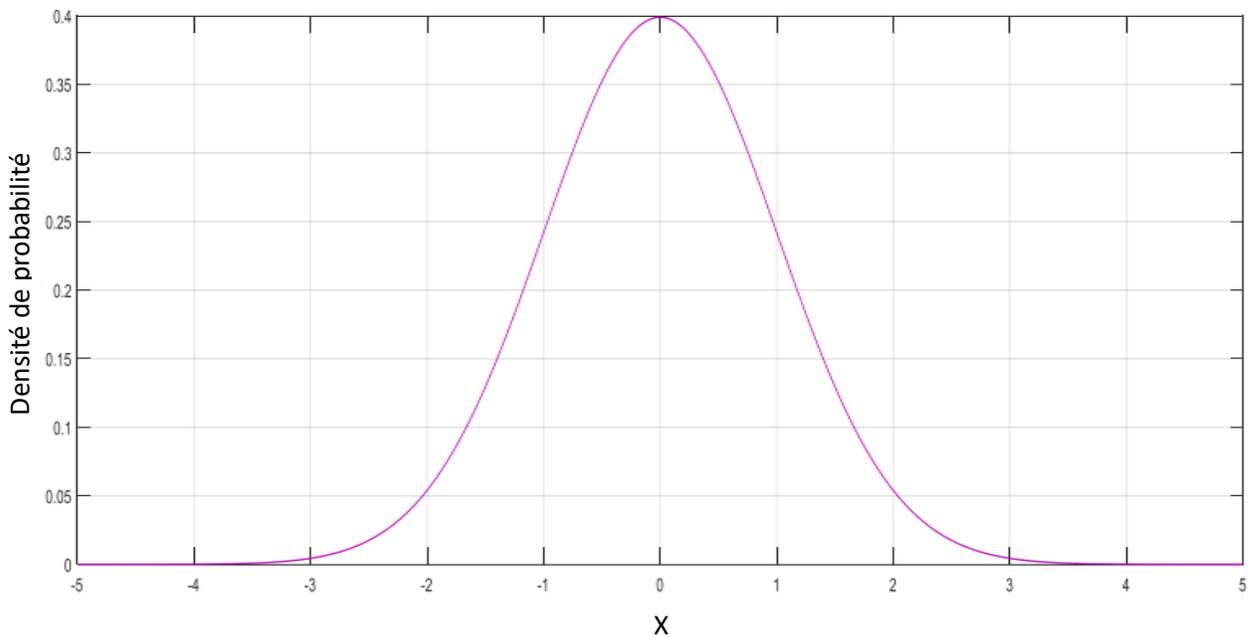
Une variable aléatoire  $X$  de moyenne  $m$  et de variance  $\sigma^2$  suit une loi gaussienne (où normale)  $N = (m, \sigma^2)$  quand sa fonction densité de probabilité s'écrit sous la forme :

$$f_x = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (1-4)$$

Une gaussienne est entièrement définie par  $m$  et  $\sigma^2$ . Quand  $m = 0$  et  $\sigma^2 = 1$ , la loi est dite normale standard alors la variable aléatoire  $Y = \sigma X + m$  suit la loi normale  $N = (m, \sigma^2)$ .

Le modèle gaussien occupe une position centrale en statistiques. Les travaux de Gauss et Laplace ont introduit cette distribution en science principalement en astronomie.

La Figure 5 représente la très familière forme en cloche de la loi gaussienne standard (centrée et réduite). On y constate que le modèle gaussien est centré autour de sa valeur moyenne et ses queues diminuent exponentiellement avec le carré de  $x$ . cette distribution offre souvent une bonne approximation de la partie centrale des données. Etant donné que dans plusieurs applications les décisions sont basées sur les queues des distributions et que les variables aléatoires suivent des distributions à skewness élevé, le modèle normale peut dans certains cas mener à des décisions erronées.



**Figure 5:** Densité de probabilité normale standard.

### 1.5.2 Modèle de Log-normal :

Le modèle gaussien ne représente pas convenablement le clutter quand la cellule de résolution et/ou l'angle de dépression sont faibles. Dans ces conditions, il est plus probable d'avoir de grandes valeurs d'amplitudes de clutter « longues queues » que celles obtenues dans un modèles gaussien. La distribution log-normal décrit le clutter non gaussien qui caractérise par une longue queue.

Pour une variable aléatoire  $X$  suit une loi log-normale, sa fonction de densité de probabilité est donnée par la formule suivante [2] :

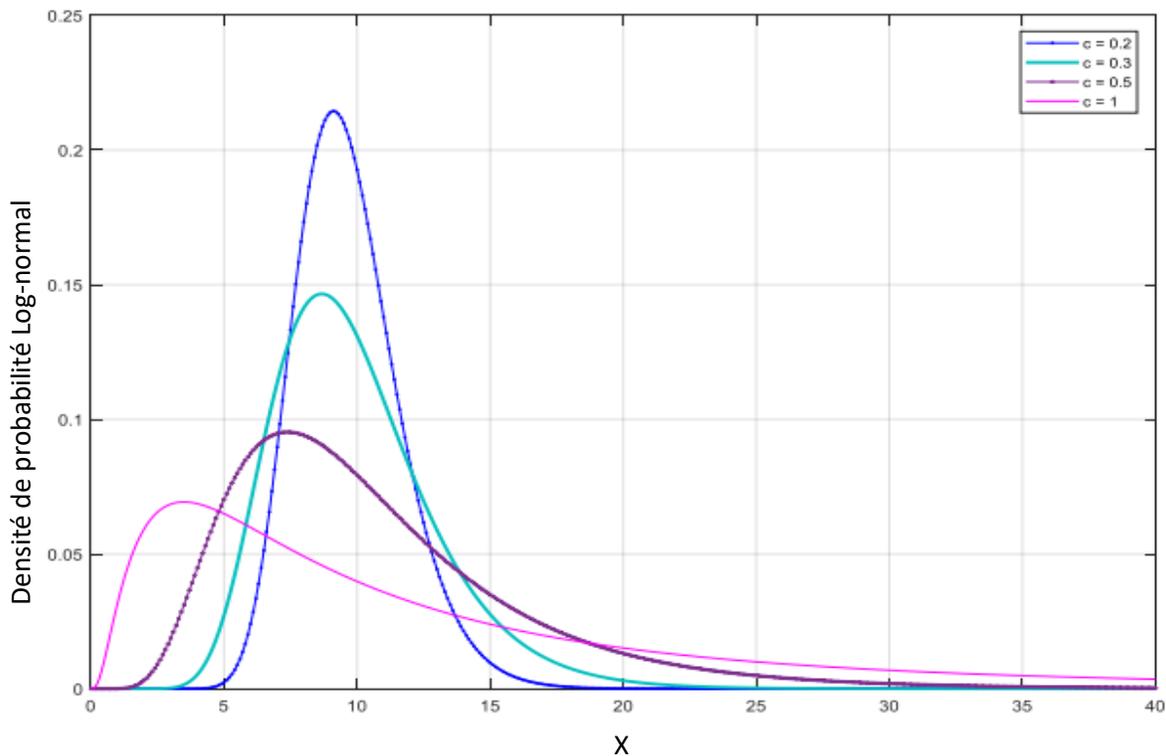
$$f_x(x) = \begin{cases} \frac{1}{b\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln[(x)-b]}{c}\right)^2\right) \\ 0 & \text{ailleurs} \end{cases} \quad (1-5)$$

Avec :

$b$ : Paramètre de l'échelle.

$c$ : Paramètre de la forme.

La distribution log-normale a été introduite depuis plus d'un siècle, par Galto [4], elle est devenue par la suite un modèle intéressant à cause de sa forme qui convient naturellement à beaucoup de modèle de données. La Figure 6 représente la loi log-normale pour  $b = 0$  et différentes valeurs de  $c$ .



**Figure 6:** Densité de probabilité log-normale.

La distribution log-normale est généralement utilisée pour représenter plus convenablement un clutter de sol ou de mer.

### 1.5.3 Modèle de Weibull :

Une variable aléatoire  $X$  suit une distribution Weibull si sa densité de probabilité s'écrit sous la forme :

$$f_x(x) = \begin{cases} \frac{c}{b} \left(\frac{x}{b}\right)^{c-1} \exp\left(-\left(\frac{x}{b}\right)^c\right) & \text{Si } x > 0 \\ 0 & \text{ailleurs} \end{cases} \quad (1-6)$$

Avec :

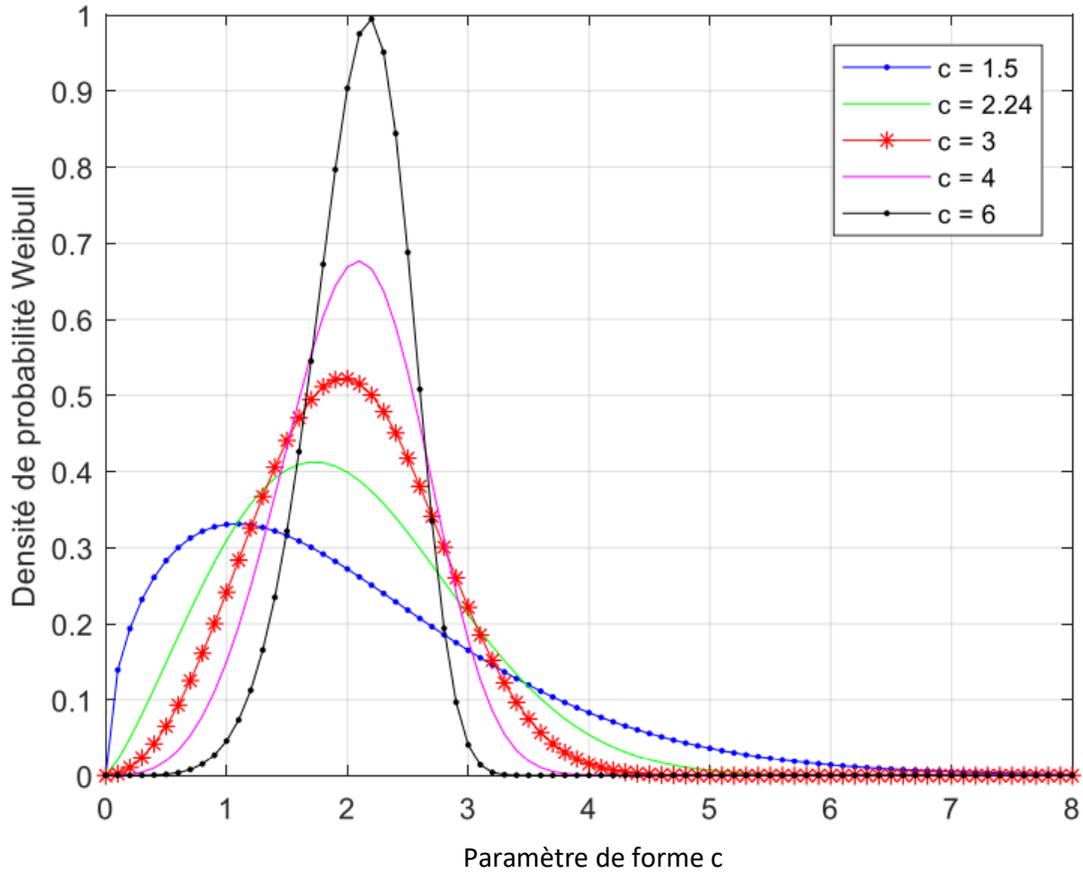
$c > 0$  : Paramètre de la forme ;

$b > 0$  : Paramètre d'échelle.

Cette distribution a été introduite à l'origine par Fisher et Tippet en 1928 [2].

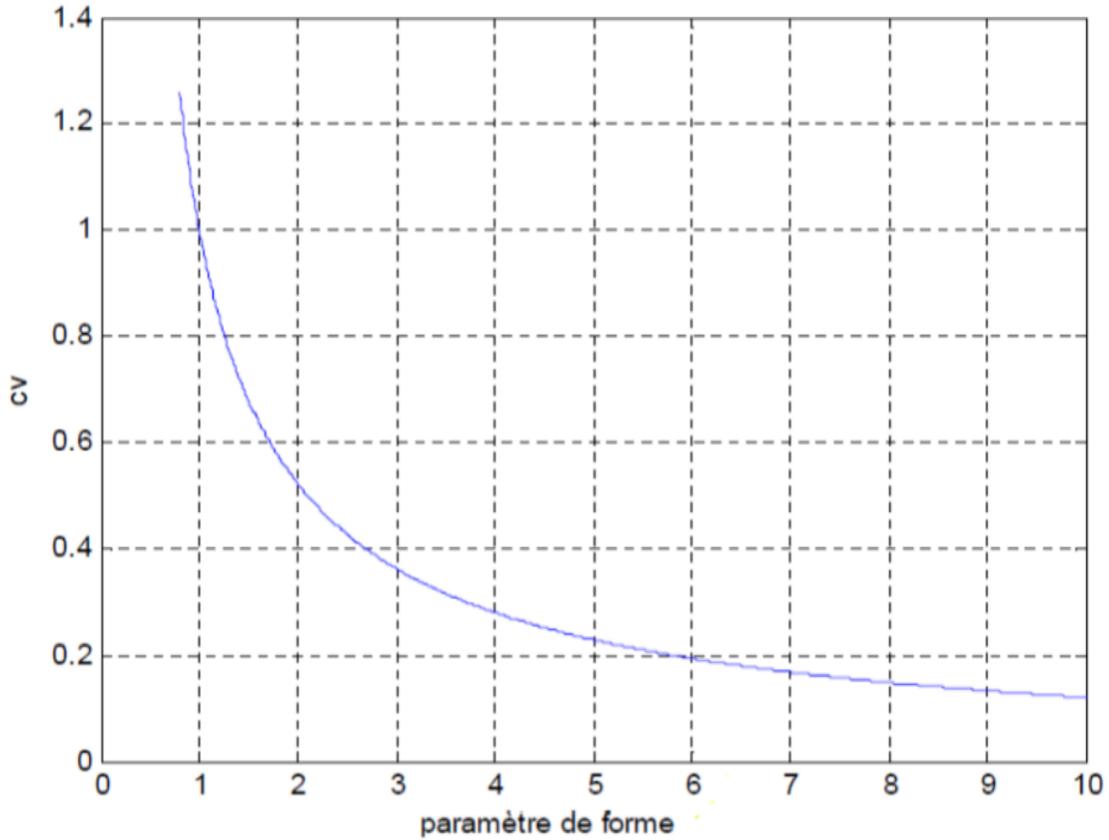
En 1939, le physicien suédois Weibull a utilisé la même distribution en se basant sur des exigences pratiques dans le domaine de l'analyse de données, mais ce n'est qu'en 1951 que l'un des travaux de ce dernier suscita un intérêt particulier parmi la communauté scientifique concernée par la modélisation statistique des données.

La figure (7) représente la densité de probabilité de Weibull correspondant à différentes valeurs du paramètre de forme.



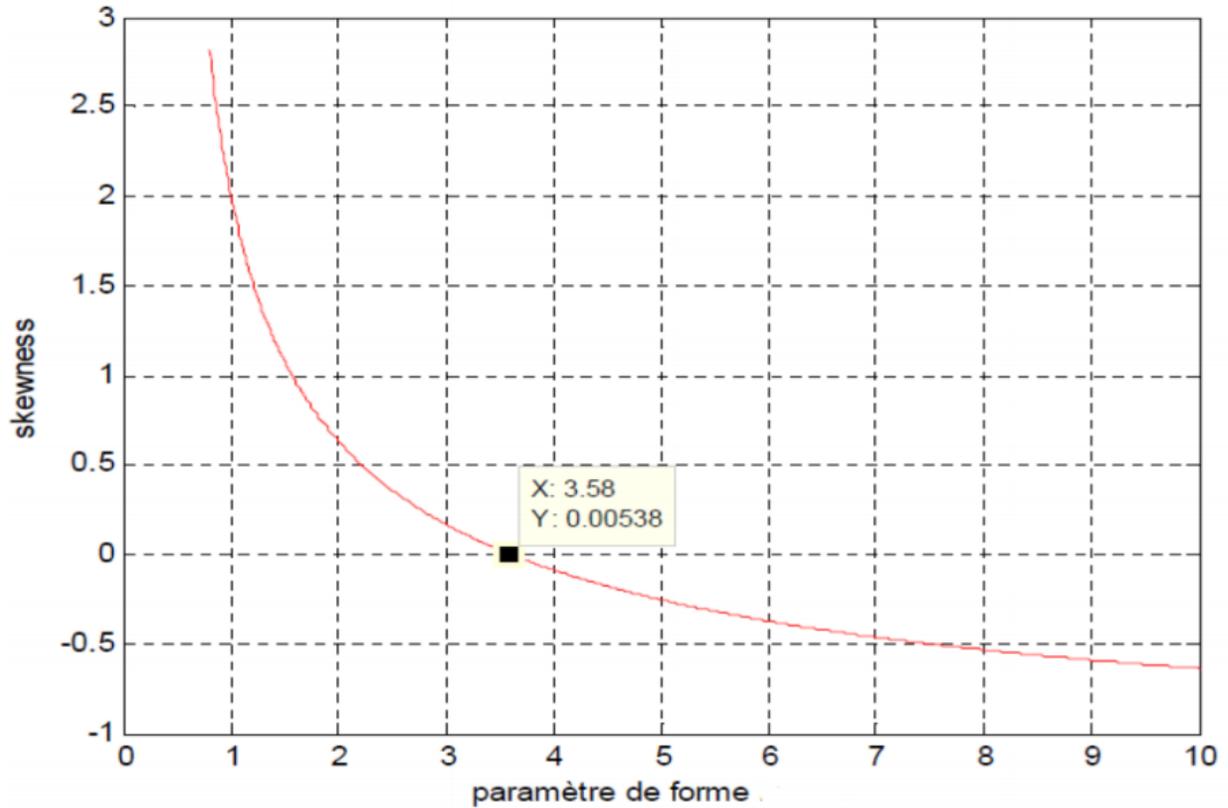
**Figure 7:** Densité de probabilité Weibull.

Les Figures 8, 9 et 10 représentent respectivement les variations du coefficient de variation et des facteurs de forme (Skewness et Kurtosis) en fonction du paramètre de forme de la densité de probabilité de Weibull.



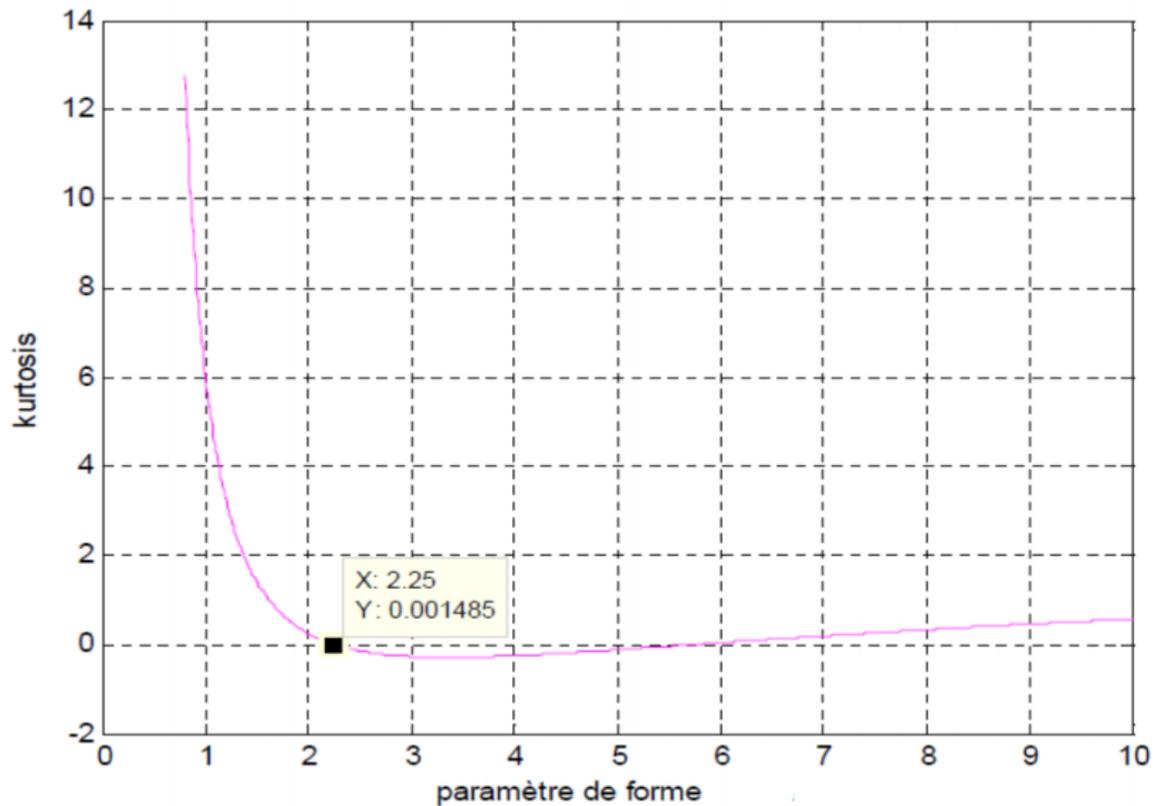
**Figure 8:** Evolution du coefficient de variation (cv) de la loi Weibull.

La Figure 9 représente les variations du premier facteur de forme (Skewness) en fonction du paramètre de forme  $c$ , montrent que pour les valeurs de  $c$  inférieures à 3.6, la densité de probabilité présente une queue plus longue sur la droite et une inclinaison vers la gauche (positively Skewed).



**Figure 9:** Variation du premier facteur de forme (Skewness) de la loi Weibull.

La Figure 10 représente les variations du second facteur de forme (Kurtosis) en fonction du paramètre de forme  $c$ , montrant que le Kurtosis croît rapidement pour les valeurs de  $c$  inférieures à 2, on constate une déviation du modèle gaussien, et des queues de plus en plus lourdes (étendues), au voisinage de zéro le Kurtosis tend vers l'infini (queue infiniment lente) et pour  $c = 2.25$ , la loi de Weibull tend à avoir l'apparence d'une loi normale (Kurtosis=0).



**Figure 10:** Variation du second facteur de forme (Kurtosis) de la loi Weibull.

Le modèle de Weibull offre de meilleures conditions pour représenter convenablement la distribution réelle du clutter dans la pratique par rapport aux distributions log-normal et Rayleigh [5].

### 1.5.4 Modèle de Rayleigh :

Quand le paramètre de forme  $c$  de la distribution de Weibull est égale à 2, la loi est dite de Rayleigh, son expression est donnée par :

$$f_x(x) = \begin{cases} \frac{x}{b^2} \exp\left(-\frac{1}{2}\left(\frac{x}{b}\right)^2\right) & \text{Si } x \geq 0 \\ 0 & \text{ailleurs} \end{cases} \quad (1-7)$$

Avec  $b$  : le paramètre d'échelle.  $e^{\left(-\frac{x^2}{2b^2}\right)}$ .

La Figure 11 représente la densité de probabilité de Rayleigh correspondant à différentes valeurs du paramètre d'échelle.

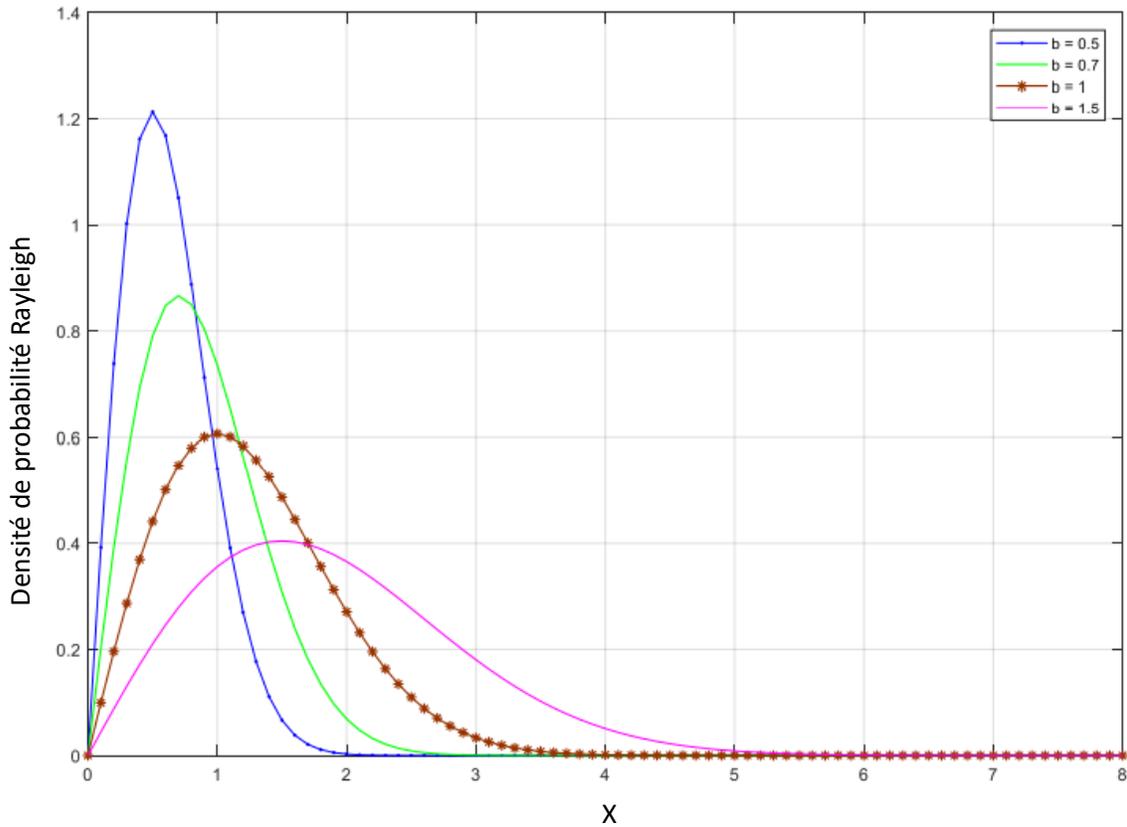


Figure 11: Densité de probabilité de Rayleigh.

Dans les applications radar le modèle Rayleigh est basé sur la supposition qu'il y a un grand nombre de réflecteurs élémentaires uniformes placés aléatoirement dans la surface de clutter illuminée par le radar [2].

### 1.5.5 Modèle de K-distribué :

Ce modèle est utilisé souvent pour modéliser le clutter de mer en utilisant deux composantes. La première, appelée «speckle», suit une loi de Rayleigh. La seconde, appelée «texture», est chi-distribuée [2].

L'amplitude  $X$  du clutter est considérée comme le produit de deux composantes  $T$  : «texture», et  $S$  : «speckle». La PDF est définie comme suit [2] :

$$f_x(x) = \begin{cases} \frac{4}{b\Gamma(\gamma)} \left(\frac{x}{a}\right)^\gamma k_{\gamma-1}\left(\frac{2x}{b}\right) & \text{Si } x > 0 \\ 0 & \text{ailleurs} \end{cases} \quad (1-8)$$

Avec :

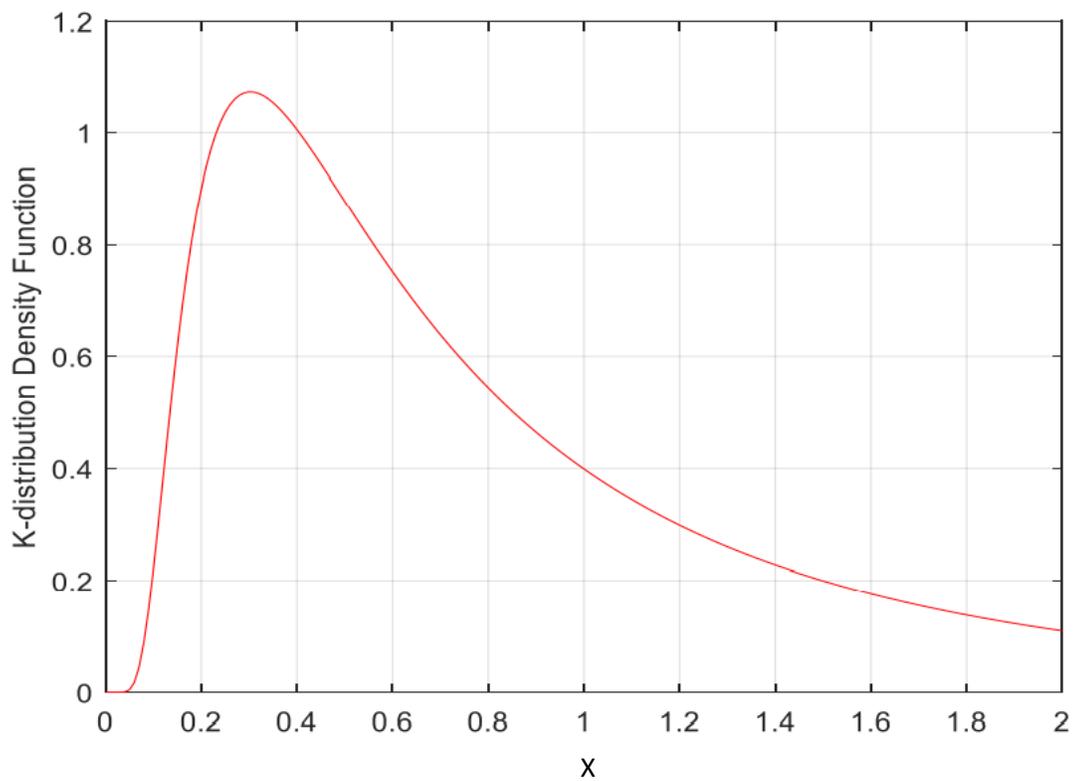
$b$  : Le paramètre d'échelle ;

$\Gamma$  : La fonction Gamma ;

$\gamma$  : Le paramètre de forme ;

$k$  : La fonction de Bessel modifiée.

La Figure 12 représente la densité de probabilité K.



**Figure 12:** Densité de probabilité K-distribué.

## Chapitre 1 : Généralités sur les systèmes Radars

Le tableau I.2 montre des exemples réels sur la modélisation de quelques types du clutter [6].

Type de radar	Largeur d'impulsion	Terre ou Mer	Modélisation de clutter
Basse résolution	2	Montagnes rocheuses	Gaussiens et Weibull
Basse résolution	3	Montagnes rocheuses	Log-normal et Weibull
Haute résolution	0.17	Foret	Log-normal et Weibull
Haute résolution	0.17	Terre cultivée	Log-normal et Weibull
Haute résolution	0.2	Mer	Log-normal, Weibull et k

**Tableau 2:** Exemples de modélisation du clutter.

### Conclusion :

Dans ce chapitre, nous avons rappelé quelques généralités sur les radars tel que ses principes de fonctionnement et leurs compositions puis on a cité les différents types des bruits et des types de fluctuations des SER, et on a terminé par la citation des modèles de distributions de clutter.

Dans le chapitre suivant, nous présentons les méthodes non-paramétriques d'estimation de densité de probabilité.

### Chapitre 2 : estimation de la fonction de densité de probabilité :

#### Introduction :

Dans ce chapitre, nous présentons les différentes techniques non-paramétriques utilisées pour estimer la fonction de densité de probabilité, des plus simples au plus évoluées.

#### 2.1 Introduction à l'estimation de la fonction de densité de probabilité :

Pour estimer les paramètres de la fonction de densité qui on suppose connue par l'estimation paramétrique, nous utilisons par exemple la méthode de l'estimation de vraisemblance ou bien la méthode de moindre carré...etc. mais si cette fonction de densité est inconnue nous utilisons l'estimation non paramétrique.

L'avantage principal de l'estimation non-paramétrique sous-jacente à un ensemble fini d'observations est de ne pas nécessiter d'hypothèses a priori sur l'appartenance de cette densité à une famille de lois connues. L'estimation ne concerne pas les paramètres permettant de sélectionner une loi, mais directement la fonction elle-même.

Soit  $x_1, x_2$  et  $x_n$  n observation équipondérées issues d'une variable aléatoire réelle  $X$  de densité de probabilité réelle  $f(x)$  inconnue. Comment on peut obtenir une estimation de  $f(x)$  à partir de la seule information contenue dans l'échantillon ?

Ce problème que l'on désigne généralement par l'estimation non-paramétrique de la densité de probabilité a fait l'objet de multiples travaux par des méthodes diverses.

##### 2.1.1 Estimation de la fonction de densité de probabilité :

La fonction de densité de probabilité est la valeur en un point  $x$  donne la probabilité d'un événement lorsque la variable aléatoire est égale à  $x$ . Elle donne une description de la distribution de la variable aléatoire à laquelle elle est associée.

La densité de probabilité permet de déterminer les probabilités liées à cette variable à partir de la fonction de distribution (la fonction de répartition).

$$P_r(a < X < b) = \int_a^b f(x) dx; a < b \quad (2- 1)$$

### 2.1.2 Propriétés de la fonction de densité de probabilité :

- La densité de probabilité est continue.
- La fonction de distribution  $\geq 0$ .
- $\int_{-\infty}^{+\infty} f_x(x) dx = 1$ .
- $P(X \leq x) = \int_{-\infty}^x f(x) dx$ .

### 2.1.3 Applications de l'estimation de fonction de densité de probabilité :

L'estimation de la fonction de densité de probabilité qui consiste à construire une estimée de cette fonction à partir d'un ensemble d'échantillons d'apprentissage intervient dans une variété de disciplines. Citons quelques exemples: en codage d'images, l'efficacité d'un code peut se mesurer en termes d'entropie, et donc en termes de densité de probabilité sous-jacente de la classe d'image considérées [7], d'où l'importance d'une estimation fidèle de la densité réelle. En modélisation et simulation des phénomènes physiques par ordinateur on dispose souvent de mesures relatives au processus à simuler, l'estimation de la densité de probabilité à partir de ces mesures permet de générer les variables aléatoires par simulation Monte Carlo.

Finalement dans le domaine de la détection radar qui nous s'intéresse ici, l'estimation de la densité de probabilité du clutter, est utilisée dans ce travail afin d'adopter une stratégie adéquate de détection adaptative (adaptation automatique du calcul du seuil de détection pour maintenir le taux de probabilité de fausses alarmes à une valeur constante) et même on peut associer à chaque modèle estimé un détecteur CFAR qui le convient.

## 2.2 Critère d'erreur :

Lorsqu'on définit un estimateur, il y a un certain nombre de critère qui permettent d'évaluer la similarité de cet estimateur noté par rapport à la vraie densité  $f$  à estimer.

Parmi les nombreux critères proposés dans la littérature, on trouve ISE (Integrated Squared Error), MSE (Mean Squared Error) et MISE (Mean Integrated Squared Error)... [8].

En effet de telles mesures présentent un double intérêt :

D'une part, dans la construction des estimateurs, la minimisation de ces mesures permet des choix optimaux des paramètres de ces estimateurs.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

D'autres parts, l'utilisation de ces mesures dans un contexte comparatives permet de prendre des décisions à la nature de la densité de probabilité estimée par rapport à un ensemble de modèles. Dans ce qui suit, on donne la définition de chaque critère :

### 2.2.1 Erreur quadratique intégrée ISE (Integrated Squar Error) :

L'erreur quadratique intégrée peut également être mesuré sur tout le domaine de variation en l'intégrant sur ce dernier pour donner l'erreur quadratique intégrée définie par :

$$ISE = [f(X), f(h)] = \int (f(X) - f_h(X))^2 \quad (2- 2)$$

### 2.2.2 Erreur quadratique moyenne MSE (Mean Squar Error) :

L'erreur quadratique moyenne pour une estimée par rapport à la vraie densité évaluée au point est donnée par :

$$\begin{aligned} MSE &= [f(x), f(h)] = E[(f(x) - f_h(x))^2] \\ &= E(f^2(x)) - 2E(f(x) \cdot f_h(x)) + E(f_h^2(x)) \\ &= E(f^2(x)) - 2E(f(x) \cdot f_h(x)) + E(f_h^2(x)) + E^2(f_h(x)) - E^2(f_h(x)) \\ &= [E(f_h(x)) - f(x)]^2 + E(f_h^2(x)) - E^2(f_h(x)) \\ &= [\text{biais}(f_h(x))]^2 + \text{var}(f_h(x)) \end{aligned} \quad (2- 3)$$

### 2.2.3 Erreur quadratique moyenne intégrée MISE (Mean Integrated Square Error) :

L'erreur quadratique moyenne intégrée est en réalité une variable aléatoire qui dépend de la densité de probabilité réelle, de l'estimée de cette densité et de l'échantillon (observation) utilisé pour calculer cette estimée [4] ce constant justifié la considération de sa valeur moyenne comme mesure d'écart.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

$$\begin{aligned}MSE(f(x), f_h(x)) &= \int MSE(f(x), f_h(x)) dx \\&= \int E[(f(x) - f_h(x))^2] dx \\&= \int [\text{biais}(f_h(x))]^2 + \text{var}(f_h(x)) dx\end{aligned}\tag{2-4}$$

### 2.2.4 Distance de Kullback- Leibler :

La distance de Kullback-Leibler ou divergence de KULLBACK est généralement utilisée pour mesurer la distance entre deux densités de probabilité. Si on considère la densité de probabilité réelle et son estimée, cette mesure est définie par [9] :

$$d_{kl} = \int_{-\infty}^{+\infty} f_h \log \left[ \frac{f_h}{f} \right] dx\tag{2-5}$$

### 2.2.5 Distance de Hellinger :

La distance de Hellinger est définie par [10].

$$d_H = \int \left[ f_h^{\frac{1}{2}} - f^{\frac{1}{2}} \right]^2 dx\tag{2-6}$$

Outre ces mesures d'écart, des tests statistiques d'ajustement basés sur d'autres mesures de distance entre la densité réelle et son estimée peuvent être construits dans cet ordre d'idées. Bickel et Rosenblatt [11] dans leurs études des estimateurs de densité de probabilité ont introduit des tests fondés sur les mesures tels que :

$$\frac{|f_{\text{estimée}} - f_{\text{réelle}}|}{\sqrt{f_{\text{réelle}}(x)}} \quad \text{Et} \quad \frac{|f_{\text{estimée}} - f_{\text{réelle}}|}{\sqrt{f_{\text{réelle}}(x)}}$$

Certains de ces mesures de distances seront utilisées dans la construction des estimateurs abordée dans ce chapitre, alors que d'autres seront utilisées comme mesure d'ajustement de la densité de probabilité estimée par rapport à un ensemble de densité modèles.

## 2.3 Estimation par histogramme :

L'histogramme est le plus ancien des estimateurs non-paramétriques de densité, l'origine des histogrammes est attribuée par John Graunt au XVII<sup>ème</sup> siècle répandant à l'objectif d'une distribution

## Chapitre 2 : Estimation de la fonction de densité de probabilité

de données. L'histogramme est une méthode graphique qui permet de fournir une description des caractéristiques des données telles que la forme, la localisation, ou la dispersion.

Elle consiste à estimer la densité de la variable aléatoire par le nombre d'occurrences de réalisations appartenant à la classe associée à la valeur  $x$ .

### 2.3.1 Histogramme :

Construire un histogramme à partir d'un ensemble d'observations  $x_1, x_2, \dots, x_n$ , qui sont les réalisations des variables aléatoires réelles indépendantes et identiquement distribuées  $X_1, X_2, \dots, X_n$  de densité inconnue sur l'intervalle de référence fini  $\Omega = [e_{min}; e_{max}]$  en  $p \in \mathbb{R}$  classes  $B_k \in \{1, \dots, p\}$  et à compter le nombre d'observations appartenant à chaque classe  $B_k$ .

Si toutes les classes de l'histogramme ont la même largeur, on dit que l'histogramme est régulier.

On note  $h \in \mathbb{R}$ , la largeur des classes qui est alors appelé le « pas » ou la fenêtre de l'histogramme.

La valeur de  $h$  est donnée par :  $h = \frac{e_{max} - e_{min}}{p}$

L'estimation d'une densité de probabilité pour un échantillon aléatoire :  $x_1, x_2, \dots, x_n$  par l'histogramme s'effectue en choisissant une largeur de pas (bins)  $h$  qui détermine le lissage de l'histogramme, en effet de petites valeurs de  $h$  produisent des histogrammes avec beaucoup de variations, tandis que les grands valeurs de  $h$  conduisent à des histogrammes plus lissés.

L'histogramme correspond alors au nombre d'observations  $x_i$  situées dans le même palier, ramené aux nombre total des observations  $n$ . Le résultat obtenu n'est pas une fonction continue et on peut l'exprimer comme suit :

$$\hat{f}_{hist}(x) = \frac{1}{nh} \sum_{i=1}^n I_{B_k}(X_i); x \in B_k \quad (2-7)$$

Avec : Le nombre d'observations appartenant à chaque classe  $B_k$ .

$$B_k = \sum_{i=1}^n I_{B_k}(X_i)$$

## Chapitre 2 : Estimation de la fonction de densité de probabilité

Où :  $I_{B_k}(X_i)$  est la fonction indice définie par :

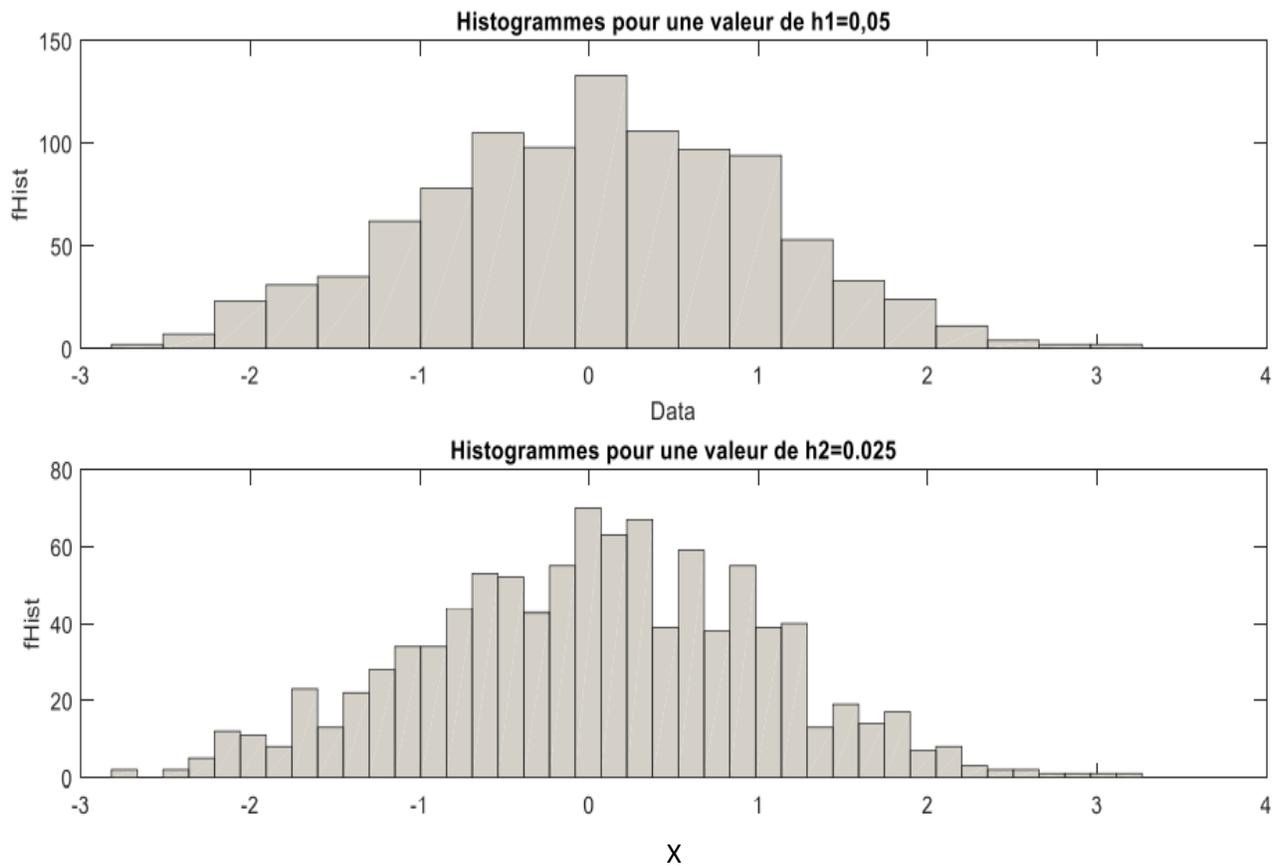
$$I_{B_k}(X_i) = \begin{cases} 1, & \text{si } x_i \in B_k \\ 0 & \text{sinon} \end{cases}$$

$B_k$  : représente le  $K^{\text{ième}}$  bin

$h$  : Est le paramètre de lissage.

Cette méthode est très simple et permet d'obtenir rapidement une idée sur la forme de la loi. En revanche, il n'est pas très judicieux de l'utiliser pour une estimation précise.

La Figure 13 présente un histogramme de 1000 observations tirées aléatoirement d'une loi normale centrée réduite  $N(0,1)$  ces observations sont réparties sur un intervalle de référence  $\Omega = [-4; +4]$ . La fenêtre de l'histogramme est fixée à  $h_1 = 0.05$  et  $h_2 = 0.05$



**Figure 13:** Histogramme de deux valeurs de  $h$  ( $h_1=0.05$  et  $h_2=0.025$ )

### 2.3.2 Choix de paramètre de lissage :

En pratique, on choisit  $h$  en fonction de  $n$ . Les règles les plus utilisées sont :

➤ **Règle de Sturges :**

Prendre le nombre de  $k$  de classes égales à  $1 + \log_2 n$ , en pratique cela revient à prendre :

$$h = \frac{x_k - x_i}{k}$$

Les  $(x_i)$  sont les valeurs d'observations d'un échantillon ordonné par ordre croissant.

La règle de Sturges a tendance de prendre des histogrammes trop lisses [12].

➤ **Règle de Scott :**

La valeur qui minimise l'erreur quadratique moyenne intégrée, MISE est donnée par :

$$h_{opt} = \frac{6}{\int f'(t)^2 dt}$$

En prenant pour la densité de loi normale  $N(\mu, \sigma^2)$ , d'après Scott [13] on peut alors montrer que :

$$h_{opt} = 3.491\sigma n^{-\frac{1}{3}}$$

En estimant  $\sigma$  par l'écart type  $S$  de l'échantillon, on aura :

$$h_{opt} = 3.491Sn^{-\frac{1}{3}}$$

### 2.4 Polygone de fréquence :

Parmi les méthodes de base, souvent non citée la méthode des polygones de fréquence, qui permet d'estimer la densité de probabilité par interpolation linéaire entre les centres des bins d'un histogramme.

L'estimée de la densité de probabilité en utilisant cette méthode est donnée par :

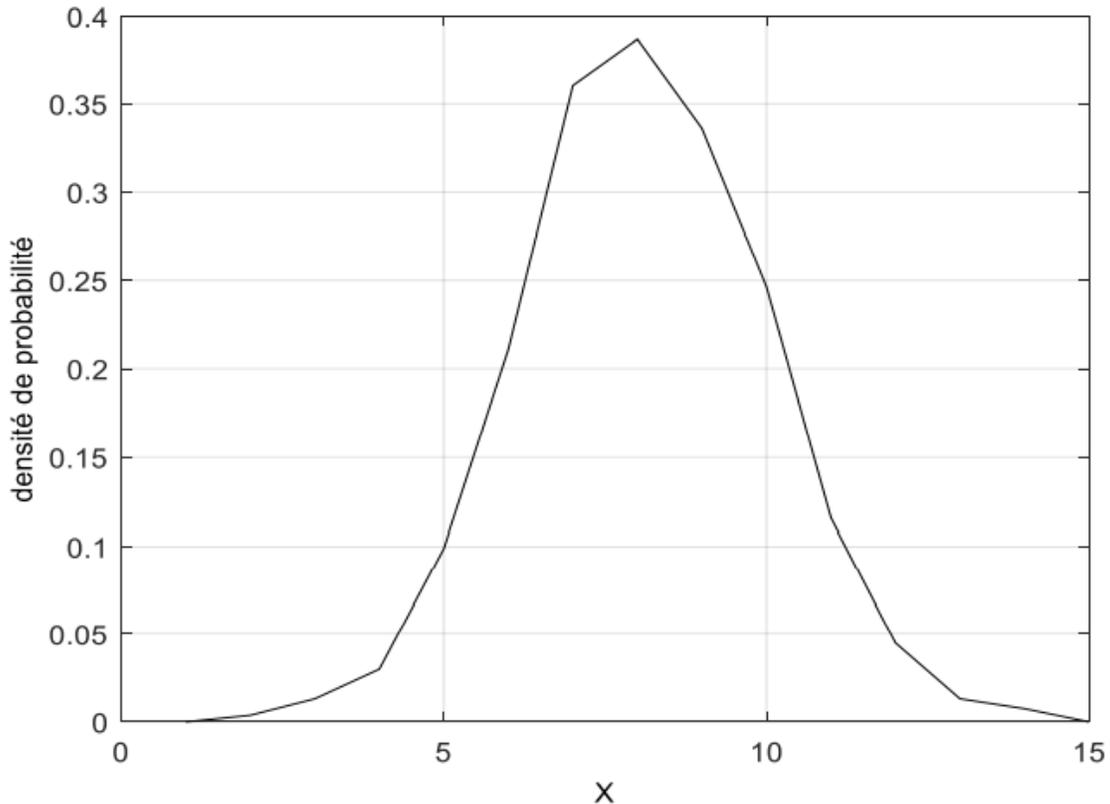
$$\hat{f}_{FP}(x) = \left(\frac{1}{2} - \frac{x}{h}\right) \hat{f}_k + \left(\frac{1}{2} + \frac{x}{h}\right) \hat{f}_{k+1} \quad (2-8)$$

Avec :  $\bar{B}_k \leq x \leq \bar{B}_{k+1}$

$\hat{f}_k, \hat{f}_{k+1}$  sont les valeurs adjacentes de l'histogramme et  $\bar{B}_k$  est le centre du  $k$ ème bin.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

La Figure 14 montre un exemple d'une estimée de la densité de probabilité relative à un échantillon gaussien calculée par la méthode des polygones.



**Figure 14:** Densité de probabilité normale estimée par polygone.

La méthode des polygones fait partie des méthodes non précises qui permettent juste de donner une idée sur la nature de la variable aléatoire, néanmoins il constitue une amélioration des histogrammes conventionnels.

Dans le même ordre d'idées, l'histogramme classique peut être amélioré en utilisant une procédure de moyennage de plusieurs versions décalées de ce dernier.

### 2.4.1 Choix du paramètre de lissage :

Les méthodes de choix du paramètre de lissage sont basées sur l'étude asymptotique de certaines des mesures d'écarts. Pour un histogramme, l'erreur quadratique intégrée moyenne asymptotique en fonction de  $h$  est donnée par [14].

$$AMISE_{Hist}(h) = \frac{1}{nh} + \frac{1}{12} h^2 R(f') \quad (2-9)$$

## Chapitre 2 : Estimation de la fonction de densité de probabilité

$R(\cdot)$  : est une mesure de rugosité définie pour une fonction  $g$  par :

$$R(g) = \int_x^{+\infty} g^2(x) dx$$

La valeur optimale de  $h$  peut être obtenue par minimisation de l'erreur quadratique intégrée moyenne asymptotique AMISE est donnée par :

$$h_{Hist}^{AMISE} = \left( \frac{6}{nR(f')} \right)^{\frac{1}{3}} \quad (2-10)$$

La densité de probabilité à estimer, donc inconnue intervient dans l'expression de la valeur optimale de  $h$ . La solution à ce problème consiste à trouver une règle de choix de  $h$  basée sur des données normalement distribuées, dans lequel cas, l'indice de rugosité est connu, puis déterminer des facteurs de correction pour le cas général.

Pour des données normalement distribuées, on a :

$$R(f') = \frac{1}{4\sigma^3\sqrt{\pi}}$$

$$h_{Hist}^{Norm} = \left( \frac{24\sigma^3\sqrt{\pi}}{n} \right)^{\frac{1}{3}} \approx 3.5\sigma n^{-\frac{1}{3}} \quad (2-11)$$

Deux règles ont été proposées pour l'estimation de  $\sigma$  :

Scott (1979-1992) proposa une règle selon laquelle  $\sigma$  est égal à l'écart type de l'échantillon.

Freedman et Diaconis (1981) proposèrent une règle basée sur l'estimation de l'interquartile IQR au lieu de l'écart type, la formule du calcul de  $h$  est alors :

$$h_{Hist}^{Freedman} = 2IQRn^{\frac{1}{3}} \approx 3.5\sigma n^{-\frac{1}{3}}$$

Pour des données jugées issues de processus à queue lente (skewness non nul), un facteur de correction a été proposé par Scott :

$$\text{facteur de correction} = \frac{\sigma 2^{\frac{1}{3}}}{e^{\frac{5\sigma^2}{4}} (\sigma^2 + 2)^{\frac{1}{3}} (e^{\sigma^2} - 1)^{\frac{1}{2}}} \quad (2-12)$$

La valeur de  $h$  doit être multipliée par ce facteur de correction.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

Pour des échantillons non gaussiens à Kurtosis non nul, Scott(1992) a démontré que la correction sur la valeur de h calculée par référence à la loi normale est négligeable.

### 2.5 Histogrammes décalés ASH (Averaged Shifted Histograms):

Cette méthode est basée sur le choix de l'origine du calcul d'un estimateur classique dans le souci de lisser ce dernier, plusieurs histogrammes sont calculés avec la même largeur h mais à partir de différentes origines, la moyenne de ces histogrammes fait alors office de l'estimée de la densité de probabilité.

La construction d'un histogramme ASH se fait de la manière suivante :

Détermination de l'ensemble des origines données par :

$$t'_0 = t_0 + 0, t_0 + \frac{h}{m}, t_0 + \frac{h}{2m}, \dots, t_0 + \frac{(m-1)h}{m} \quad (2- 13)$$

Calcul des histogrammes relatifs aux m origines.

Calcul de l'histogramme ASH par :

$$\hat{f}_{ASH}(x) = \frac{1}{m} \sum_{i=1}^m \hat{f}_i(x) \quad (2- 14)$$

Sur la figure 15 un exemple d'une densité de probabilité estimée en utilisant un moyennage d'une cinquantaine de versions décalées d'histogrammes conventionnels.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

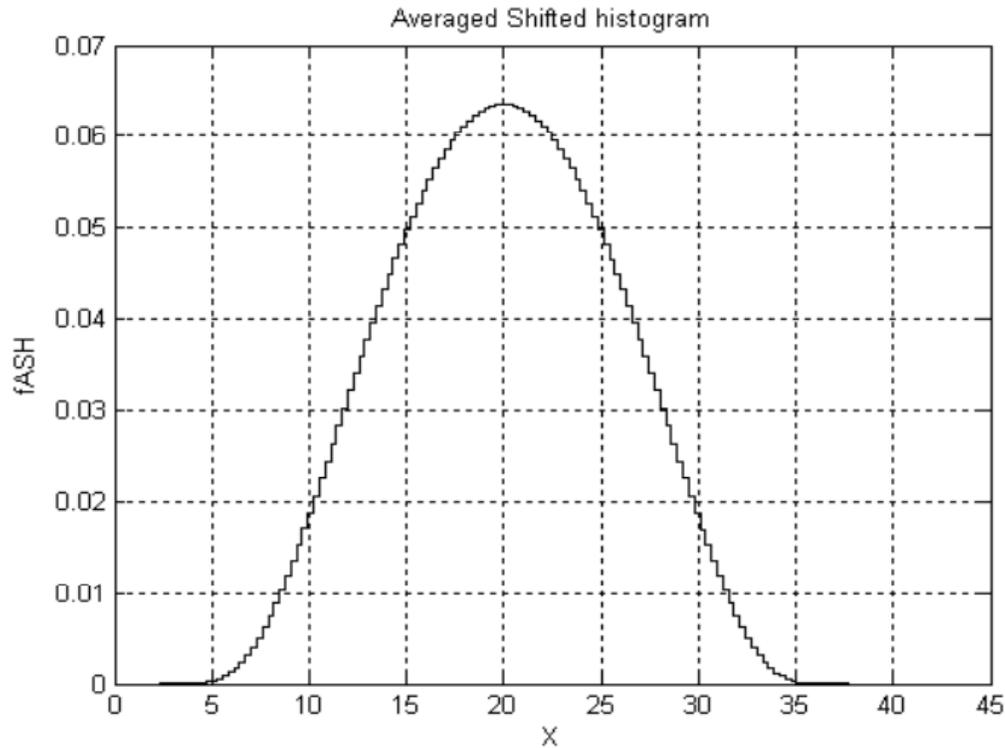


Figure 15 : Densité de probabilité normale estimée par ASH.

### 2.6 Estimation non paramétrique par la méthode du noyau :

#### 2.6.1 Introduction :

La méthode du noyau est l'une des méthodes d'estimation non paramétrique la plus utilisée. Rosenblatt (1956) [15], suivi par Parzen (1962) [14] ont proposé une classe d'estimateurs à noyau d'une densité de probabilité. Cet estimateur est une fonction de deux paramètres : le noyau  $K$  et le paramètre de lissage  $h$ .

Le succès rencontré par cet estimateur s'explique par sa simplicité, sa flexibilité et aussi ses propriétés de convergence. Il laisse l'utilisateur une grande latitude non seulement dans le choix du noyau, mais aussi dans le choix du paramètre de lissage  $h$ .

Soit  $X$  une variable aléatoire de densité de probabilité inconnue  $f$ . supposons que nous avons  $n$  observateurs  $x_1, x_2, \dots, x_n$  provenant de  $X$ . le problème consiste à trouver un estimateur pour la fonction  $f$  à partir de cet échantillon issu de  $X$  pour cela, l'approche non paramétrique est la plus adéquate lorsqu'on ne possède aucune information précise sur la forme et la classe de la vraie densité. Dans

## Chapitre 2 : Estimation de la fonction de densité de probabilité

cette approche, ce sont les observations qui vont nous permettre de déterminer un estimateur pour la densité  $f$ .

Dans cette section, on s'intéresse à la méthode du noyau pour l'estimation de la densité de probabilité. L'estimateur à noyau sera présenté ainsi que ses différentes propriétés statistiques.

### 2.6.2 Définition du noyau de Parzen-Rosenblatt :

En 1956, Rosenblatt [15] a proposé le premier estimateur à noyau pour la densité de probabilité  $f(x)$ . Six ans après, cet estimateur a été généralisé par Parzen (1962) [14], à partir de cette date, cet estimateur a pris le nom de l'estimateur de Parzen-Rosenblatt.

L'idée de l'estimateur par la méthode du noyau consiste à évaluer la densité  $f(x)$  au point  $x$  en comptant le nombre d'observations tombées dans un certain voisinage de  $x$  sur  $\mathbb{R}$ . En 1962, Parzen a généralisé cet estimateur en remplaçant la fonction  $w$  (noyau uniforme) par une fonction noyau  $K$  satisfaisant la condition suivant :

$$\int_{-\infty}^{+\infty} k(u) du = 1$$

Généralement,  $K$  est une densité de probabilité. Donc l'estimateur à noyau de Parzen est définie par :

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right) \quad (2-15)$$

Où :  $h$  est le paramètre de lissage et  $k$  la fonction noyau définie sur  $\mathbb{R}$  et vérifiant les conditions suivantes:

$$\int_{-\infty}^{+\infty} k(y) dy = 1; \int_{-\infty}^{+\infty} yk(u) du = 0; \int_{-\infty}^{+\infty} y^2 k(u) du = \sigma_k^2 \neq 0$$

### 2.6.3 Propriétés de l'estimateur à noyau :

Cette partie est consacrée à la présentation de quelques résultats théoriques [8] sur les propriétés mathématique de l'estimateur à noyau de Parzen-Rosenblatt donné par l'équation

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)$$

Ces propriétés mathématiques sont :

## Chapitre 2 : Estimation de la fonction de densité de probabilité

- L'espérance, le biais et la variance (voir annexe A).
- Le comportement asymptotique du biais et de la variance.
- Le MSE et le MISE de l'estimateur.
- La convergence d'un estimateur à noyau au sens du MSE et MISE.

\* **L'espérance :**

$$\begin{aligned}\mathbb{E}\{f_h(x)\} &= \mathbb{E}\left[\frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)\right] = \frac{1}{nh} \mathbb{E}\left[\sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)\right] \\ &= f(x) + \frac{h^2}{2} f'' \sigma^2(k)\end{aligned}\quad (2-16)$$

\* **Le biais :**

$$\begin{aligned}\text{var}[f_h(x)] &= \text{var}\left[\frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)\right] = \text{var}\left[\frac{1}{n^2 h^2} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)\right] \\ &= \frac{1}{n^2 h^2} \sum_{i=1}^n \mathbb{E}\left[k\left(\frac{x-x_i}{h}\right)\right]^2 - \frac{1}{n^2 h^2} \sum_{i=1}^n \left(\mathbb{E}\left[k\left(\frac{x-x_i}{h}\right)\right]\right)^2 \\ &= \frac{f(x)}{nh} \int_{-\infty}^{+\infty} k^2(y) dy - \frac{f'(x)}{n} \int_{-\infty}^{+\infty} y k^2(y) dy + \frac{h f''(x)}{2n} \int_{-\infty}^{+\infty} y^2 k^2(y) dy \\ &\quad - \frac{1}{n} [f(x) \text{biais}(f_h(x))]^2\end{aligned}\quad (2-17)$$

### 2.6.4 La convergence de l'estimateur à noyau :

Parmi toutes les qualités que peut avoir un estimateur, on s'intéresse souvent à sa consistance, c'est-à-dire, au fait qu'un estimateur  $f_h$  converge ou non vers  $f$ . La convergence d'un estimateur peut être faible (en probabilité) ou forte (presque sûrement ou en moyenne quadratique). On donne quelques résultats de convergence des estimateurs à noyaux :

✓ **Convergence en moyenne quadratique :**

Le MSE d'un estimateur à noyau est donné par la formule suivante :

$$MSE(f(x), f_h(x)) = \frac{h^4}{4} (f''(x))^2 \sigma_2^2 + \frac{1}{nh} f(x) \int_{-\infty}^{+\infty} k^2(y) dy \quad (2-18)$$

## Chapitre 2 : Estimation de la fonction de densité de probabilité

Le théorème 2 de Parzen [14] montre que si :  $\lim_{n \rightarrow \infty} h(n) = 0$  Et  $\lim_{n \rightarrow \infty} nh = \infty$  et  $K$  satisfait aux conditions suivantes :

$$\sup |k(y)| < \infty \quad \text{Et} \quad \lim_{y \rightarrow \infty} y|k(y)| = 0,$$

$$\int |k(y)| dy < \infty \quad \text{Et} \quad \int k(y) dy = 1$$

Alors l'estimateur  $f_h(x)$  est consistant en moyenne quadratique, c'est-à-dire :

$$\lim_{n \rightarrow \infty} MSE(f(x), f_h(x)) = 0 \quad \text{Pour tout } x \text{ pour lequel la densité } f \text{ est continue.}$$

### ✓ Convergence en moyenne quadratique intégrée :

Le théorème 3 de Parzen [14] montre que si  $k$  est un noyau de Parzen-Rosenblatt, et on a :

$$\lim_{n \rightarrow \infty} h(n) = 0 \quad \text{Et} \quad \lim_{n \rightarrow \infty} nh = \infty$$

$$\text{Alors } \lim_{n \rightarrow \infty} MSE(f(x), f_h(x)) = 0 \quad \text{pour } f \text{ définie sur } \mathbb{R}.$$

Cette partie nous a permis d'identifier les paramètres de la méthode de Parzen-Rosenblatt et de souligner la nécessité et l'intérêt du choix du couple  $(k, h)$ . Le noyau  $k$  peut-être discret ou continu et selon la symétrie du domaine de définition.

Il existe deux catégories principales de noyaux : les noyaux symétriques et les noyaux asymétriques.

### 2.6.5 La consistance d'un estimateur :

La consistance d'un estimateur de la densité de probabilité il peut être considéré en un point comme il peut être étudié sur sa totalité. La consistance en un point a été étudiée par Parzen, qui moyennant certaines hypothèses sur le noyau  $K$  et a largeur de la bande  $h$ , a montré que l'estimée tend vers la densité réelle quand la taille de l'échantillon tend vers l'infini. Les hypothèses considérées sont :

#### ✓ Sur noyau :

$$\int |k(y)| dy < \infty \quad \text{Et} \quad \int k(y) dy = 1$$

$$\lim_{y \rightarrow \infty} |yk(y)| \rightarrow 0$$

## Chapitre 2 : Estimation de la fonction de densité de probabilité

✓ **Sur la largeur de bande :**

$$h_n \rightarrow 0$$

$$\lim_{n \rightarrow \infty} nh_n \rightarrow \infty \quad \text{Quand } n \rightarrow \infty$$

L'étude de la consistance globale de l'estimateur nécessite de spécifier dans quelle sens l'estimée approxime la densité réelle [4].

### **2.6.6 Choix du noyau (exemples de noyaux) :**

L'estimateur non paramétrique d'une densité de probabilité par la méthode du noyau nécessite le choix du noyau  $K$ . dans cette partie nous allons faire une brève définition de quelques noyaux usuels :

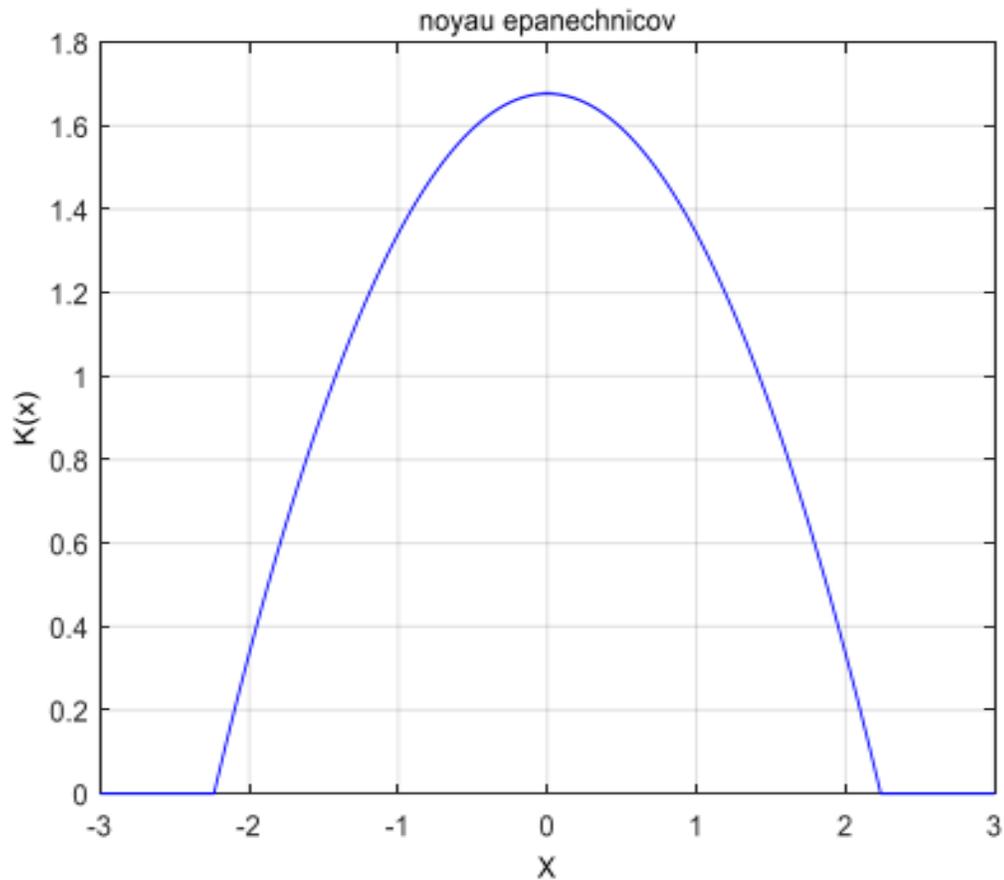
## Chapitre 2 : Estimation de la fonction de densité de probabilité

Noyaux	Définition
Noyau uniforme (Rosenblatt)	<p>Ce noyau a été proposé par Rosenblatt en 1956 [8], l'avantage de ce noyau est la simplicité de sa forme. Il s'écrit sous la forme :</p> $k(u) = \begin{cases} \frac{1}{2} & \text{si }  u  < 1 \\ 0 & \text{sinon} \end{cases}$
Noyau box (boite)	$k(u) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{si } -\sqrt{3} \leq  u  \leq \sqrt{3} \\ 0 & \text{sinon} \end{cases}$
Noyau triangulaire	<p>Ce noyau a un avantage par rapport au noyau de Rosenblatt, il est continu partout ce qui conduit à une estimation de continue. Ce noyau s'écrit sous la forme :</p> $k(u) = \begin{cases} (1 -  u ) & \text{si } -1 \leq u \leq 1 \\ 0 & \text{sinon} \end{cases}$
Noyau Cosine	$k(u) = \begin{cases} \frac{\pi}{4} \cos\left(\frac{\pi u}{2}\right) & \text{si } -1 \leq u \leq 1 \\ 0 & \text{sinon} \end{cases}$
Noyau Sinus	$k(u) = \begin{cases} \frac{1}{2\pi} \left(\frac{\sin \frac{x}{2}}{\frac{x}{2}}\right)^2 & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$
Noyau Gaussien	<p>L'avantage du noyau gaussien est que plus la valeur de est élevée plus on élargit la fenêtre, ce qui a un effet de lissage globale important, mais le cout de calcul dans le cas de ce noyau est très élevée du fait de son support infini. Ce noyau s'écrit sous la forme :</p> $k(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \quad \forall u \in \mathbb{R}$
Noyau de biweight (bi-poids)	<p>Ce noyau est le plus intéressant car il donne une estimation dérivable partout tout en étant simple à mettre en œuvre. En fait, il s'agit du noyau le plus simple parmi les noyaux de forme polynomial dérivable partout. Ainsi, il assure le lissage locale de la fonction .ce noyau est d'une forme très proche du noyau gaussien, il est donc préférable. Il s'écrit sous la forme :</p> $k(u) = \begin{cases} \frac{15}{16} (1 - u^2)^2 & \text{si } -1 \leq u \leq 1 \\ 0 & \text{sinon} \end{cases}$
Noyau triweight (tri-poids):	$k(u) = \begin{cases} \frac{35}{32} (1 - u^2)^3 & \text{si } -1 \leq u \leq 1 \\ 0 & \text{sinon} \end{cases}$
Noyau Epanechnikov (parabolique)	<p>En 1969, Epanechnikov [4] , a donné la forme du noyau défini par :</p>

## Chapitre 2 : Estimation de la fonction de densité de probabilité

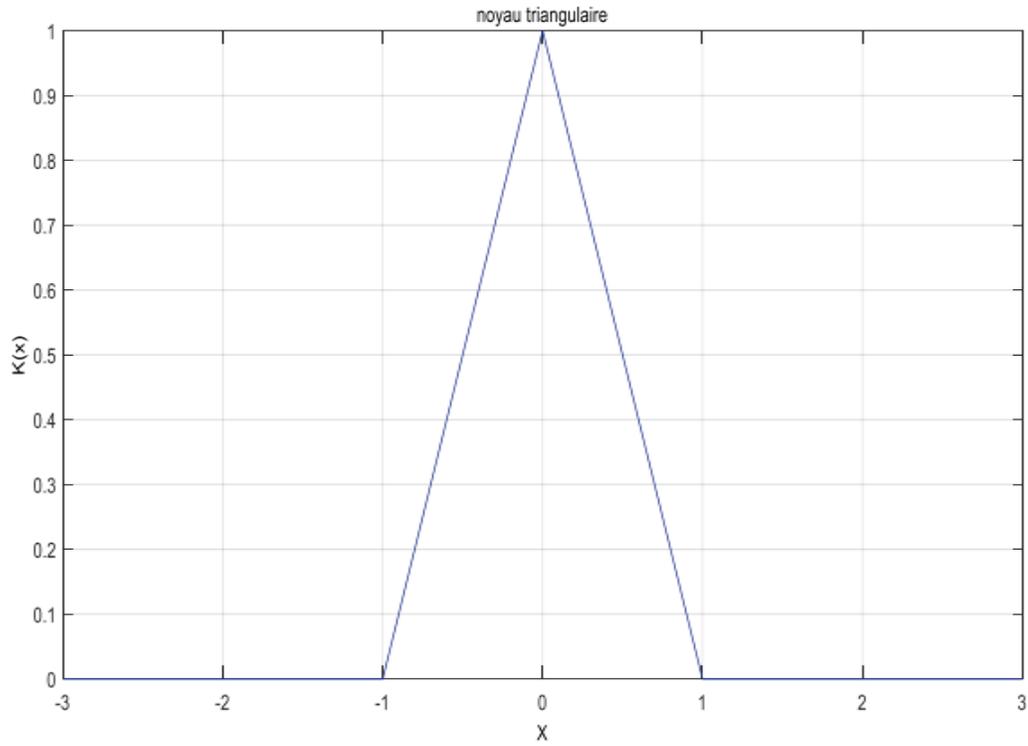
	$k_E(u) = \begin{cases} \frac{3}{4\sqrt{5}} \left(1 - \frac{x^2}{5}\right) & \text{si } x \in [-\sqrt{5}, \sqrt{5}] \\ 0 & \text{sinon} \end{cases}$
Noyau de Silverman	$k(u) = \frac{1}{2} e^{\left(\frac{- t }{\sqrt{2}}\right)} \sin\left(\left(\frac{- t }{\sqrt{2}}\right) + \frac{\pi}{4}\right)$

**Tableau 3:** Exemples de noyaux utilisés pour estimer la fonction de densité de probabilité.

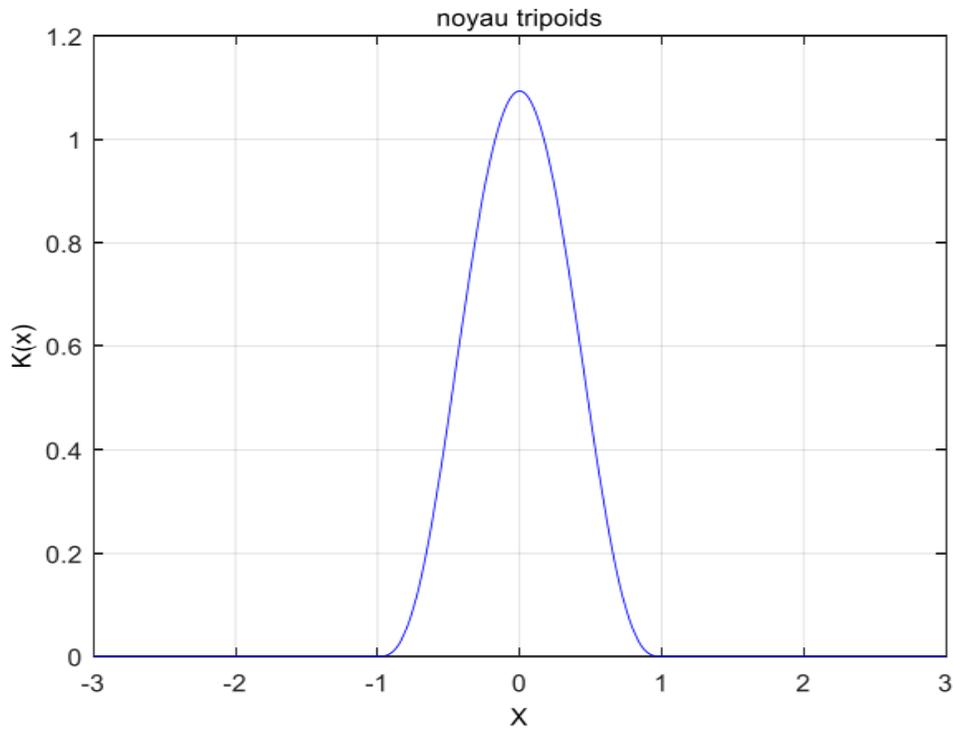


**Figure 16:** exemple de noyau d'epanechnikov utilisé pour estimer les fonctions de densité de probabilité.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

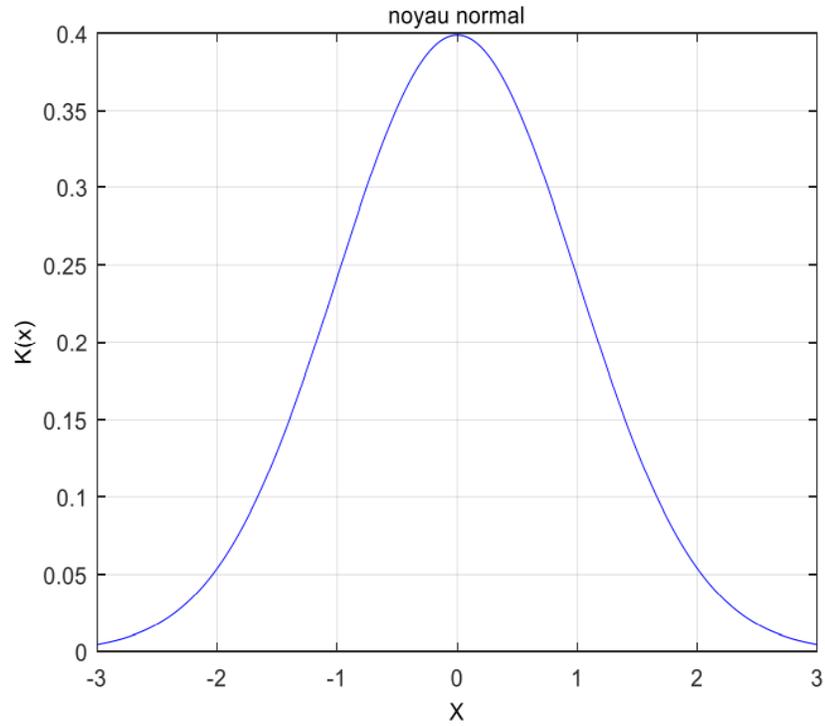


**Figure 17:** exemple de noyau triangulaire utilisé pour estimer les fonctions de densité de probabilité.

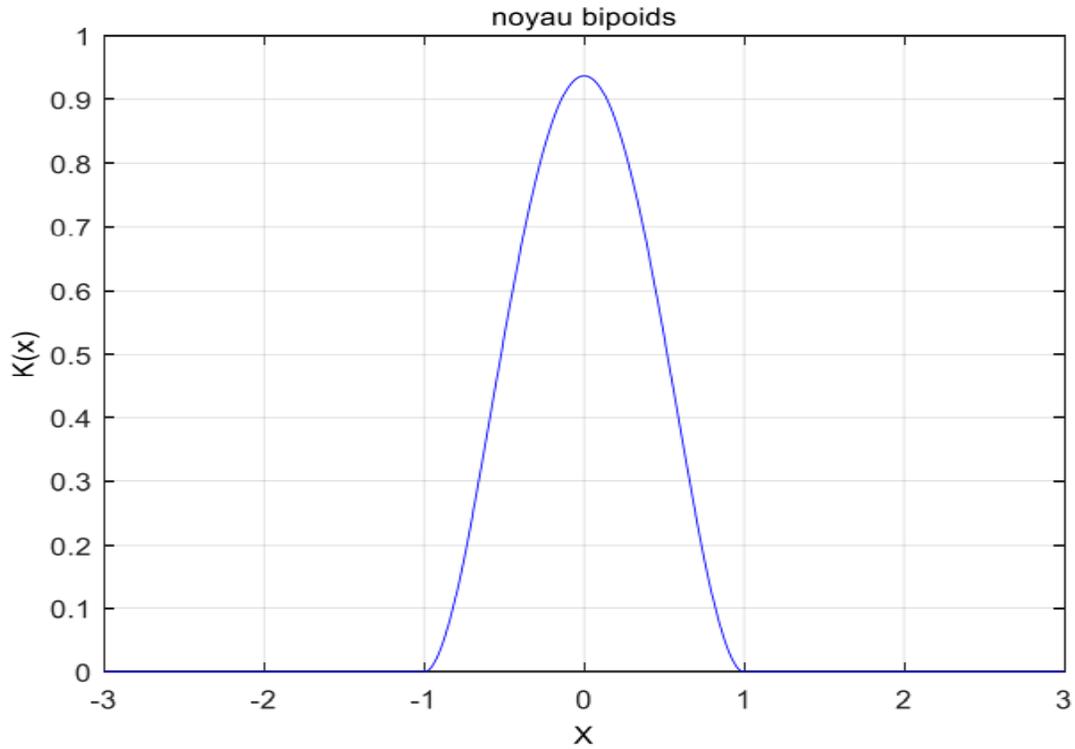


**Figure 18:** exemple de noyau tripoids utilisé pour estimer les fonctions de densité de probabilité.

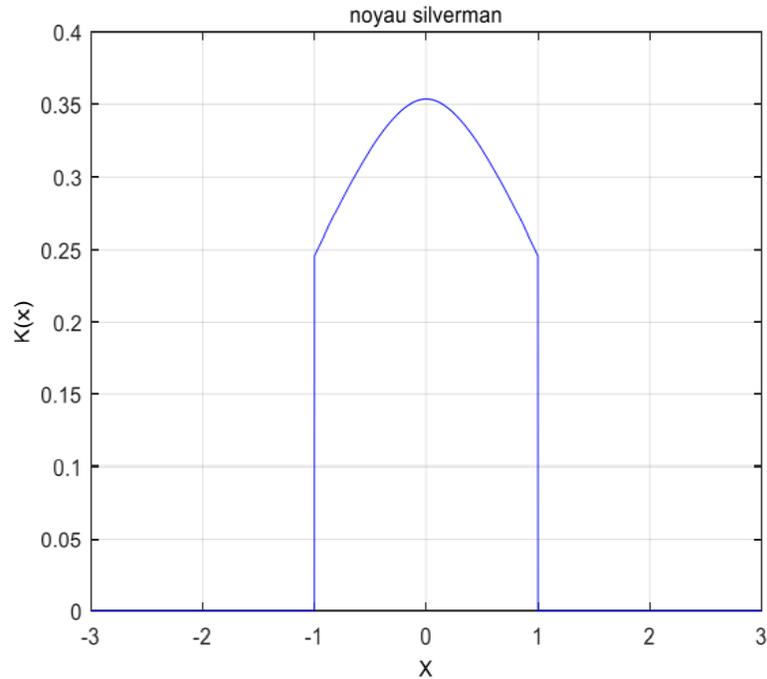
## Chapitre 2 : Estimation de la fonction de densité de probabilité



**Figure 19:** exemple de noyau normal utilisé pour estimer les fonctions de densité de probabilité.



**Figure 20:** exemple de noyau bipoids utilisé pour estimer les fonctions de densité de probabilité.



**Figure 21:** Exemple de noyau silverman utilisé pour estimer les fonctions de densité de probabilité.

### 2.6.7 Choix du paramètre de lissage :

Le paramètre de lissage est le second élément de la méthode d'estimation à noyau. Ce paramètre est indispensable pour la convergence de l'estimateur à noyau et donc l'efficacité du lissage et la qualité de l'estimation. Plusieurs méthodes pour choisir ce paramètre. Ces techniques sont regroupées en deux classes.

- **La première classe :**

La majorité des méthodes de cette classe ont été proposées avant 1990.

- ✓ **Méthode de validation croisée par moindre carrée :**

Rudemo (1982) [16] et Bozman (1984) [17] ont donné l'idée de cette méthode. En utilisant la formule développée de l'erreur quadratique intégrée  $ISE(h)$  on choisit le paramètre de lissage qui minimise cette erreur.

$$ISE(h) = \int [f(x) - f_h(x)]^2 dx = \int f_h^2(x) dx - 2 \int f(x) f_h(x) dx + \int f^2(x) dx \quad (2-19)$$

Le troisième terme de la formule (2-19) ne dépend pas de h, donc on peut choisir h de façon à ce qu'il minimise le critère de la validation croisée.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

### ✓ Méthode de validation croisée biaisée :

Cette méthode a été proposée par Scott et Terrell (1987) [15]. Le paramètre de lissage choisi par cette méthode est la valeur de  $h$  qui minimise un estimateur du critère AMISE :

$$AMISE[f(x), f_h(x)] = \frac{1}{nh} R(g) + \frac{h^4}{4} R(f'') \Sigma^2(k)$$

$R(\cdot)$  : Est une mesure de rugosité défini pour une fonction par :

$$R(g) = \int g^2(x) dx$$

On peut estimer le AMISE si l'on estime ( $F''$ ).

#### • La deuxième classe :

##### ✓ Règle du pouce (Rule of Thumb):

L'idée de cette méthode revient de Deheuvels (1979) [4] Le choix du paramètre de lissage par cette méthode consiste à remplacer la partie inconnue  $R(f'')$  dans l'expression de  $h_{AMISE}$  donné par l'équation  $h_{AMISE} = \left[ \frac{R(K)}{nR(f'')\sigma^2} \right]^{\frac{1}{5}}$ , en prenant pour référence une densité normale de moyenne 0 et de variance  $\sigma^2$ , on a alors :  $R(f''(x)) = \int (f''(x))^2 dx = \frac{3}{8\sqrt{\pi}} \sigma^{-\frac{1}{5}}$ .

De plus, si en utilise un noyau gaussien, alors la valeur de  $h_{AMISE}$  que l'on note dans ce cas  $h_{opt}$  :

$$h_{opt} = (4\pi)^{-\frac{1}{10}} \left[ \frac{3}{8\sqrt{\pi}} \sigma \right] n^{-\frac{1}{5}} = \left( \frac{4}{3} \right)^{\frac{1}{5}} \sigma n^{-\frac{1}{5}} = 1.06 \sigma n^{-\frac{1}{5}} \quad (2- 20)$$

## 2.7 Estimateur à noyau adaptatif :

Cette méthode consiste à prendre en référence la famille des lois normales pour attribuer une valeur à la densité  $f$  et d'en déduire une valeur  $h$  de référence qui va servir à la construction d'un estimateur pilote avec lequel la densité de probabilité sera estimée une première fois. Cette estimée sera ensuite utilisée pour construire un estimateur à largeur de bande variable d'un point à un autre, de façon à utiliser des noyaux plus larges dans les zones à faible densité, ainsi une observation sur la queue aura son effectif exhibé à travers une échelle plus étendue que celle sur la partie principale de la

## Chapitre 2 : Estimation de la fonction de densité de probabilité

distribution, une précision bien utile à notre application vu que les fausses alarmes sont beaucoup plus présentes sur les queues que sur les parties principales [4].

### 2.7.1 Définition d'un estimateur à noyau adaptatif :

L'estimateur à noyau adaptatif opère en deux étapes :

- ✓ Construction d'un estimateur pilote en utilisant une largeur de bande globale obtenue est donnée par (2-20) en considérant une densité normale dans l'expression de  $h_{opt}$ .

L'estimateur pilote est alors donné par :

$$f_h = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_{opt}} K\left(\frac{x - X_i}{h_{opt}}\right) \quad (2-21)$$

- ✓ Définir un facteur de largeur de bande locale donnée par :

$$\lambda_i = \left\{ \frac{f_h(X_i)}{g} \right\}^{-\alpha}$$

Avec :

$g$  : Moyenne géométrique des  $f_h(X_i)$

$\alpha$  : Paramètre de sensibilité qui satisfait :  $0 \leq \alpha \leq 1$

Un choix de  $\alpha = \frac{1}{2}$  donne de bons résultats [10].

Calculer l'estimateur à noyau adaptatif par :

$$f_h = \frac{1}{n} \sum_{i=1}^n \frac{1}{h\lambda_i} K\left(\frac{x - X_i}{h\lambda_i}\right)$$

Divers travaux [4] ont montré l'insensibilité de la méthode au choix de l'estimateur pilote. L'utilisation d'une technique pour optimiser la valeur de  $h$  n'est pas justifiée, elle entrainera une consommation en temps de calcul, sans apport notable sur le résultat final.

Dans cette section de ce chapitre, nous déroulons les résultats relatifs aux différentes techniques d'estimation des fonctions de densité de probabilité dans un contexte comparatif.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

On entamera la première partie, par la présentation des performances de certaines techniques d'estimation de densité de probabilité. Nous allons exposer la suprématie du choix du paramètre de lissage (largeur de bande) sur le choix des noyaux lui-même dans le cas de l'estimateur à noyau, puis des exemples de densité estimées par diverses méthodes.

Les performances des estimateurs (estimateur à noyau fixe, estimateur à adaptatif) seront ensuite examinées et comparées, ce qui permettra le choix de l'estimateur à utiliser.

### 2.8 Simulation et comparaison des résultats :

#### 2.8.1 Introduction :

Dans cette partie nous allons présenter les résultats pour l'estimation non-paramétrique de la densité de probabilité du clutter associée à des critères d'erreurs entre l'estimée et les différents modèles générés. Nous avons choisi d'introduire cette section par exprimer l'importance du choix du paramètre de lissage (la largeur de la bande) sur le choix du noyau lui-même dans le cas d'estimateur à noyau, ensuite des exemples de densités estimées par diverses méthodes seront présentés.

Puis nous proposons une comparaison visuelle et une analyse des performances des estimateurs (estimateur à noyau fixe et à noyau adaptatif) ce qui permettra d'optimiser le choix de l'estimateur à utiliser.

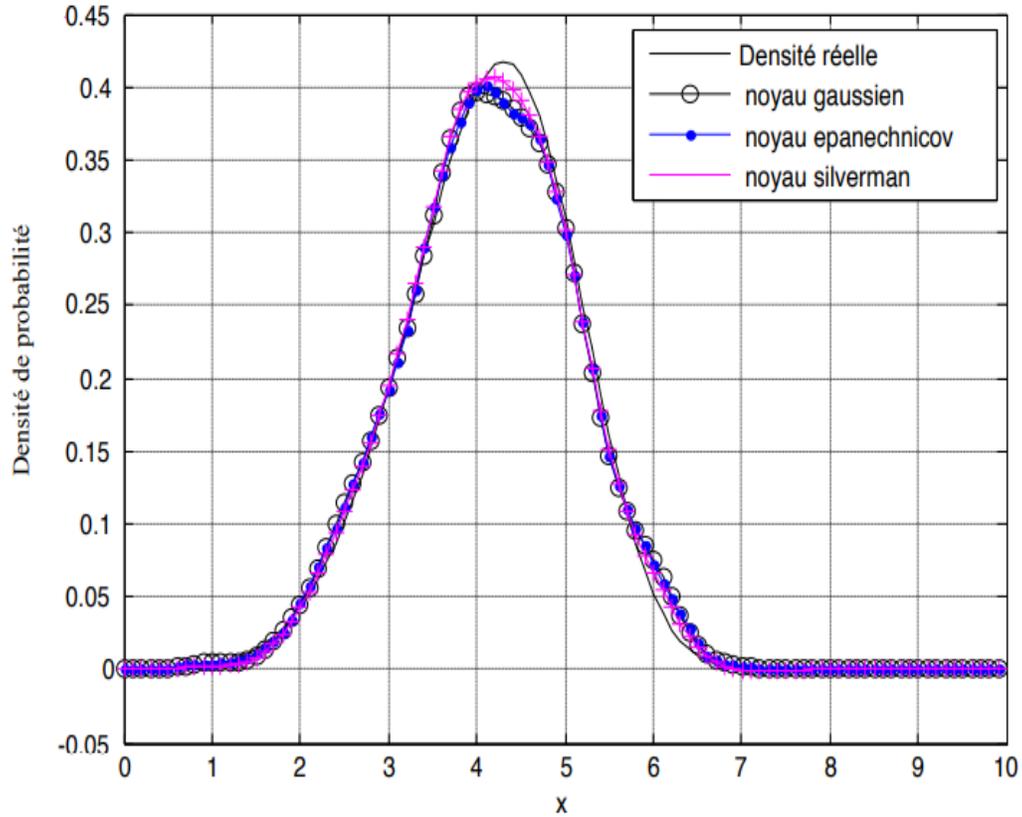
#### 2.8.2 Le choix de langage de programmation :

Notre simulation a été faite par le logiciel MATLAB. Sa particularité réside dans la richesse de sa bibliothèque notamment dans le traitement du signal. Il offre à l'utilisateur un environnement interactif, facile à manipulé et qui nous aide beaucoup dans le processus d'apprentissage. Les étapes de notre procédé de simulation sont:

- ✓ La densité de probabilité est estimée en générant un échantillon aléatoire d'une telle taille suivant une loi de Weibull (l'expression est donnée en (1-6)) ayant le paramètre de forme  $c=4.5$  et le paramètre d'échelle.
- ✓ La densité de probabilité est estimée en utilisant différents noyaux avec la même valeur du paramètre de lissage (valeur donnée en (1-20)).

### 2.8.3 Estimateur à noyau, que choisir le noyau ou sa largeur :

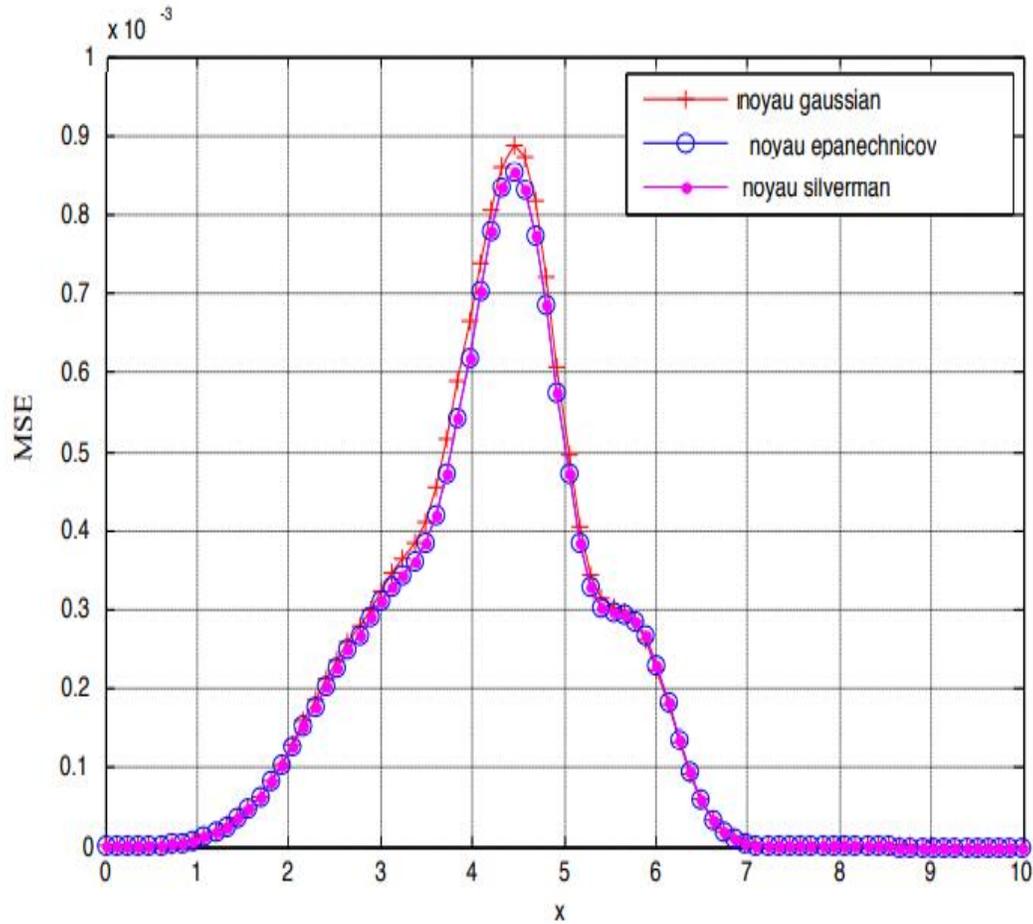
Les résultats de ces estimées ainsi que la densité réelle sont montrées sur la Figure 22



**Figure 22:** Densité de probabilité estimée avec différents noyaux.

On trouve une régularité des résultats montrées sur la Figure 23 représentative de l'erreur quadratique moyenne (critère d'erreur) entre les estimées et la densité réelle.

## Chapitre 2 : Estimation de la fonction de densité de probabilité

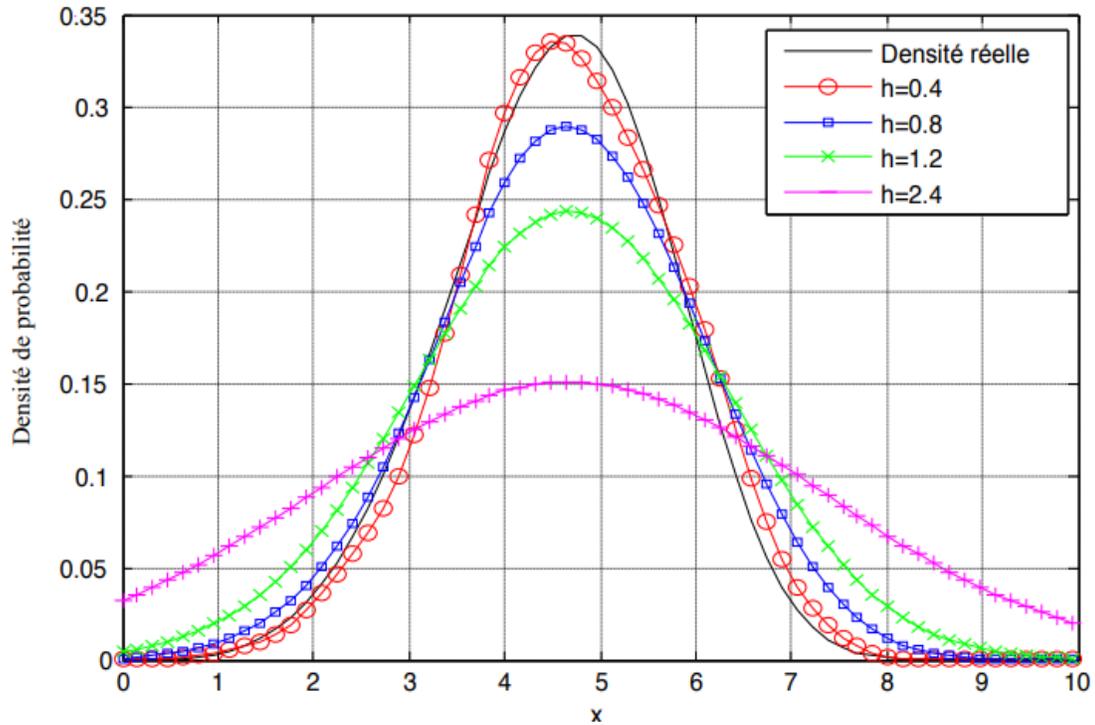


**Figure 23:** Erreur quadratique moyenne pour les différents noyaux.

Ces résultats permettent de conclure l'inutilité d'orienter les travaux à améliorer les performances des estimateurs à noyau vers le choix des largeurs des eux-mêmes.

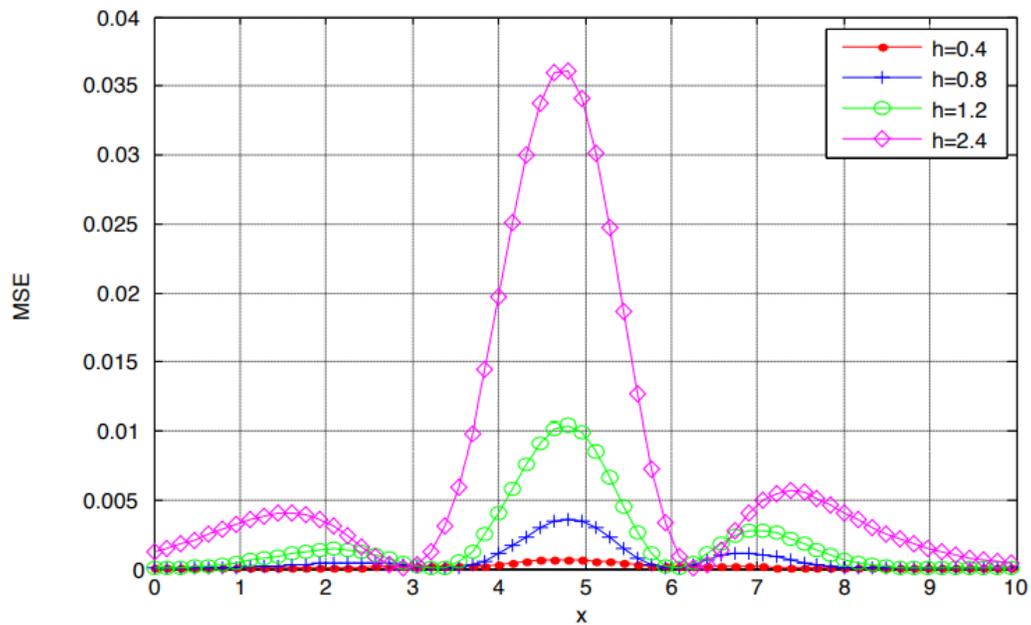
La Figure 24 exprime des fluctuations significatives de l'estimée de la densité de probabilité pour les différentes valeurs du paramètre de lissage et en utilisant le même noyau (noyau gaussien).

## Chapitre 2 : Estimation de la fonction de densité de probabilité



**Figure 24:** Densité de probabilité estimée avec différentes largeurs de bandes.

Les résultats de la mesure de l'erreur quadratique moyenne entre l'estimée et la densité réelle montrés dans la Figure 25 sont relatives à ce constat.



**Figure 25:** Erreur Quadratique moyenne pour les différents paramètres de lissage.

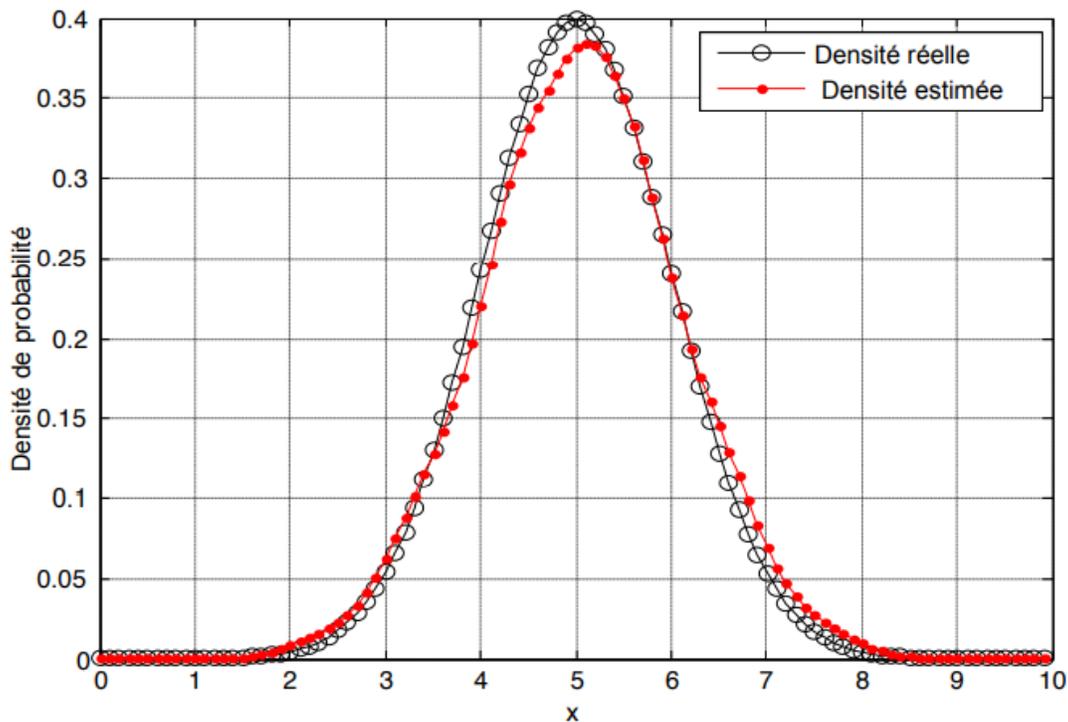
## Chapitre 2 : Estimation de la fonction de densité de probabilité

Ces résultats permettent de conclure qu'il soit préférable de bien choisir le paramètre de lissage au lieu de choisir les noyaux pour avoir un bon estimateur.

### 2.8.4 Exemples de densités estimées par les différents estimateurs :

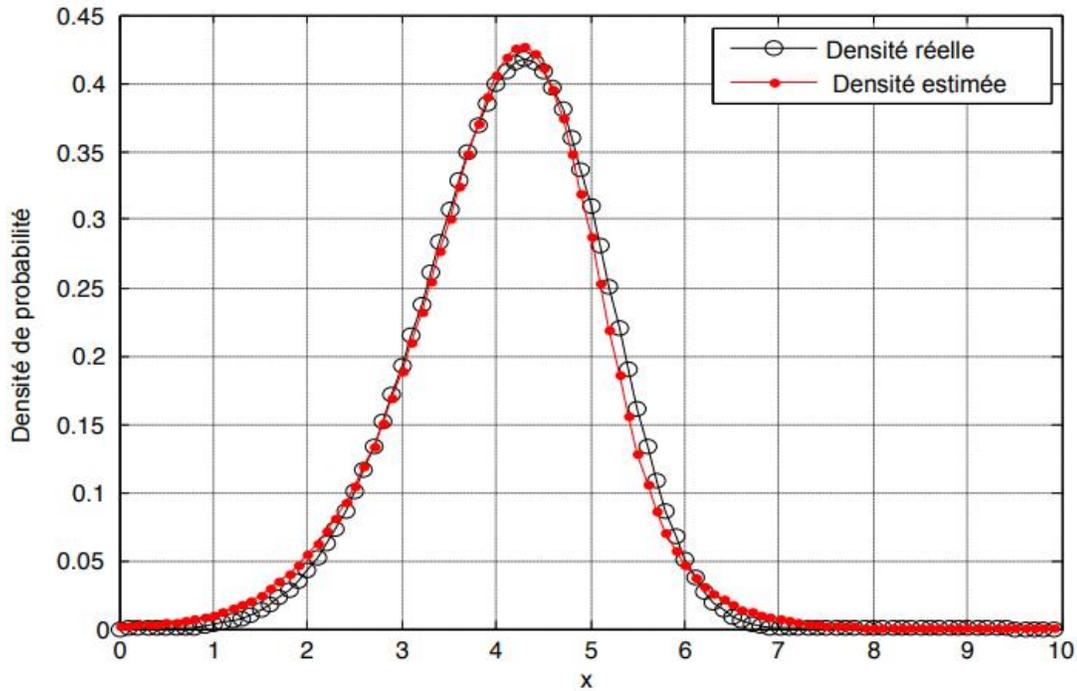
On donne des exemples de densités de probabilité estimées en utilisant les principales méthodes tel que : le noyau fixe et le noyau adaptatif. La densité de probabilité est estimée en utilisant un noyau normal avec une valeur de  $h$  donnée en (2-20).

Les exemples choisis ici sont compatibles aux modèles de clutter présentés dans le premier chapitre. Les résultats obtenus sont présentés dans les Figures 26 et 27.



**Figure 26:** Densité de probabilité Normale ( $m = 0, \sigma^2 = 1$ ) stimée avec un noyau à largeur fixe.

## Chapitre 2 : Estimation de la fonction de densité de probabilité



**Figure 27:** Densité de probabilité Weibull ( $c = 4.5, b = 1$ ) estimée avec un noyau à largeur adaptative.

Les résultats présentés par ces exemples ne sont pas juger les performances des estimateurs.

### 2.8.5 Comparaison des performances des estimateurs :

Après avoir la primauté du choix du paramètre de lissage sur celui du noyau, et après garantir le meilleur choix de l'estimateur qui sera exploité dans le schéma final du détecteur, nous mettons au cours d'exprimer une comparaison visuelle des performances des deux estimateurs considérés (estimateur à noyau fixe et à noyau adaptatif) en fonction de la taille de l'échantillon.

Les étapes de notre procédé de simulation sont :

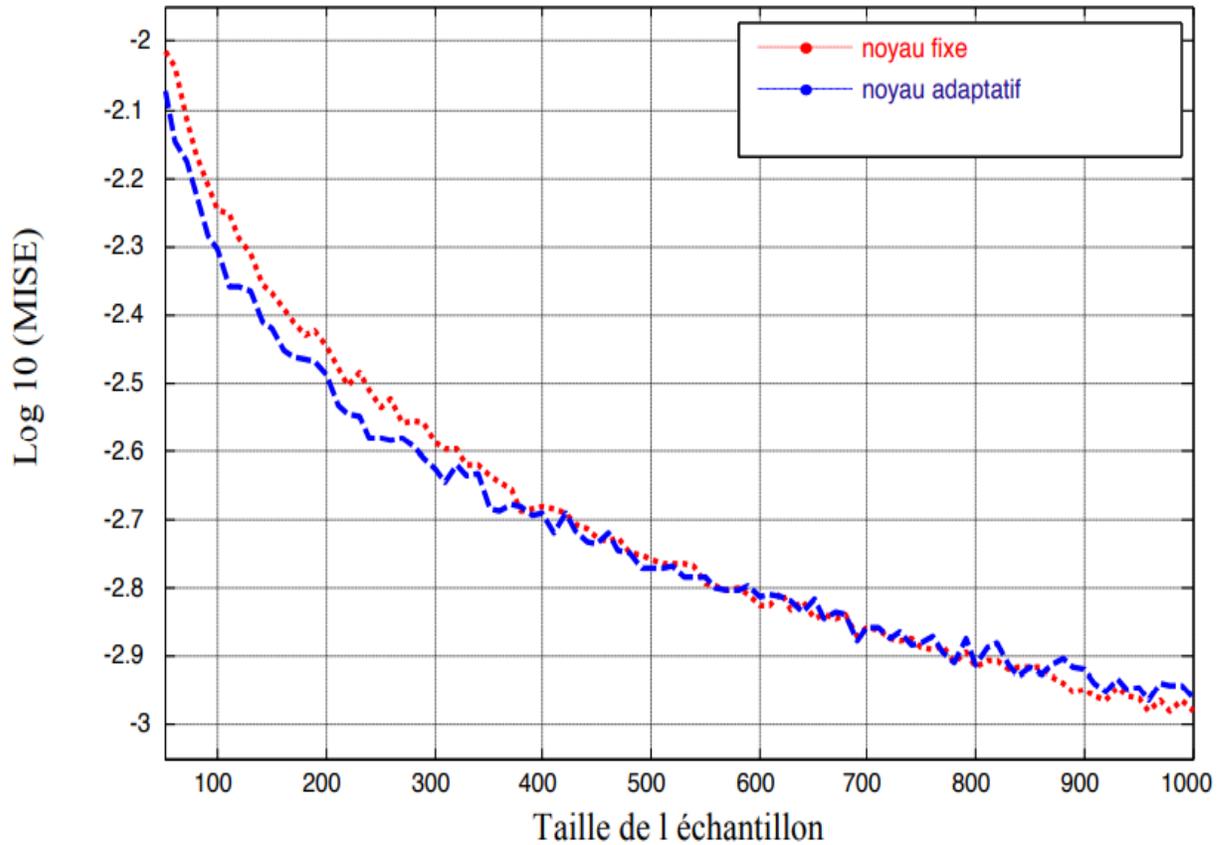
- La densité de probabilité est estimée en générant un échantillon aléatoire d'une taille égale à 1000 utilisant les lois suivantes :
  - La loi Normale (l'expression est donnée en (1-4)) de moyenne  $m = 5$  et de variance  $\sigma^2 = 1; (N(5,1))$  ;
  - La loi de Weibull (l'expression est donnée en (1-6)) ayant les paramètres de forme  $c = 4.5$  et le paramètre d'échelle  $b = 5$ .
  - La loi de Log-normale (l'expression est donnée en (1-5)) ayant les paramètres de forme  $c = 0.3$  et le paramètre d'échelle  $m = 1.2$

## Chapitre 2 : Estimation de la fonction de densité de probabilité

- La densité de probabilité est estimée en utilisant différents noyaux avec la même valeur du paramètre de lissage (valeur donnée en (2-20)).

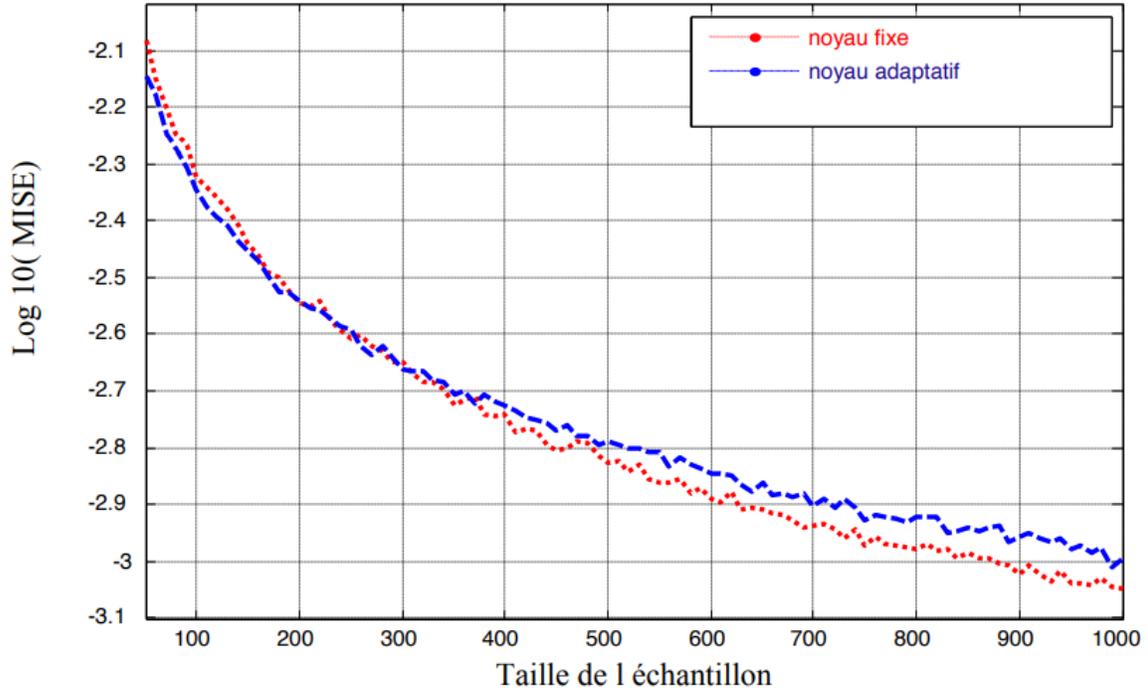
Un critère d'évaluation montré dans [4] s'appelle l'erreur quadratique moyenne intégrée (MISE), son expression est donnée par (2-4).

Les résultats obtenus sont présentés dans les Figures 28, 29, et 30

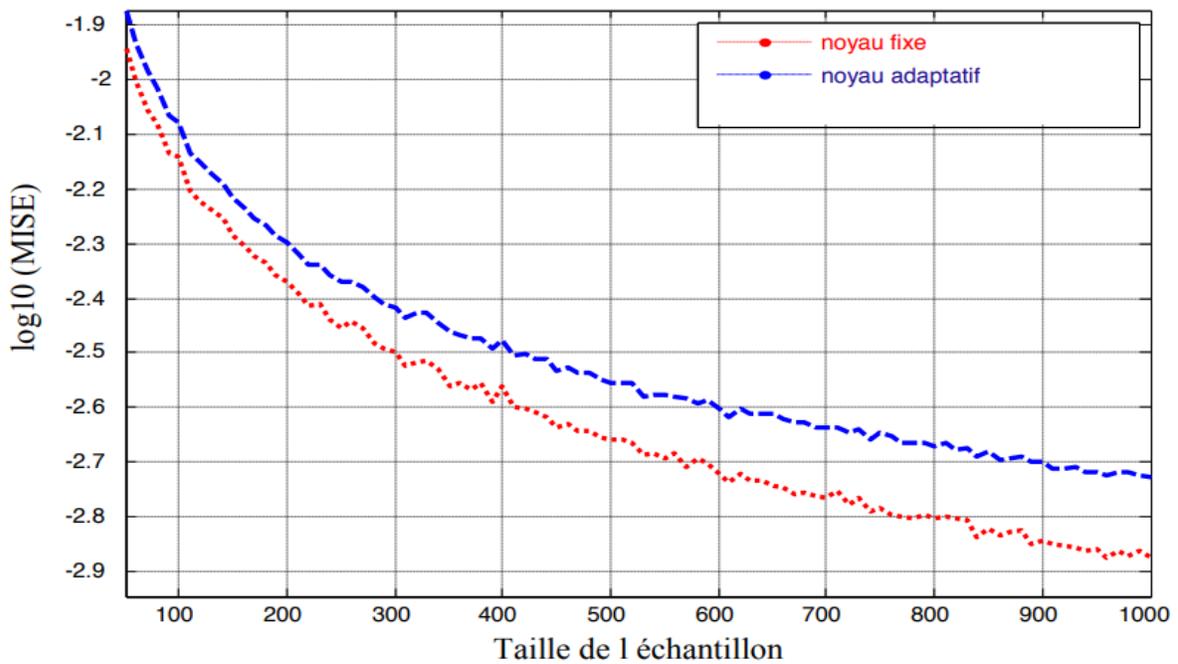


**Figure 28:** Variations de la MISE en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Normale ( $N(5,1)$ ).

## Chapitre 2 : Estimation de la fonction de densité de probabilité



**Figure 29 :** Variations de la MISE en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité de Weibull ( $c=5$ ,  $b=5$ ).



**Figure 30:** Variations de la MISE en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Log-normale ( $m=1.2$ ,  $c=0.3$ ).

## Chapitre 2 : Estimation de la fonction de densité de probabilité

### Discussions des résultats :

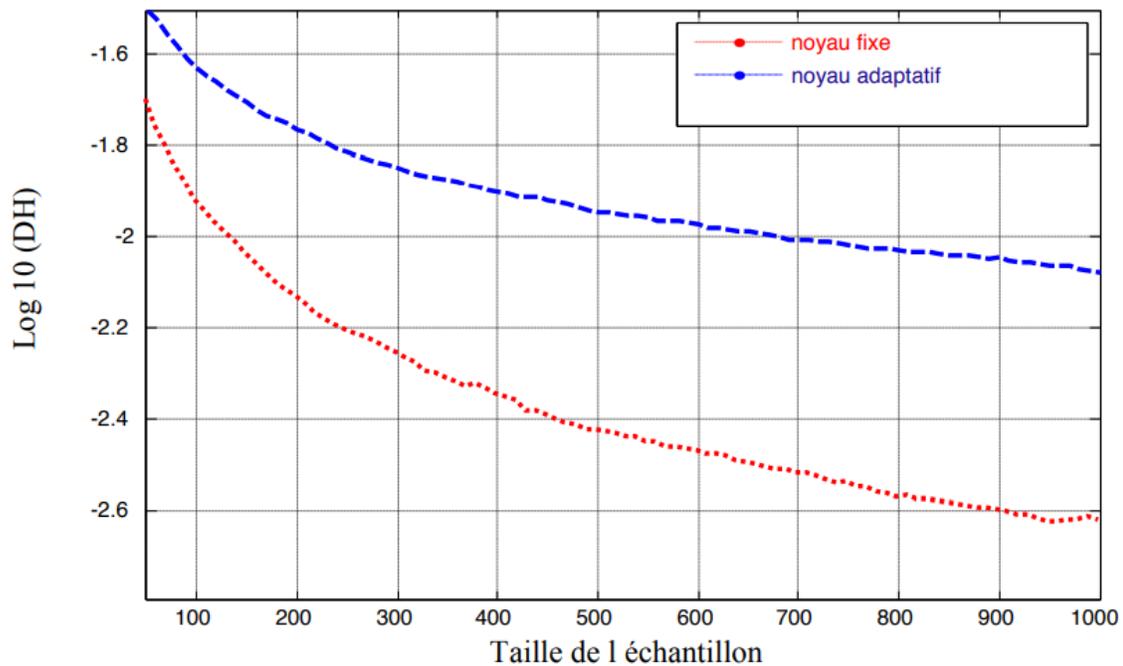
Concernant les courbes relatives aux variations des erreurs quadratiques moyennes intégrées (MISE) :

Le cas de l'estimation d'une densité de probabilité Normale, l'estimateur par noyau adaptatif est plus performant pour des échantillons de taille inférieure à 550. Pour des échantillons de taille supérieure à 550 les deux estimateurs tendent à avoir les mêmes performances.

Le cas de l'estimation d'une densité de probabilité Weibull, les deux estimateurs avoir les mêmes performances pour des échantillons de taille inférieure à 350 et l'estimateur à noyau fixe plus performant pour des échantillons de taille supérieure à 350.

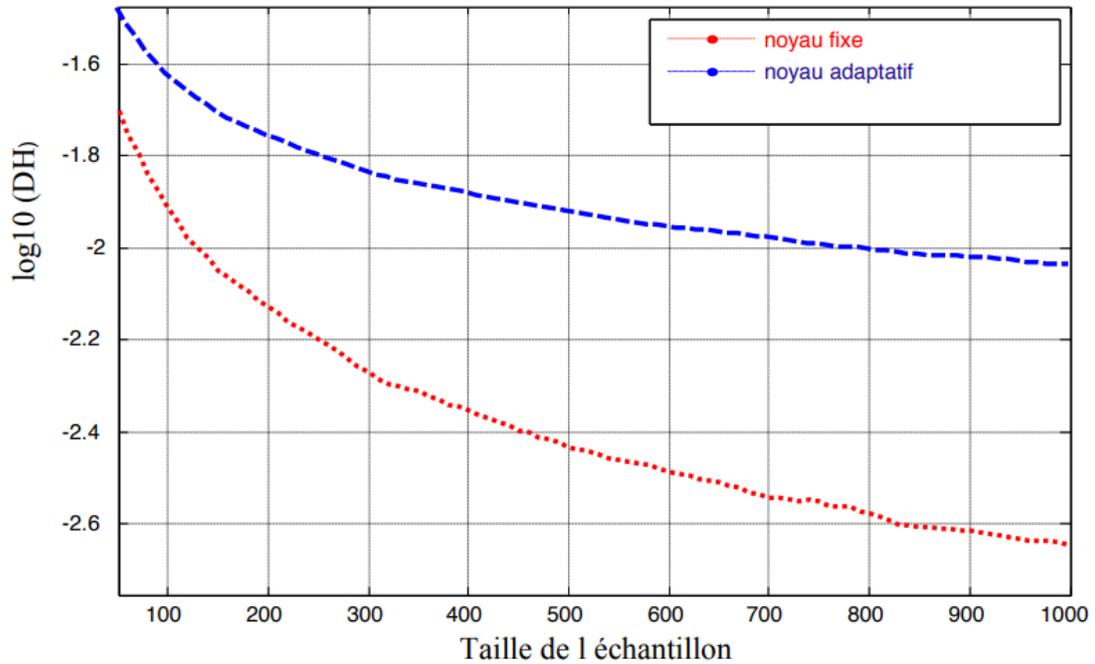
Le cas de l'estimation d'une densité de probabilité de Log-normale, l'estimateur par noyau fixe est plus performant que le noyau adaptatif quel que soit la taille de l'échantillon utilisé.

Ensuite nous utilisons une mesure d'écart montré et son expression est donnée par II.6, s'appelle distance de Hellinger en fonction de la taille de l'échantillon. Les résultats obtenus sont présentés dans les Figures 31, 32 et 33.

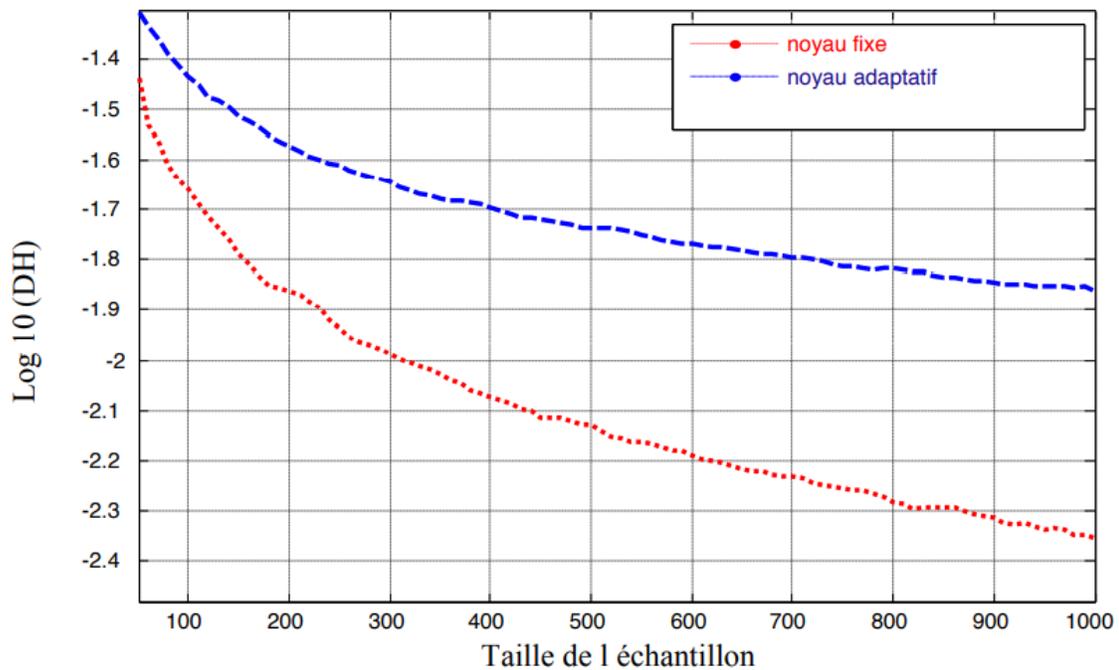


**Figure 31:** Variations de la distance de Hellinger en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Normale (N(5,1)).

## Chapitre 2 : Estimation de la fonction de densité de probabilité



**Figure 32:** variations de la distance de Hellinger en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Weibull ( $c=4.5$ ,  $b=5$ ).



**Figure 33:** Variations de la distance de Hellinger en fonction de la taille de l'échantillon dans le cas de l'estimation d'une densité de probabilité Log-normale ( $m=1.2$ ,  $c=0.3$ ).

## Chapitre 2 : Estimation de la fonction de densité de probabilité

### Discussions des résultats :

Les courbes relatives aux variations de la distance de Hellinger montrent que l'estimateur noyau fixe est soit le plus performant que l'estimateur à noyau adaptatif.

### Conclusion :

Le travail autour du deuxième chapitre s'articule aux méthodes non paramétriques d'estimation de la fonction de densité de probabilité, plus les critères d'erreurs entre fonctions. Puis nous introduisons les techniques d'estimations des plus simples aux plus évoluées. La finalité de ce chapitre, est consacrée aux résultats, ces derniers s'articulent en deux sections : Dans la première section : des résultats visuelles d'estimation de fonctions de densité de probabilité à travers des exemples de densité en utilisant les différents estimateurs. Dans la deuxième section : une étude comparative de ces estimateurs afin de choisir le meilleur.

Dans le chapitre suivant, nous présentons les principales techniques de détection radar existants ensuite nous analysons les performances des détecteurs CA, GO, SO, et l'OS dans différentes situations d'environnement de fonctionnement radar.

### Chapitre 3 : Technique de détection Radar :

#### Introduction :

Dans ce chapitre, nous présentons les principales techniques de détection radar existantes, en commençant par la détection classique à seuil fixe puis le principe de la détection adaptative. Ensuite, nous analysons les performances des détecteurs CA, GO, SO, et l'OS dans différentes situations d'environnement de fonctionnement radar; notamment dans un milieu homogène et en présence de cibles interférentes. La finalité de ce chapitre est consacrée à la proposition d'une architecture d'exploitation des techniques d'estimation de la PDF à la détection CFAR.

#### 3.1 Définition de détection :

La détection est définie comme étant l'opération qui consiste à prendre une décision sur la présence ou l'absence de cibles. Le principe de base de la détection d'une cible est de comparer le signal reçu noyé dans un clutter, après quelques traitements, à un seuil calculé a priori. Si le signal composite (signal plus clutter) dépasse ce seuil, la présence de la cible est déclarée. Dans le cas contraire, l'absence de cible est déclarée. La détection radar classique est basée sur un seuillage fixe et nécessite un opérateur radar expérimenté pour la prise de la décision. La non stationnarité de l'environnement fait qu'une petite augmentation dans la puissance du bruit conduit à augmenter sensiblement la probabilité de fausse alarme, ce qui peut provoquer non seulement la saturation de l'opérateur du PPI (écran de visualisation), mais aussi celle du système radar lui-même. Les chercheurs dans le domaine de détection ont pensé à une détection adaptative, c'est-à-dire à seuillage adapté au niveau du bruit et des signaux indésirables tout en maintenant une probabilité de fausse alarme constante

#### 3.2 Rappels sur la théorie de la détection radar :

Les systèmes radar utilisent une antenne pour détecter des cibles dans l'espace autour de cette antenne. Pour chaque cible, on désire savoir sa position, sa direction et sa vitesse. On considère ici le cas simple où on détermine uniquement la position de la cible. Le modèle classiquement utilisé en détection radar est le suivant [6] :

$$Y(t) = k * s(t - t_0) + C(t) \quad (3- 1)$$

$$t_0 = \frac{2R}{c} \quad (3- 2)$$

Avec :

$Y(t)$  : Le signal reçu ;

$C(t)$  : Le bruit ;

$t_0$  : Le retard et  $k$  un paramètre qui dépend de l'altitude de la cible, de ses propriétés de réflexion, et de la distance ;

$c$  : est la vitesse de propagation des ondes électromagnétiques.

Pour détecter la cible, le système doit donc résoudre un problème de décision entre les deux hypothèses suivantes :

- ✓  $H_0: Y(t) = C(t)$  : Hypothèse indiquant que le signal est dû au bruit seulement ;
- ✓  $H_1: Y(t) = s(t) + C(t)$  : Hypothèse indiquant que le signal est dû à un écho plus le bruit.

A cause de la fluctuation des cibles et la nature aléatoire du clutter, le problème de la détection est étudié comme un processus stochastique.

### 3.2.1 Critères de décision :

#### 3.2.1.1 Tests de Bayes :

L'utilisation du critère de Bayes nécessite principalement deux hypothèses  $P(H_1)$  et  $P(H_0)$  Le but de ce critère est de déterminer une règle de décision qui minimise le cout moyen.

La règle de Bayes se résume par comparaison du rapport de vraisemblance  $\Lambda (y)$  à un seuil  $\eta$ :

$$\Lambda (y) = \frac{p(y|H_1)}{p(y|H_0)} \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \eta = \frac{p(H_0)(C_{10}-C_{00})}{p(H_1)(C_{01}-C_{11})} \quad (3- 3)$$

Avec :

$\Lambda (y)$ : est le rapport de vraisemblance;

$p(y|H_1)$  : est la fonction de densité de probabilité de la cellule sous test sachant que le signal « s » est présent (cible présente);

$p(y|H_0)$  : est la fonction de densité de probabilité de la cellule sous test sachant que le signal « s » est absent (cible absente);

$p(H_1)$  et  $p(H_0)$  sont la probabilité a priori que la cible est présente et la probabilité a priori que la cible est absente, respectivement;

$C_{ij}; i, j = 0, 1$  Coût associé à la décision  $D_i$  sachant que l'hypothèse  $H_j$  est vraie ( $D_i/H_j$ )

### 3.2.1.2 Tests de Neyman-Pearson :

Pour construire le test de Bayes, c'est-à-dire, le coût moyen d'une décision, il faut connaître les probabilités à priori  $P_i$ , Pour beaucoup d'applications, c'est une contrainte de connaître ces valeurs. Dans pareils cas on utilise la probabilité de détection  $p_d$  et la probabilité de fausse alarme  $P_{fa}$ .

Les tests de Neyman-Pearson (N-P) agissent directement sur les mesures de performance

$$p_d = p_R\left\{ \text{décider } \frac{D_1}{H_1} \right\}$$

$$P_{fa} = p_R\left\{ \text{décider } \frac{D_1}{H_0} \right\}$$

$$P_m = p_R\left\{ \text{décider } \frac{D_0}{H_1} \right\}$$

$P_m$  Est la probabilité de non détection (miss). On veut, usuellement, avoir une valeur de  $p_d$  le plus grand possible, et en même temps, une valeur de  $P_{fa}$  la plus petite possible.

Le critère de Neyman-Pearson vise à maximiser la probabilité de détection  $p_d$  (Décider  $H_1$  lorsque  $H_1$  est vraie, avec la probabilité  $p_d = \left(\frac{H_1}{H_1}\right) = 1 - p_{nd}$ ) pour un taux de fausses alarmes  $P_{fa} = \alpha$  fixée (dans la pratique la plus faible possible).

$L(\lambda)$  : Le test optimal pour le critère de Neyman-Person. Cela s'obtient facilement par la méthode des multiplicateurs de Lagrange en maximisant la fonction suivante :

$$L = p_d + \lambda(P_{fa} - \alpha) \tag{3- 4}$$

Dans ce test, le seuil de détection, est déterminé pour avoir  $P_{fa} = \alpha$  fixée, et se calcule en résolvant l'une des deux équations suivantes :

$$P_{fa} = P(\Lambda(y; H_0) > \eta) = \alpha$$

$$P_{fa} = \int P(y/H_0) dy \quad [D_1; \infty] \quad (3-5)$$

$D_0$  : est l'ensemble des  $y$  contenus dans le domaine de décision de l'hypothèse  $H_0$  . Il est souvent difficile d'obtenir une expression analytique de ces équations, et le calcul de  $\eta$  s'effectue soit en résolvant numériquement (3-6) ou (3-7), soit par Monte-Carlo.

Une fois le seuil de détection déterminé, le calcul de  $p_d$  peut s'effectuer de deux manières différentes :

$$p_d = P(\Lambda(y; H_1) > \eta) = \alpha \quad (3-6)$$

$$P_d = \int P(y/H_1) dy \quad [D_1; \infty] \quad (3-7)$$

$D_1$  : Est l'ensemble des  $y$  contenus dans le domaine de décision de l'hypothèse  $H_1$ .

### 3.3 Détection à seuil fixe :

La sortie du récepteur est considérée comme un processus aléatoire dont les échantillons sont caractérisés par une fonction de densité de probabilité (PDF). L'existence du bruit parasite rend impossible la reconnaissance du signal utile sans risque d'erreur. C'est pourquoi les problèmes liés à la détection d'un signal noyé dans un bruit relèvent des problématiques de la détection. La détection est définie comme étant le test de deux hypothèses  $H_0$  et  $H_1$  :

$H_0$  : représente l'hypothèse de l'existence du bruit seul.

$H_1$  : représente l'hypothèse de l'existence du signal utile en présence du bruit.

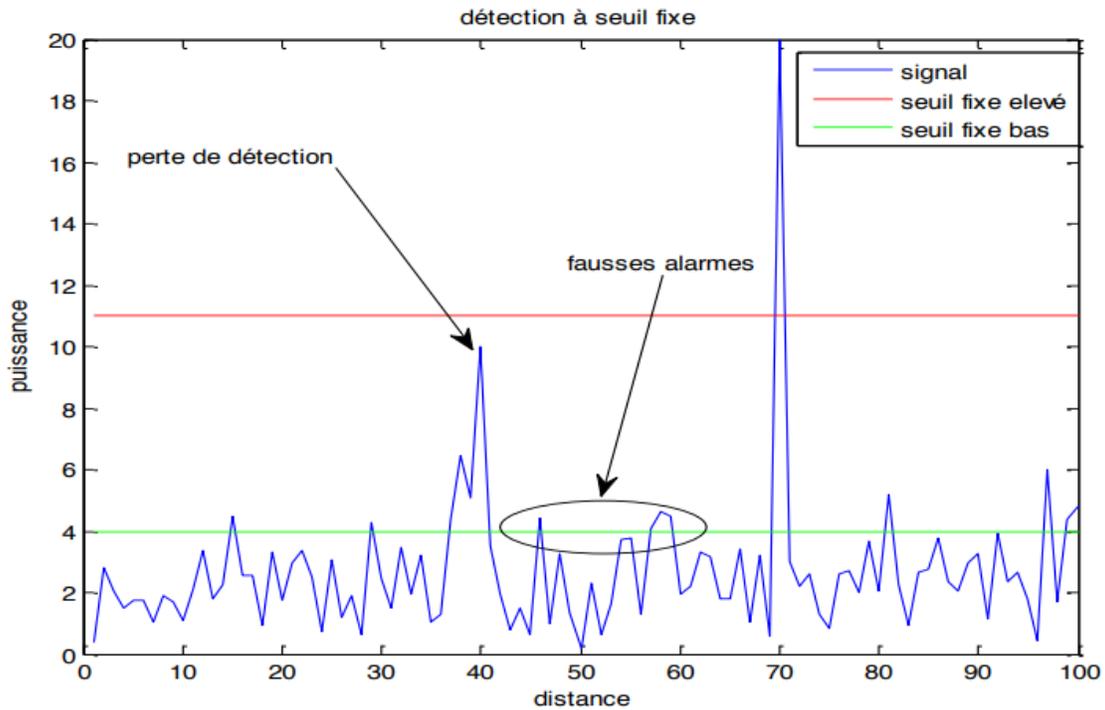
1) Lorsque  $H_0$  est vraie et  $H_1$  est reconnue ; l'erreur est appelé « fausse alarme ».

2) Lorsque  $H_1$  est vraie et  $H_0$  est reconnue ; l'erreur est appelée « perte de détection » (missed detection).

Le tableau 4 représente quatre situations possibles.

	vraie	$H_0$	$H_1$
reconnue			
$H_0$		Pas d'erreur	perte détection
$H_1$		Fausse alarme	Pas d'erreur

**Tableau 4:** Détection et test de deux hypothèses  $H_0$  et  $H_1$ .



**Figure 34:** Principe de détection avec un seuil fixe.

Un seuil trop haut entraînera une perte la détection, et un seuil trop bas un taux de fausse alarme trop élevé

### 3.4 Détection à seuil adaptatif CFAR :

Dans un radar, le bruit accompagnant le signal utile à la sortie est soumis à des variations de niveau pour de nombreuses raisons. Ces variations ont pour effet de faire varier la probabilité de fausse alarme  $P_{fa}$ .

Pour un seuil fixe  $Q_f$ , la  $P_{fa}$  est donnée par la relation suivante:

$$P_{fa} = \exp\left(\frac{-Q_f^2}{2\sigma^2}\right) \quad (3- 8)$$

Soit  $P_{fa\ nom}$  la probabilité de fausse alarme nominale correspondant à une puissance du bruit  $\sigma$ .

$$P_{fa\ nom} = \exp\left(\frac{-Q_f^2}{2\sigma_{nom}^2}\right) \quad (3- 9)$$

Des équations (3-8) et (3-9), on aura:

$$P_{fa} = (P_{fa\ nom}) \frac{\sigma_{nom}^2}{\sigma^2} \quad (3- 10)$$

### 3.4.1 Analyse du détecteur CA-CFAR :

Les détecteurs CFAR estiment le niveau du clutter à partir de la moyenne arithmétique des cellules de référence. Autrement dit, afin de maintenir une probabilité de fausse alarme désirée, le calcul du seuil adaptatif est obtenu par la multiplication du facteur d'échelle par le niveau moyen du clutter dans les cellules. Ces types de détecteurs sont conçus pour donner de bonnes estimations de la puissance du clutter, dont la fonction de densité de probabilité est de type exponentielle et Rayleigh.

Le détecteur CA-CFAR (Cell Averaging CFAR), proposé par H.M Finn and R.S Johnson [18] en 1968, prend, comme son nom l'indique, la moyenne des échantillons dans la fenêtre de référence comme estimation du niveau de bruit dans cette fenêtre. Ce détecteur est destiné à travailler dans un environnement homogène, car ces performances en termes de détection se dégradent dans un environnement à cibles multiples qui présente un niveau d'interférence important.

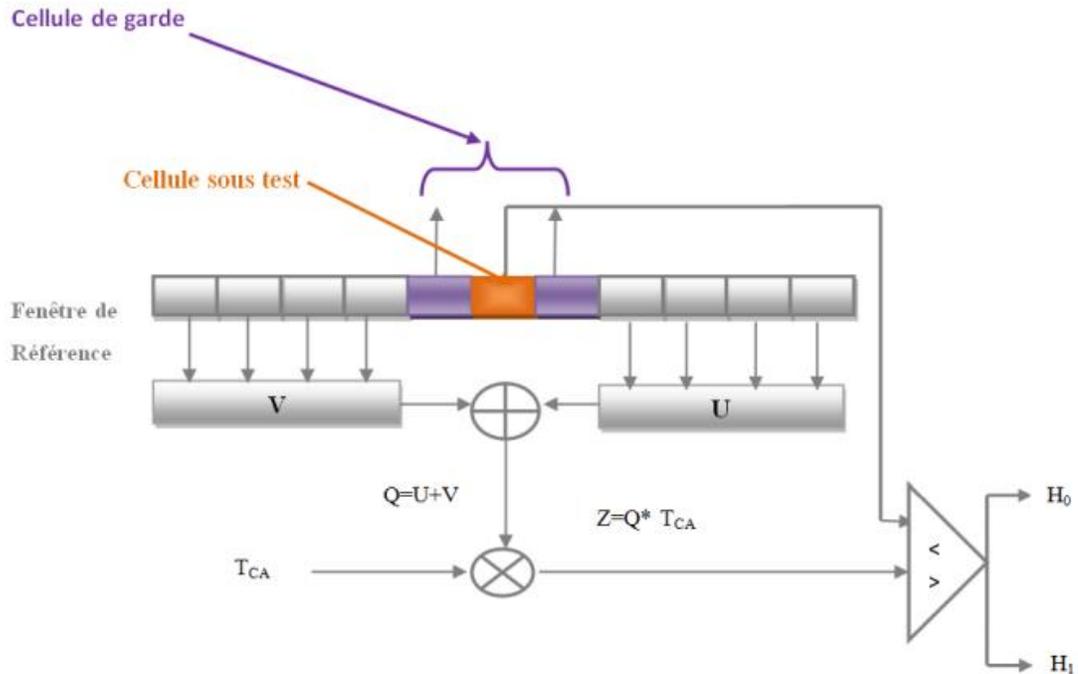


Figure 35: Schéma fonctionnel du détecteur CA-CFAR.

Une fenêtre de référence de  $N$  cellules est subdivisée en deux sous fenêtres de  $N/2$  cellules situées de part et d'autre de la cellule sous test. Le calcul de seuil adaptatif est obtenu par la somme arithmétique des cellules références multipliée par un facteur d'échelle  $T$ , afin de maintenir la probabilité de fausse alarme  $P_{fa}$  constante. Les cellules de garde, adjacentes à la cellule de test, sont exclues de ce calcul car elles peuvent contenir une partie de la puissance de la cible.

La décision est une décision partielle, elle est effectuée par la comparaison du signal provenant de la cellule sous test et le seuil adaptatif calculé. Si l'écho de la cellule sous test dépasse ce seuil, une cible est déclarée présente dans l'espace d'observation. La Figure 30 montre l'architecture de détecteur CA-CFAR.

➤ **Etude probabiliste de CA-CFAR :**

Pour déterminer la probabilité de détection et la probabilité de fausse alarme d'une seule cible dans le détecteur CA-CFAR, nous supposons que :

Une cible primaire est présente dans la cellule sous test (hypothèse  $H_1$  ), l'expression de la probabilité de détection du CA-CFAR :

$$P_d = [1 + \frac{T}{1+S}]^{-N} \quad (3- 11)$$

En posant S=0 dans (3-12), nous déduisons l'expression de la probabilité de fausse alarme :

$$P_{fa} = [1 + T]^{-N} \quad (3- 12)$$

L'expression du facteur d'échelle T est tirée directement de l'expression (3-12) comme suit :

$$T = (P_{fa})^{-\frac{1}{N}} - 1 \quad (3- 13)$$

Si le nombre N des cellules distance de la cellule de référence devient grand, les probabilités de détection et de fausse alarme deviennent  $N \rightarrow \infty$  [18].

$$P_d = \exp(-\frac{T}{1+S}) \quad (3- 14)$$

$$P_{fa} = \exp(-T) \quad (3- 15)$$

Les équations (3-14) et (3-15) sont les expressions qui décrivent les performances du détecteur optimum de Neyman-Pearson. Ainsi, pour un environnement homogène, le détecteur CA-CFAR est le détecteur optimum, vu que sa probabilité de détection se rapproche de celle du détecteur de Neyman-Pearson, dans le cas où N est très grand.

À partir de (3-14), on peut tirer l'expression de  $P_d$  en fonction du  $P_{fa}$  du détecteur de Neyman-Pearson :

$$P_d = P_{fa}^{\frac{1}{1+S}} \quad (3- 16)$$

### 3.4.2 Analyse du détecteur GO-CFAR :

Le GO-CFAR (Greatest Of CFAR) est le premier type de détecteur CFAR modifié. Il a été proposé par Hansen et Sawyer [19].

L'estimation du niveau du bruit est le maximum des sommes des sorties des deux sous fenêtres U et V. Pour ce détecteur, Q est donnée par [19] :

$$Q = \max (U, V) \quad (3- 17)$$

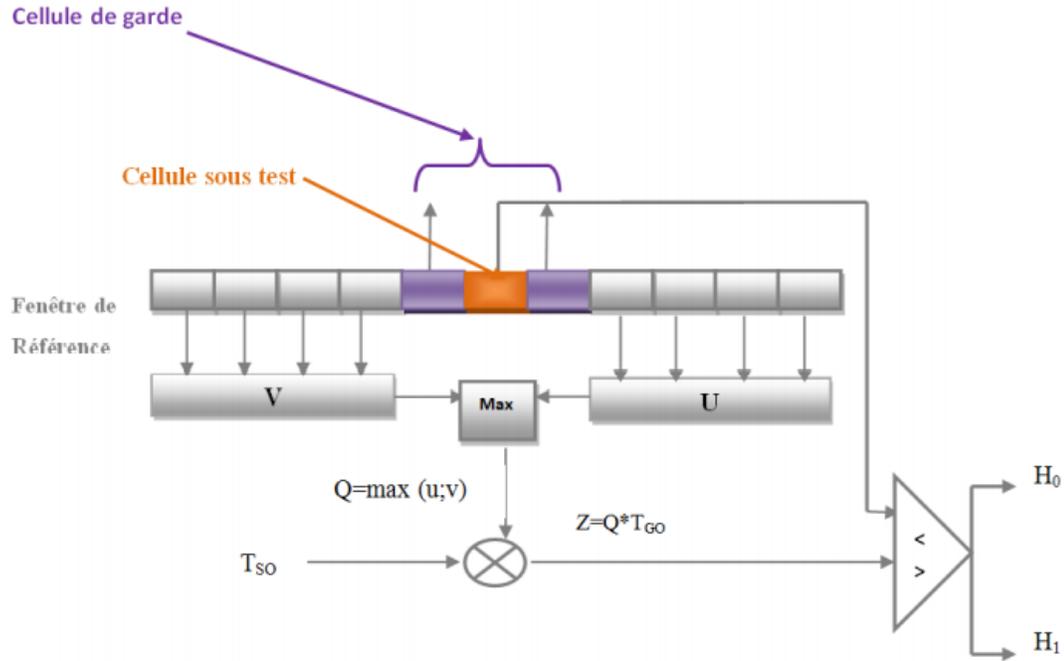


Figure 36: Schéma fonctionnel du détecteur GO-CFAR.

➤ **Etude probabiliste de GO-CFAR :**

$$P_d = 2\left(1 + \frac{T_{GO}}{1+S}\right)^{-N/2} - 2 \sum_{j=0}^{\frac{N}{2}-1} \frac{\left(\frac{N}{2}-1+j\right)!}{\left(\frac{N}{2}-1\right)!j!} \left(2 + \frac{T_{GO}}{1+S}\right)^{-\left(\frac{N}{2}+j\right)} \quad (3-18)$$

Avec :

$S$  : Le rapport signal sur bruit ;

$T_{GO}$  : Le facteur d'échelle du détecteur GO-CFAR ;

$N$  : Nombre de cellules de référence du détecteur GO-CFAR.

Pour trouver la probabilité de fausse alarme, il suffit de poser  $S = 0$  dans la formule (3-18) de  $P_d$ , nous obtenons :

$$P_{fa} = 2\left(1 + T_{GO}\right)^{-N/2} - 2 \sum_{j=0}^{\frac{N}{2}-1} \frac{\left(\frac{N}{2}-1+j\right)!}{\left(\frac{N}{2}-1\right)!j!} \left(2 + T_{GO}\right)^{-\left(\frac{N}{2}+j\right)} \quad (3-19)$$

### 3.4.3 Analyse du détecteur SO-CFAR :

Le **SO-CFAR** (Smallest Of CFAR) est un autre type modifié de détecteur **CFAR** [20] l'estimation du niveau du bruit est le minimum des sommes des sorties des deux fenêtres U et V. Pour ce détecteur, Q est donnée par :

$$Q = \min(U, V) \quad (3- 20)$$

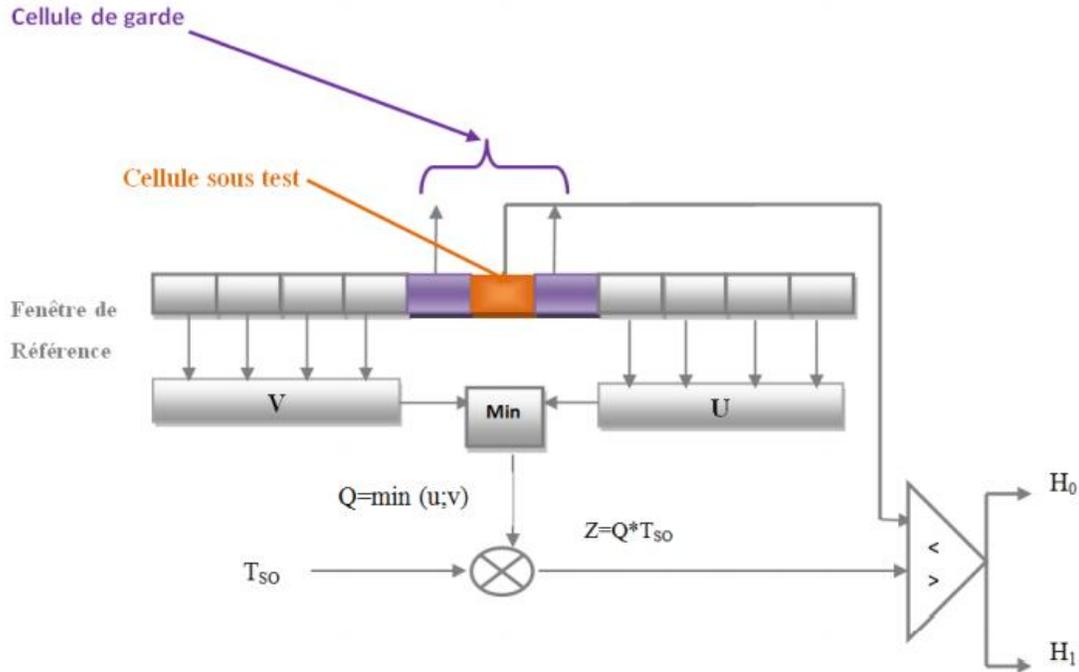


Figure 37: Schéma fonctionnel du détecteur SO-CFAR.

➤ **Etude probabiliste de SO-CFAR :**

La probabilité de détection  $P_d$  d'un détecteur **SO-CFAR** est donnée par l'expression suivante [20] ;

$$P_d = 2 \sum_{j=0}^{\frac{N}{2}-1} \frac{\left(\frac{N}{2}-1+j\right)!}{\left(\frac{N}{2}-1\right)!j!} \left(2 + \frac{T_{SO}}{1+S}\right)^{-\left(\frac{N}{2}+j\right)} \quad (3- 21)$$

Pour trouver la probabilité de fausse alarme  $P_{fa}$ , il suffit de poser  $S=0$  dans la relation (3-21) de  $P_d$ , on obtient :

$$P_{fa} = 2 \sum_{j=0}^{\frac{N}{2}-1} \frac{\left(\frac{N}{2}-1+j\right)!}{\left(\frac{N}{2}-1\right)!j!} \left(2 + T_{SO}\right)^{-\left(\frac{N}{2}+j\right)} \quad (3- 22)$$

### 3.5 Comparaison des performances des détecteurs CA, GO et SO-CFAR :

Dans cette section, les performances des détecteurs **CA**, **GO** et **SO-CFAR** seront étudiées dans différents environnements. La comparaison est basée sur le tracé de la probabilité de détection en fonction du rapport signal sur bruit en utilisant la méthode Monte-Carlo. Les simulations sont faites avec 16 cellules de référence et une probabilité de fausse alarme de  $10^{-5}$ .

Trois cas d'environnements sont envisagés: cas d'absence des cibles interférentes, cas de bord de clutter et cas de présence des cibles interférentes.

#### 3.5.1 Cas d'absence de cibles interférentes :

Nous avons étudié les performances du **CA**, **GO** et **SO-CFAR** à l'aide du logiciel de simulation MATLAB, en générant un signal radar aléatoire simulant l'environnement dans le cas d'absence de cibles interférentes, puis appliquer les différents types de détecteur : **CA**, **GO** et **SO-CFAR**

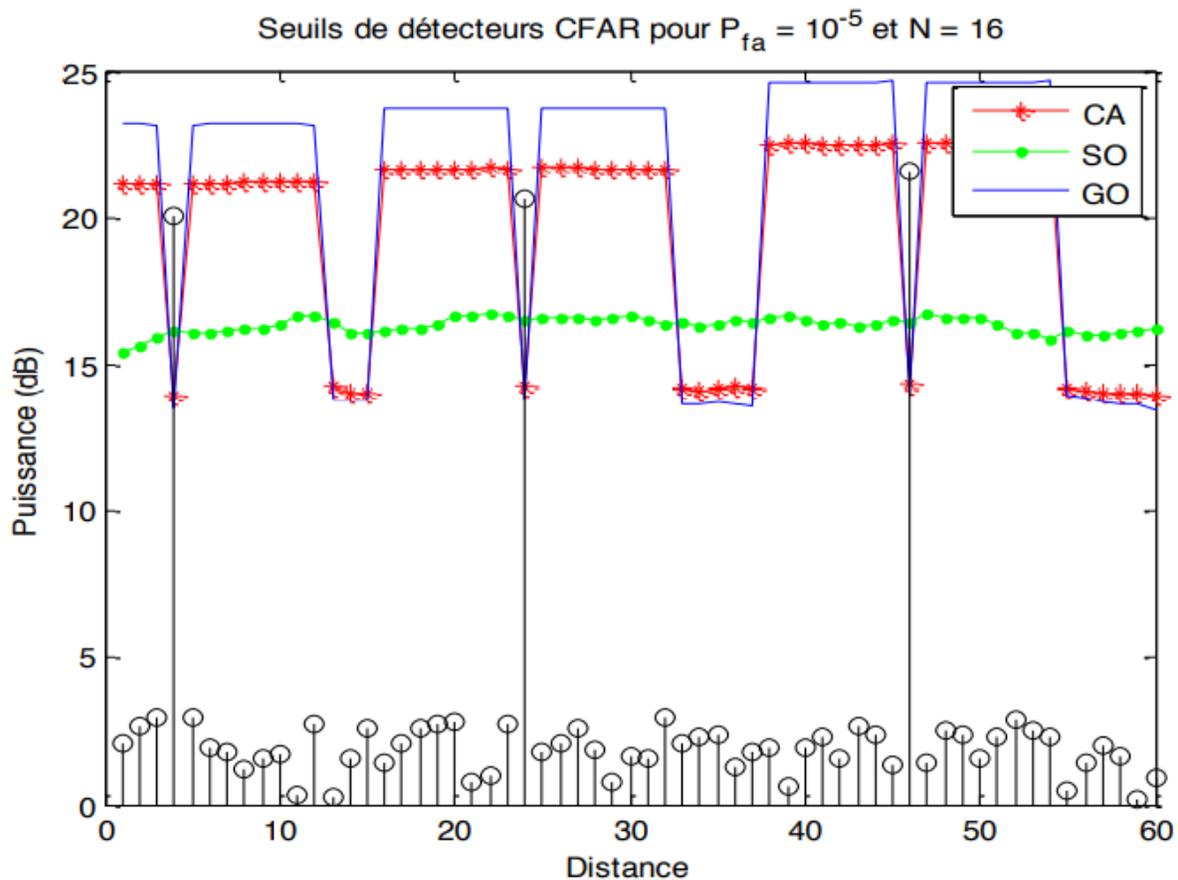
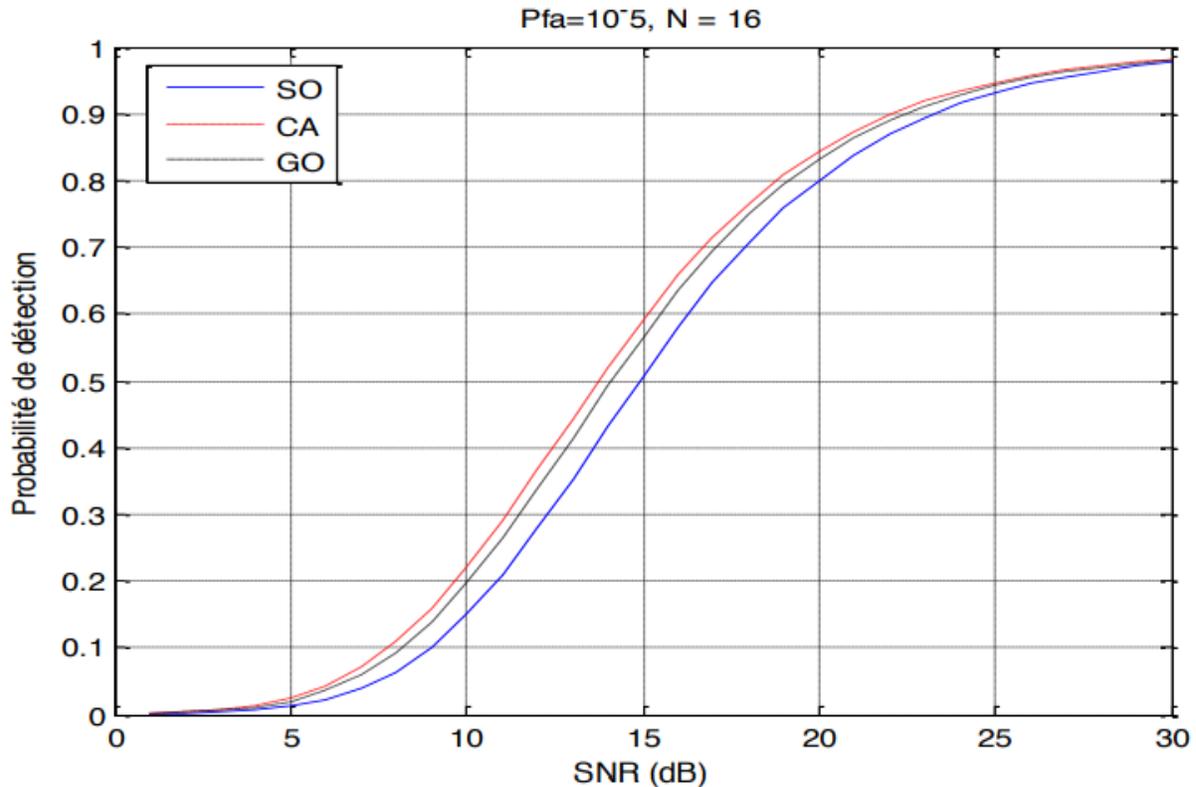


Figure 38: Seuils CA, GO et SO-CFAR en absence de cibles interférentes.

La simulation montre que les trois détecteurs présentent de bonnes performances dans un milieu homogène.

Pour une comparaison précise, nous simulons la probabilité de détection en fonction du SNR pour chaque détecteur.



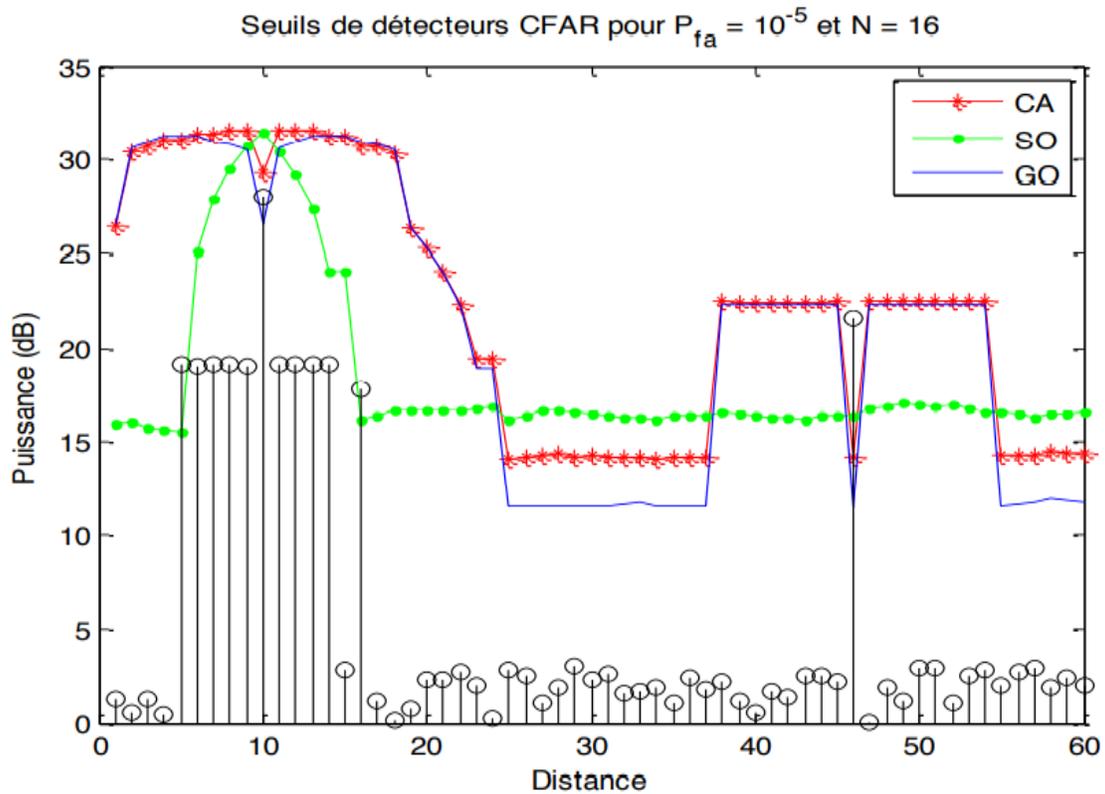
**Figure 39:**  $P_d$  en fonction du SNR pour CA-CFAR, SO-CFAR et GO-CFAR en absence de cibles interférentes.

De la Figure 35, nous constatons que le **CA-CFAR** est le plus adapté à cette situation. Le détecteur SO-CFAR présente les plus grandes pertes de détection, quant au détecteur **GO-CFAR**, il présente de légères pertes de détection relativement au détecteur **CA-CFA**

### 3.5.2 Cas d'un bord de clutter :

Le bord de clutter peut être défini comme une transition dans la puissance de clutter dû à la réflexion d'une partie de l'énergie émise par un réflecteur naturel ou artificiel. Nous avons simulé les performances du **CA**, **GO** et **SO-CFAR** en présence d'un bord de clutter pour une  $P_{fa}=10^{-5}$  et

$N = 16$ .



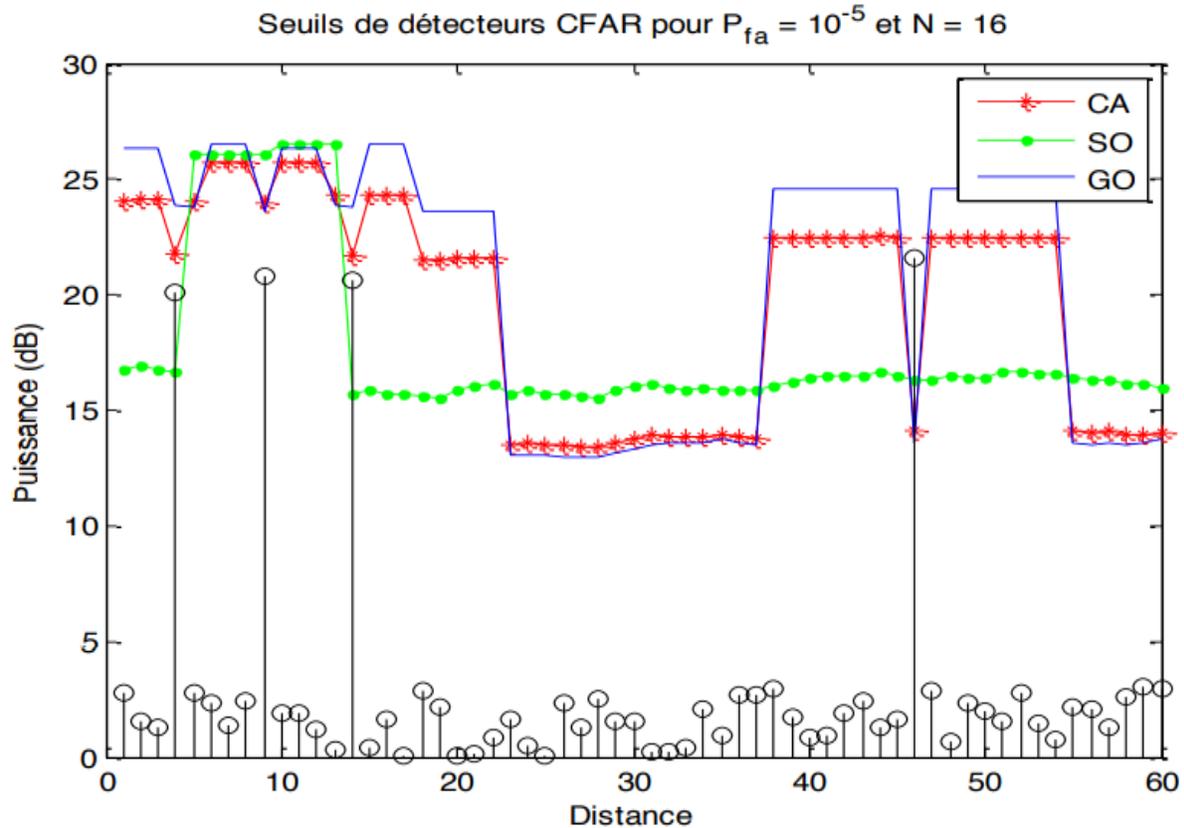
**Figure 40:** Seuils CA, GO et SO-CFAR en présence d'un bord de clutter.

D'après la Figure 40, nous constatons que :

- Le détecteur CA-CFAR n'est pas performant en présence de bord de clutter.
- Le détecteur GO-CFAR est plus performant que le SO-CFAR dans le cas de bord de clutter si la cellule sous test est noyée dans le clutter et le contraire si la cellule sous test se trouve dans la zone claire.

### 3.5.3 Cas de présence de cibles interférentes :

Dans la Figure 41 nous avons étudié le comportement des détecteurs CA, GO et SOCFAR en présence des cibles interférentes, pour une  $P_{fa}=10^{-5}$  et  $N = 16$ .



**Figure 41:** Seuils CA, GO et SO-CFAR en présence des cibles interférente.

Nous constatons que :

- Le détecteur CA-CFAR n’est pas performant en présence de cibles interférentes.
- Le détecteur GO-CFAR n’est plus performant dans ce cas, car il sélectionne toujours la fenêtre possédante la plus grande puissance, ce qui augmente le seuil et par conséquent détériore la probabilité de détection.
- Le détecteur SO-CFAR est plus performant que le GO-CFAR et le CA-CFAR si les cibles interférentes sont dans un seul côté. Ceci est dû au fait qu’il sélectionne toujours la fenêtre qui présente le plus faible niveau de puissance ce qui permet de ne pas prendre en compte les échos des cibles interférentes dans l’estimation du seuil.

A présent, nous avons présenté les avantages de la détection adaptative par rapport à celle classique. En faisant une analyse des détecteurs CA, GO et SO-CFAR, nous avons vu leurs limites en présence de cibles interférentes. Pour cela, plusieurs détecteurs CFAR ont été proposés pour opérer dans des milieux d’interférents. Parmi ces détecteurs, nous trouvons l’OS-CFAR.

### 3.6 Détection radar en présence d'interférences :

#### 3.6.1 Détecteur OS-CFAR :

Ce type de détecteur a été introduit en premier par Rohling [21] pour remédier aux problèmes rencontrés par les détecteurs CA, SO et GO-CFAR en présence de cibles interférentes.

Pour le détecteur OS-CFAR Les N échantillons de la fenêtre de référence sont d'abord arrangés en ordre croissant suivant leurs amplitudes pour obtenir la suite d'échantillons ordonnés suivante :

$$X_1 \leq X_2 \leq X_k \leq X_{N-1} \leq X_N \quad (3-23)$$

Avec  $X_1$  est le plus petit échantillon et  $X_N$  le plus grand. L'échantillon d'ordre K est ensuite choisi pour estimer le seuil.

$$Q = X_k \quad (3-24)$$

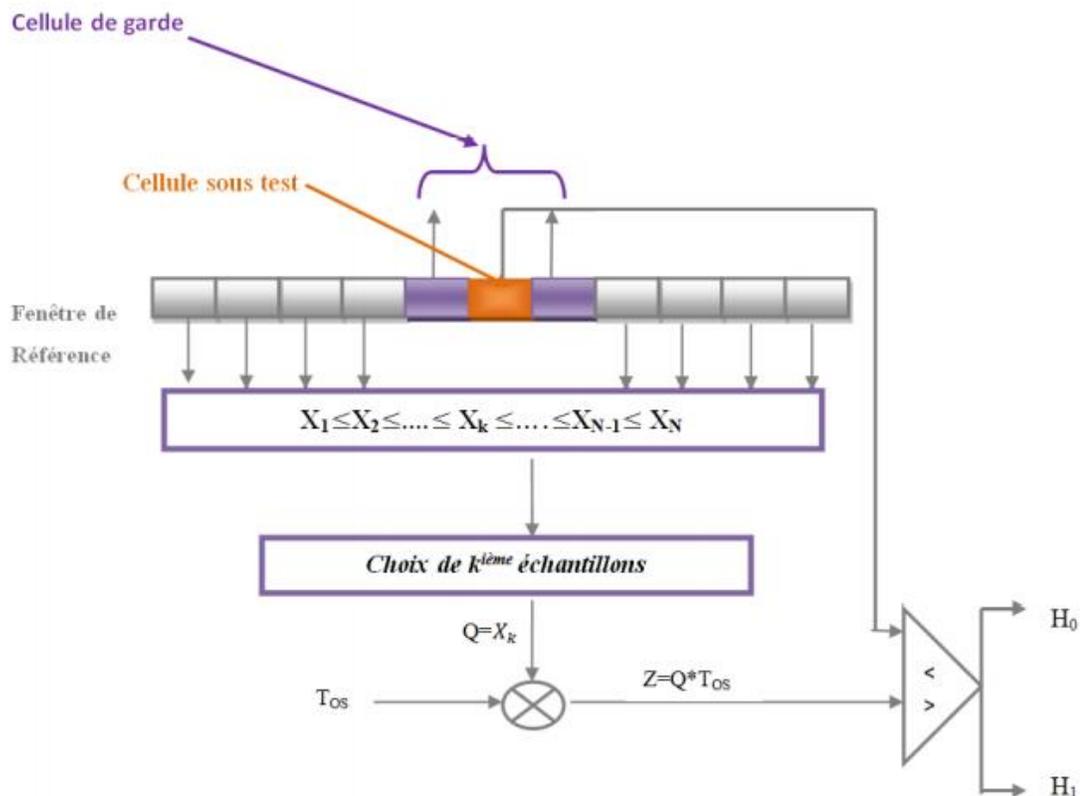


Figure 42: Schéma fonctionnel du détecteur OS-CFAR.

Ce détecteur offre une certaine immunité aux cibles interférentes. En effet, les échos provenant de ces dernières occupent les positions les plus hautes dans la suite ordonnée et ne seront pas pris en compte dans l'estimation de la puissance du seuil.

Il est évident que, si  $m$  est le nombre de cibles interférentes qui apparaissent dans la fenêtre de référence, il faut choisir le paramètre  $K$  tel que  $k \leq N - m$ . Si cette condition n'est pas vérifiée, l'échantillon choisi correspondra à un écho d'une cible interférente et de ce fait la détection de la cible primaire se trouve compromise. L'analyse de l'effet du paramètre  $K$  sur les performances de détection dans un clutter de type gaussien a été étudiée et la valeur préconisée pour  $K$  égale à  $3 \frac{N}{4}$  [21].

### ➤ Etude probabiliste détecteur OS-CFAR :

Rohling a montré que la probabilité de détection de l'OS-CFAR peut s'exprimer sous la forme suivante :

$$P_d = k \binom{N}{K} \frac{(T_d + N - K)!(K-1)!}{(T_d + N)!} \quad (3-25)$$

Avec :

$$\binom{N}{K} = \frac{N!}{(N-K)!K!} \quad ; \quad T_d = \frac{T_{OS}}{1+S}$$

$S$  : Le rapport signal à bruit ;

$K$  : L'ordre de la cellule que nous allons prendre pour l'estimation du seuil ;

$T_{OS}$  : Le facteur d'échelle.

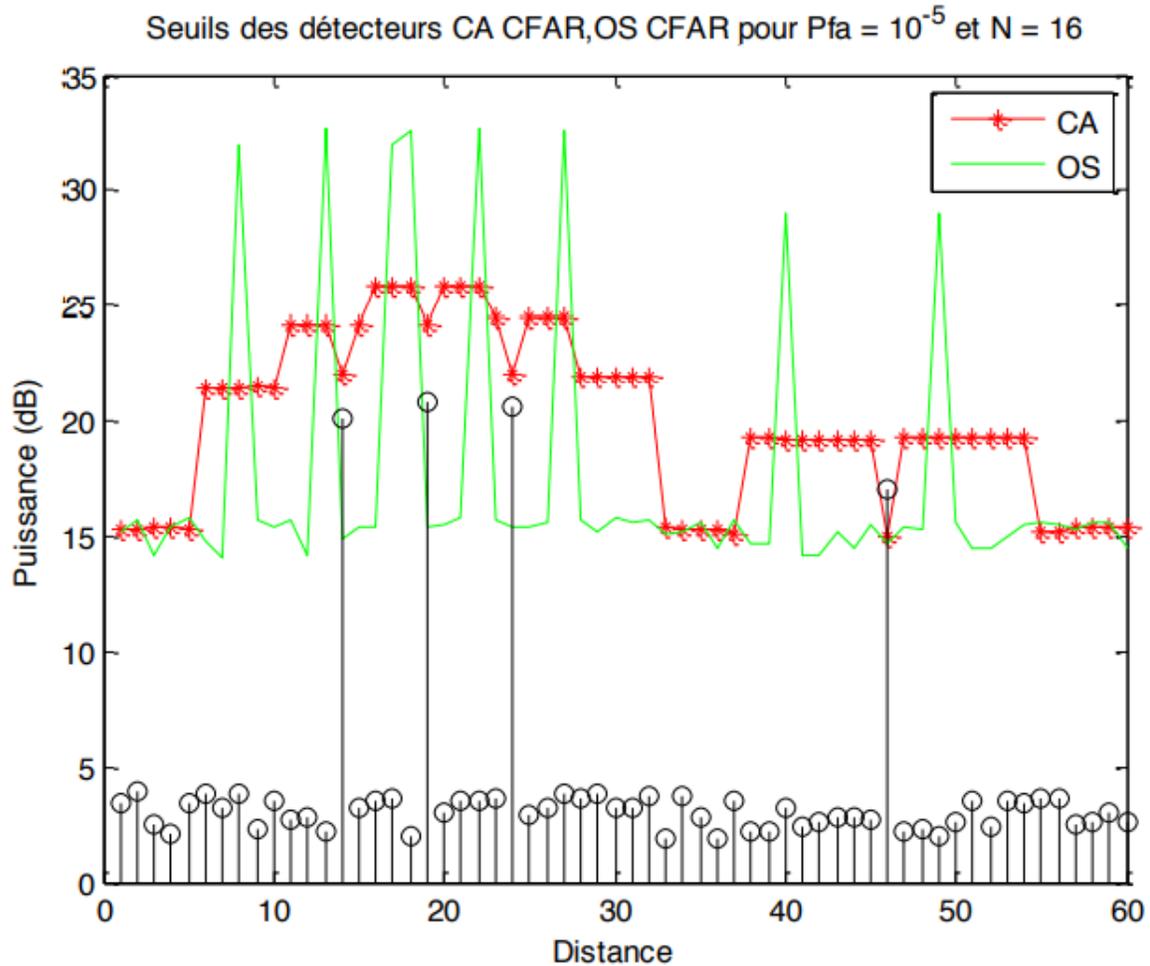
En remplaçant  $S$  par zéro dans l'expression (3-25), nous trouvons l'expression de  $P_{fa}$

$$P_{fa} = k \binom{N}{K} \frac{(T_{OS} + N - K)!(K-1)!}{(T_{OS} + N)!} \quad (3-26)$$

Dans un détecteur OS-CFAR, la probabilité de détection est en fonction du rapport signal sur bruit  $S$ , le nombre de cellules de référence  $N$ , l'ordre  $K$  et du facteur d'échelle  $T_{OS}$ . Ce facteur d'échelle  $T_{OS}$  peut être déterminé à partir de la résolution de l'équation (3-26) de probabilité de fausses alarmes ou bien à partir de la courbe  $P_{fa}$  en fonction de  $T_{OS}$ .

➤ Performances du détecteur OS-CFAR :

Nous avons simulé le fonctionnement de l'OS-CFAR et CA-CFAR en présence de cibles interférentes. Les paramètres de simulation ont été effectués pour  $N=16$  et  $P_{fa}=10^{-5}$ .



**Figure 43:** Seuil CA et OS-CFAR en présence des cibles interférentes.

Nous constatons que les deux détecteurs CA-CFAR et OS-CFAR détectent la cible dans le milieu homogène. Cependant, les cibles interférentes présentes aux cellules distances: 14, 19 et 25, elles sont ratées par CA-CFAR et détectées par l'OS-CFAR.

Le détecteur OS-CFAR présente de meilleures performances en présence de cibles interférentes, mais il présente des pertes CFAR supplémentaires dans un milieu homogène comparativement au détecteur CA-CFAR.

La simulation montre que le CA-CFAR présente de meilleures performances en termes de probabilité de détection que l'OS-CFAR dans un milieu homogène sans présence de cibles interférentes.

Dans cette partie nous avons effectué l'étude de l'OS-CFAR avec une comparaison de ses performances avec celles de CA-CFAR, le premier a de meilleures performances au présence de cibles interférentes et le deuxième dans un milieu homogène.

### 3.7 Proposition d'une architecture d'exploitation des techniques d'estimations de la PDF à la détection CFAR :

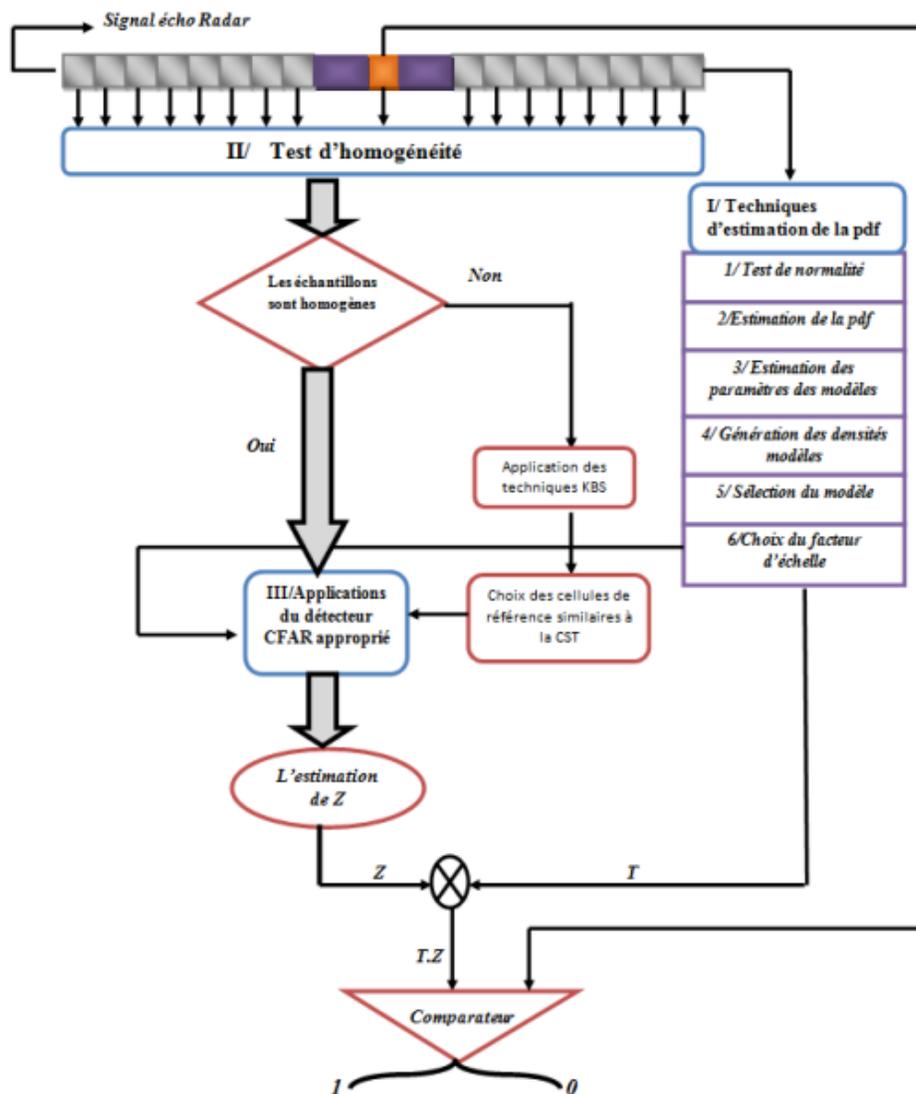


Figure 44: Organigramme basée sur l'estimation de la fonction de densité de probabilité.

### 3.7.1 Explication de l'approche proposée :

Cette approche est formée de trois étapes essentielles

#### 3.7.1.1 Techniques d'estimation de la PDF :

- **Test de normalité :**

Le test de normalité permet de décider l'hypothèse concernant une loi de probabilité est convenable avec la réalisation d'un échantillon autrement dit test de normalité est la vérification de gaussienneté du clutter.

- **Estimation de la PDF :**

On choisit un estimateur (estimateur par noyau fixe, estimateur par noyau adaptatif, et estimateur par mixture...) permettent le choix convenable de l'estimateur qui présente les meilleurs performances.

- **Estimation des paramètres des modèles :**

Il y a deux méthodes pour estimer les paramètres des densités de probabilité à partir des données : la méthode du maximum de vraisemblance et la méthode du moment.

- **Génération des densités modèles :**

Une fois les paramètres des différentes densités estimés, peuvent être générés sous forme de vecteurs qui seront comparés à la densité estimée. Les expressions des densités (Weibull, log-normale, ...etc.). En fonction de leurs paramètres sont données dans le premier chapitre.

- **Décision et sélection d'une densité modèle :**

Cette partie peut être fondée sur des mesures d'écart et le choix du modèle le plus proche dans le sens des mesures d'écart. Enfin, pour la construction de détecteurs basés sur les méthodes d'estimation des fonctions de densité de probabilité, à partir des données de la cellule sous test et à travers l'utilisation de la valeur appropriée du multiplicateur de seuil quel que soit le CFAR adopté.

- **Choix du facteur d'échelle :**

Ce choix est obtenu en utilisant des équations ou des valeurs tabulées selon le type du modèle de clutter obtenu.

### 3.7.1.2 Test d'homogénéité :

Dans cette étape, y a deux chemins à envisager :

Le premier cas : si les échantillons de la PDF estimée sont homogènes ; c.-à-d sont indépendants, identiquement distribués (IID), alors on passe à l'étape suivante, sinon on doit passer aux techniques KBS (Knowledge Based Systems) pour connaitre les cellules qui ont la même IID que la cellule sous test et enlever la non homogénéité du clutter.

### 3.7.1.3 Application du détecteur CFAR approprié :

Pour cette étape on applique pour chaque type du clutter estimé (défini par sa propre PDF) un détecteur CFAR qui le convient.

Par exemple : on applique le détecteur CA-CFAR pour les clutter de modèle de Rayleigh et exponentiel, et le détecteur ML « maximum likelihood » et logt- CFAR pour le clutter à densité de probabilité Weibull et log-normal.

### Conclusion :

Dans ce chapitre, nous avons analysé les détecteur CFAR : CA, GO, SO et l'OS-CFAR dans plusieurs situations d'environnement. Dans le cas d'un environnement homogène, le CA-CFAR est le plus adapté. Ses performances sont proches de celles du détecteur idéal de Neyman-Pearson lorsque le nombre  $N$  est important. Lorsque l'environnement n'est plus homogène (bord de clutter et cibles interférentes), les performances se dégradent d'une manière considérable.

En présence de l'effet de bord de clutter, si la cellule sous test est noyée dans ce dernier, le détecteur GO-CFAR est plus performant. Par contre si cette cellule sous test se trouve dans la partie claire, le détecteur SO-CFAR est plus robuste.

En présence de cibles interférents dans l'une des demi-fenêtres  $U$  ou  $V$ , le détecteur SO-CFAR est le mieux indiqué. Par contre si les deux demi-fenêtres contiennent des cibles interférentes, aucun des trois détecteurs n'offre une bonne détection.

Pour cette dernière situation, le plus adapté est le détecteur OS-CFAR. Ce type de détecteur donne de meilleurs résultats dans le cas de cibles multiples quel que soit leur position dans la cellule de référence. L'OS-CFAR peut tolérer un nombre maximal de cibles interférentes égal à  $N-k$ , un bon choix du paramètre  $k$  est exigé pour avoir de bonnes performances.

## **Conclusion générale :**

Au cours de notre travail de fin d'études, nous avons traité les techniques de l'estimation non-paramétrique de la fonction de densité de probabilité, nous avons utilisé méthode de (Parzen-Rosenblatt) et celle du noyau adaptatif pour estimer la densité de probabilité.

Ces méthodes sont particulièrement utilisées pour évaluer la performance de l'estimateur.

Le premier chapitre a été consacré à l'introduction des notions abordées le long du mémoire, une intervention particulière a été accordée aux propriétés statistiques des différents modèles de clutter de façon à souligner les particularités de chaque modèle.

Le deuxième chapitre a été dédié à l'estimation des fonctions de densités de probabilité à partir d'un échantillon. Trois techniques sont présentées, à savoir technique basée sur l'histogramme, le polygone de fréquence, et les histogrammes décalées qui sont des méthodes graphiques les plus simples par rapport aux méthodes des noyaux fixes et les noyaux adaptatifs sont exposés comme méthodes plus évoluées. Les résultats de performance de ces estimateurs sont montrés par des simulations.

Le dernier chapitre est dévolu à la détection adaptative CFAR et comporte deux parties:

Dans la première partie nous avons introduit des rappels sur la théorie de détection et l'état de l'art de la détection CFAR.

Dans la seconde partie nous proposons une approche sous forme d'un organigramme qui permet d'exploiter d'une façon efficace l'estimation de la densité de probabilité à la détection CFAR.

Notre humble travail se termine par la proposition de l'utilisation d'un estimateur à noyau fixe qui permet d'estimer les paramètres de densités de probabilité.

## Annexes :

### Annexe A :

#### ❖ Propriété de l'estimateur à noyau :

L'espérance :

$$E(f_n(x)) = E\left[\frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)\right]$$

$$= \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)$$

$$E(f_n(x)) = \frac{1}{h} \int k\left(\frac{x-u}{h}\right) f(u) du$$

Changement de variable :

$$y = \frac{x-u}{h} \Rightarrow dy = \frac{-1}{h} du \Rightarrow du = -h dy$$

$$u = -yh + x$$

$$E(f_n(x)) = \int k(y) f(x-yh) dy$$

Le développement limite de Taylor à l'ordre 2 au point  $y = 0$  de  $f(x-yh)$

On obtient

$$f(x-yh) = f(x) - \frac{hy}{1!} f'(x) + \frac{h^2 y^2}{2!} f''(x) + o^2(h^2)$$

Donc  $E(f_n(x))$  sera donnée par :

$$E(f_n(x)) = \frac{1}{h} \int k(y) \left[ f(x) - \frac{hy}{1!} f'(x) + \frac{h^2 y^2}{2!} f''(x) + o^2(h^2) \right] dy$$

$$= f(x) \int k(y) dy - hf'(x) \int yk(y) dy + \frac{h^2}{2!} f'' \int y^2 k^2(y) dy + o^2(h^2)$$

Si le noyau est symétrique par rapport à 0 :

$$\int k(y) dy = 1; \int yk(y) dy = 0; \int y^2 k^2(y) dy \neq 0 = \sigma^2$$

Alors :

$$E(f_n(x)) = f(x) \frac{h^2}{2!} f''(x) \sigma^2 + o^2(h^2)$$

**Biais :**

$$\begin{aligned} \text{biais}(f(x)) &= E[f_n(x)] - f(x) \\ &= \int yk(y)f(x - hy) dy - f(x) \\ &= f(x) \frac{h^2}{2} f''(x) \sigma^2 + o(h^2) - f(x) \\ &= \frac{h^2}{2} f''(x) \sigma^2 + o(h^2) \end{aligned}$$

**Variance :**

$$\begin{aligned} \text{var}[f_n(x)] &= \text{var}\left[\frac{1}{nh} \sum_{i=1}^n k\left(\frac{x - x_i}{h}\right)\right] \\ &= \frac{1}{n^2 h^2} \sum_{i=1}^n E\left[k\left(\frac{x - x_i}{h}\right)\right]^2 - \frac{1}{n^2 h^2} \left(E\left[k\left(\frac{x - x_i}{h}\right)\right]\right)^2 \\ &= \frac{1}{nh^2} \int \left(k\left(\frac{x - u}{h}\right)\right)^2 f(u) du - \frac{1}{n} \left[\frac{1}{h} \int k\left(\frac{x - u}{h}\right) f(u) du\right]^2 \end{aligned}$$

**Changement de variable :**

On prend :

$$\frac{x - u}{h} = y \Rightarrow u = x - hy \Rightarrow du = -h dy$$

$$\text{var}[f_n(x)] = \frac{1}{nh} \int (k(y))^2 f(x - hy) dy - \frac{1}{n} \left[ \int k(y) f(x - hy) dy \right]^2$$

Le développement limite de Taylor à l'ordre 2 au point  $y = 0$  de  $f(x - hy)$ . On obtient :

$$f(x - yh) = f(x) - \frac{hy}{1!} f'(x) + \frac{h^2 y^2}{2!} f''(x) + o(h^2)$$

$$\begin{aligned} \text{var}[f_n(x)] &= \frac{1}{nh} \int (k(y))^2 \left[ f(x) - \frac{hy}{1!} f'(x) + \frac{h^2 y^2}{2!} f''(x) + o(h^2) \right] dy \\ &\quad - \frac{1}{n} \left[ \int k(y) \left[ f(x) - \frac{hy}{1!} f'(x) + \frac{h^2 y^2}{2!} f''(x) + o(h^2) \right] dy \right]^2 \end{aligned}$$

$$\begin{aligned} \text{var}[f_n(x)] &= \frac{1}{nh} \int k^2(y) dy - \frac{h}{nh} f' \int y k^2(y) dy + \frac{h^2}{2nh} f'' \int y^2 k^2(y) dy + o(h^2) \\ &\quad - \frac{1}{n} \left[ f(x) \int k(y) dy - h f'' \int y k(y) dy + \frac{h^2}{2} f'' \int y k^2(y) dy + o(h^2) \right] \end{aligned}$$

$$\begin{aligned} &= \frac{1}{nh} f(x) \int k^2(y) dy - \frac{f'}{n} \int y k^2(y) dy + \frac{h f''}{2n} \int y^2 k^2(y) dy + o(h^2) - \frac{1}{n} \left[ f(x) + \frac{h^2}{2} f'' \sigma^2 \right. \\ &\quad \left. + o(h^2) \right] \end{aligned}$$

$$\text{var}[f_n(x)] = \frac{f(x)}{nh} \int k^2(y) dy$$

## 🚩 Annexe B :

### ❖ Rappel sur les propriétés statistiques :

Dans ce qui suit, on considère une variable aléatoire  $X$  ayant la fonction de densité de probabilité PDF.

#### 1) Notion d'espérance mathématique :

Un concept de base pour définir les propriétés d'une distribution est l'espérance d'une Fonction  $h$  de la variable aléatoire  $X$  qui est donnée par :

$$E[h(x)] = \int_{-\infty}^{+\infty} h(x) f(x) dx$$

#### 2) Moments :

Le moment d'ordre r par rapport à l'origine correspond à :  $h(x) = X_r$

$$\mu_r^0 = \int_{-\infty}^{+\infty} x^r f(x) dx$$

Le moment d'ordre r par rapport à la moyenne correspond à :

$$h(x) = [X - E\{x\}]^r$$

$$\mu^r = \int_{-\infty}^{+\infty} (x - \mu_1^0)^r f(x) dx$$

### 3) Moyenne :

La moyenne correspond à la valeur autour de laquelle se situent toutes les mesures possibles sur la variable aléatoire X, elle correspond au moment d'ordre un (r=1) par rapport à l'origine et constitue une mesure de localisation.

$$\mu = E[X] = \int_{-\infty}^{+\infty} x f(x) dx$$

### 4) Variance :

La variance est équivalente au moment d'ordre deux (r= 2) par rapport à la moyenne, c'est une mesure de dispersion du modèle particulièrement utile. La racine carrée de la variance est appelée écart type.

$$\sigma^2 = E[\{X - E[X]\}^2] = \int_{-\infty}^{+\infty} (x - \mu_1^0)^2 f(x) dx$$

### 5) Facteurs de forme :

En plus des propriétés décrites pour les mesures de localisation et de dispersion, deux autres facteurs sont utilisés pour caractériser la forme d'une distribution : le Skewness et le Kurtosis.

#### ➤ Skewness (coefficient d'asymétrie) :

Ce facteur indique l'étendue de l'asymétrie ou l'inclinaison que présente la loi par rapport à sa moyenne. Une valeur de skewness positive est synonyme d'une queue à droite de la densité de probabilité plus longue que celle à gauche (inclinaison vers la gauche), une valeur négative est synonyme du contraire, tandis qu'une valeur nulle implique la symétrie. Le skewness est défini par:

$$\gamma_1 = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{E[X - \mu_1]^3}{E[(X - \mu_1)^2]^{3/2}}$$

➤ **Kurtosis (coefficient d'aplatissement) :**

Ce facteur est un indice du caractère impulsif d'une densité. Des valeurs élevées de Kurtosis signifient des densités à queues plus lourdes (lentes) que la gaussienne alors que les valeurs négatives indiquent l'inverse. La gaussienne est caractérisée par un Kurtosis nul. Le Kurtosis est défini par :

$$\gamma_2 = \frac{\mu_4}{(\mu_2)^2} - 3 = \frac{E[X - \mu_1]^4}{E[(X - \mu_1)^2]^2} - 3$$

En somme pour le skewness on retient ce qui suit :

- Skewness positif : queue à droite et inclinaison vers la gauche.
- Skewness nul : courbe symétrique.
- Skewness négatif : queue à gauche et inclinaison vers la droite.

Pour le Kurtosis on retient ce qui suit :

- Aplatissement positif : courbe aigue.
- Aplatissement nul : courbe normale.
- Aplatissement négatif : courbe aplatie.

## Bibliographie :

- [1] M.Hamadouche, "Détection Adaptive CA-CFAR et CMAP-CFAR de Cibles Radar dans du clutters Gaussien et Weibull Distribuée", constantine: université de constantine, 2000.
- [2] Barkat.M, "Signal Detection and Estimation", London: Second Edition,Artech House, Boston, September 2005.
- [3] V. Hansen, "Constant False Alarm Rate Processing in Search Radars", London: Proceedings of IEE International Radar Conference, London, pp. 325-332,, October 1973.
- [4] S. B.W, "density estimation", chapman and Hall /CRC, 1986.
- [5] S. D.C, "Radar detection in weibull clutter", Artech House,, 1980.
- [6] S. Chabbi, "Détection Adaptative CFAR à Censure automatique Basée sur les Statistiques d'Ordre en Milieux non Gaussiens", Constantine: Mémoire de Magister,université de constantine, 2008.
- [7] B. Toufik, "ESTIMATION DE LA FONCTION DE DENSITE DE PROBABILITE : APPLICATION CFAR", mémoire magister Ecole militaire polytechnique, 2009.
- [8] D. C. Besma, "EQUIVALENCE DU CHOIX DE LA FENETRE POUR L'ESTIMATION DES FONCTIONS DE DENSITE ET D'INTENSITE PAR LA METHODE DES NOYAUX", mémoire de fin de cycle Université A/MIRA de Bejaia facultés de science exactes, 2015.
- [9] I. .. A. P. a. c. .O., "Statistiques d'ordre supérieur pour le Statistiques d'ordre supérieur pour le", masson, 1997.
- [10] M. Basseville, "Distance measures for signal processing and pattern recognition", INRIA-PENNES, September 1998.
- [11] P. r. Bickel, "On some global measures of the deviatio ofdensity function estimates", Annales statistiques., 1973.
- [12] G.Saporta, "Probabilité, analyse des données et statistiques", paris: Technip, 1990.

- [13] D. T. a. J. Thompson, "Kernel density Estimation revisited, non linear analysis, Theory method", U.S.A: Baylor College of Medicine. Houston, , 1977.
- [14] J. D.Bosq., "Théorie de l'estimation fonctionnelle", Economicaedition, 1967.
- [15] D. G.R.TERELL, "Biased and unbiased cross-validation in density estimation", journal of the American Statistical Association, 1987.
- [16] M. Rudemo, "Empirical choice of histograms and kernel density estimator", Scandinavian journal of statistics, 1956, et bozman1984.
- [17] S. D.C., "Automatic detection and radar data processing", Artech house, 1980.
- [18] H. M. F. a. R. S. Johnson, "Adaptive detection mode with threshold control as a function of spatially sampled clutter-level estimates", RCA Review, Vol. 29, N°3, September 1968.
- [19] V. G. S. J. H. Hansen, "Delectability Loss Due to Greatest of Selection in a Cell-Averaging CFAR", IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-16, January 1980.
- [20] G. Trunk, "Range Resolution of Targets Using Automatic Detectors", IEEE Transactions on Aerospace and Electronic Systems, Volume: AES-14, N° 5, September 1978..
- [21] H. Rohling, "Radar CFAR Thresholding in Clutter and Multiple Target Situations", IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-19,, 1983.