



Université ABBES LAGHROUR Khenchela
Faculté des Sciences et de la Technologie
Département de Génie Industriel
جامعة عباس لغرور خنشلة
كلية العلوم والتكنولوجيا
قسم الهندسة الصناعية



N° Série :

Mémoire de fin d'étude

Pour l'obtention du diplôme de Master

Filière : Télécommunications

Spécialité : Systèmes des Télécommunications

THEME

**Détection d'interférences pour les systèmes
de navigation par satellites en utilisant les
méthodes d'optimisation multi-objectif
NSGAI**

Réalisé par : MAHREZ Imane

BADAOUI Manar

Soutenu le : 28/06/2022 *Devant le jury composé de:*

Pr. BEDRA Sami

Dr. DOUAK Fouzi

Dr. KHALFAOUI Fatima

Président

Encadreur

Examinatrice

Université Abbes Laghrour-Khenchela

Université Abbes Laghrour-Khenchela

Université Abbes Laghrour-Khenchela

Promotion 2021/2022

Je dédie ce travail à :

A ma mère, pour son amour et ses sacrifices

*A mon père, pour son soutien et la confiance qu'il m'a
accordé*

A vous mes frères : Karim, Amine et Islam

*Qui m'avez toujours soutenu et encouragé durant ces années
d'études*

A mon oncle qui a toujours été avec moi - رحمه الله -

A tous les membres de ma famille et tous mes amis

A tous mes collègues de promotion

Imane

Je dédie ce mémoire à :

*Mes très chers parents qui ont toujours été là pour moi, et qui m'ont donné un
magnifique modèle de labeur et de persévérance.*

Mes sœurs et mes frères,

*Mes neveux et nièces " Meriem, Moataz, Malak, Enes, Maria,
Chahed, Abd-Elmoemin."*

*À celui qui m'a toujours encouragé et soutenu moralement, mon cher Mari
Taha.*

À ma partenaire à ce travail Imane.

*Aux personnes qui m'ont encouragé et supporté durant ce travail. Ils sont
nombreux et tous présents dans mon cœur.*

À ceux qui ont attendu avec patience les fruits de leur bonne éducation.

À Tous mes collègues de promotion,

Manar

Remerciements

Nous tenons à remercier ALLAH tout puissant qui nous a donné la santé, le courage, et la volonté pour réaliser ce travail.

*Nous remercions notre encadreur Monsieur Dr. **Fouzi Douak**, enseignant à l'université de Khenchela, pour son suivi, pour son aide précieux.*

*Nous adressons nos plus sincères remerciements à Monsieur Pr. **Sami Bedra**, enseignant à l'université de Khenchela, pour nous avoir fait l'honneur de présider le jury de notre mémoire. Nous remercions aussi à Madame Dr. **Fatima Khalfaoui**, enseignante à l'université de Khenchela, qui a accepté d'examiner ce travail.*

Nous tenons à remercier particulièrement tous les enseignants et les responsables de notre département, qui ont veillé sur notre formation.

Nos remerciements vont également à tout le personnel du département du génie industriel pour leur gentillesse et leur disponibilité.

Enfin, nous exprimons notre remerciement à tous ceux qui ont contribué de près ou de loin à l'aboutissement de ce travail.

Résumé

La méthode proposée dans ce travail, a pour but de proposer de nouveaux outils et concepts basés essentiellement sur les techniques d'optimisations pour la détection et la suppression des interférences dans les récepteurs GNSS.

Notre méthode consiste à combiner la puissance de classifieur OCSVM (SVM à une seule classe) avec une technique d'optimisation heuristique basée sur les algorithmes génétiques uni-objectif GA et multi-objectif NSGAI (Nondominated Sorting Multi-objective Genetic Algorithm) pour filtrer les bases de données contenant des valeurs aberrantes (interférences), dans l'espace de prédiction et dans l'espace des caractéristiques. On optimise simultanément de nombreuses fonctions objectives complémentaires pour une adaptation automatique et donc une meilleure détection des interférences.

Les principales fonctions objectives à optimiser sont : l'utilisation d'un critère lié à l'espace des caractéristiques des échantillons qui est la distance Bhattacharyya, et trois critères qui affectent directement sur les performances de prédiction de classifieur OCSVM : la racine de l'erreur quadratique moyenne, le coefficient de corrélation et l'erreur de classification.

Les résultats que nous avons obtenus, après la simulation sont assez satisfaisants du point de vue amélioration des performances de l'identification et de la suppression des interférences par rapport au classifieur OCSVM classique.

Mots clés: GNSS, Interférence, Les algorithmes génétique, Fonctions de fitness, GA, NSGAI, Front de Pareto, Méthodes de classification, OCSVM.

Abstract

The proposed method in this work, aims to propose new tools and concepts based essentially on optimization techniques for the detection and the suppression of interference in GNSS receivers.

Our method consists in combining the power of the OCSVM (One class support vector machine) classifier with a heuristic optimization technique based on single-objective GA and multi-objective genetic algorithms NSGAI (Nondominated Sorting Multi-objective Genetic Algorithm) to filter the dataset containing outliers (interferences), in the both cases the prediction space and the feature space. Many additional fitness functions are simultaneously optimized for an automatic adaptation and therefore a better detection of interferences.

The main fitness functions to be optimized are: the use of a criterion that applied to the feature space of the samples which is the Bhattacharyya distance, and three criteria which are directly affect in the prediction performance of the OCSVM classifier: the root mean square error, the correlation coefficient and the classification error.

The results that we obtained, after the simulation, are quite satisfactory from the point of view of improving the performance of identification and suppression of interferences compared to the classic OCSVM classifier.

Keywords: GNSS, Interference, Genetic algorithms, Fitness functions, GA, NSGAI, Pareto front, Classification methods, OCSVM.

ملخص

تهدف الطريقة المقترحة في هذا العمل إلى اقتراح أدوات ومفاهيم جديدة تستند أساساً إلى تقنيات التحسين لاكتشاف وإزالة التداخل في مستقبلات GNSS.

تتكون طريقتنا من الجمع بين قوة المصنف OCSVM (تصنيف SVM فئة واحدة) مع تقنية تحسين تعتمد على الخوارزميات الجينية أحادية الهدف GA و متعددة الأهداف NSGAII لهدف تصفية قواعد البيانات التي تحتوي على القيم المتطرفة (التداخلات) ، في كلتا الحالتين في فضاء التنبؤ بالإضافة إلى فضاء الخصائص. تم تحسين العديد من الوظائف في وقت واحد من أجل التكيف التلقائي وبالتالي اكتشاف التداخلات بشكل أفضل.

الوظائف الرئيسية المطلوب تحسينها هي: استخدام معيار يتم تطبيقه على فضاء الخصائص للعينات وهي مسافة Bhattacharyya ، وثلاثة معايير تؤثر بشكل مباشر في أداء التنبؤ لمصنف OCSVM: جذر متوسط الخطأ التربيعي ومعامل الارتباط وخطأ التصنيف.

النتائج التي حصلنا عليها، بعد المحاكاة، مرضية تمامًا من وجهة نظر تحسين أداء تحديد وإزالة التداخلات مقارنة بمصنف OCSVM الكلاسيكي.

الكلمات المفتاحية: GNSS ، التداخل ، الخوارزميات الجينية ، وظيفة التقييم ، GA ، NSGAII ، Pareto front ، طرق التصنيف ، OCSVM

Liste des tableaux

Tableau IV.1. Paramètres du signal E5a1.	63
Tableau IV.2. Nombre total des échantillons apprentissage, validation et test.....	69
Tableau IV.3. Paramètres du GA et NSGAI1.....	73
Tableau IV.4. Les résultats obtenus pour les différents cas d'ISR (10 à 60 dB).	75
Tableau IV.5. Les résultats obtenus de GA-RMSE pour les différents cas d'ISR 10 à 60 dB.	77
Tableau IV.6. Les résultats obtenus par les différentes méthodes d'optimisation GA-r, GA- ERR et GA-D.	80
Tableau IV.7. Les résultats obtenus par les différentes méthodes d'optimisation multi- objectif NSGAI1-D-r, NSGAI1-D-RMSE, NSGAI1-D-ERR, NSGAI1-RMSE-ERR et NSGAI1-r-ERR.....	84

Liste des figures

Figure I.1. Principe de trilatération.....	6
Figure I.2. Segment du système GNSS	7
Figure I.3. Phase d'opérationnalité des différents systèmes GNSS existants.....	10
Figure I.4. Structure générales des signaux GNSS.....	11
Figure I.5. Chaîne de récepteur GNSS.	11
Figure I.6. Schéma global d'un récepteur GPS.	12
Figure I.7. Exemples de résultat d'acquisition.	13
Figure I.8. Principe d'acquisition par recherche série.....	14
Figure I.9. Acquisition en mode parallèle du code et de la phase.....	15
Figure I.10. Recherche en parallèle du Doppler et du retard de code.	16
Figure I.11. Schéma fonctionnel d'une boucle de poursuite du code et de la porteuse.	17
Figure I.12. Différentes type d'interférences GNSS.	17
Figure I.13. Les brouilleurs.	22
Figure II.1. Exemple de codage binaire d'une deux chromosome.....	27
Figure II.2. Exemple de codage réel d'une deux chromosome.....	28
Figure II. 3. Croisement avec un point de crossover.....	29
Figure II.4. Croisement avec 2 points de crossover.	29
Figure II.5. Illustration de la mutation par inversion.	30
Figure II.6. Illustration de la mutation par échange.	30
Figure II.7. Illustration de la mutation insertion.	31
Figure II.8. Schéma général d'AG.	32
Figure II.9. Représentation d'un problème multi-objectif.	34
Figure II.10. Algorithme Génétique Multi objectifs.	35
Figure II.11. Exemples de frontière de Pareto.	36
Figure II.12. Front de Pareto et rangs de Pareto pour un problème de minimisation de deux objectifs.	37
Figure II.13. Dominance de Pareto et optimalité de Pareto.	38

Figure II.14. Illustration du classement par front dans le NSGA.....	38
Figure II.15. Illustration du calcul de l' i distance du NSGA II.....	40
Figure II.16. Algorithme NSGAIL.....	41
Figure II.17. Organigramme de l'algorithme NSGAIL.	42
Figure III.1. Schéma d'un modèle supervisé.	46
Figure III.2. Schéma d'un modèle non supervisé.	46
Figure III.3. L'architecture d'un réseau de neurones de type MLP.....	47
Figure III.4. Structure d'un réseau de neurone multicouche.	48
Figure III.5. Fonction de transfert seuil.	49
Figure III.6. Fonction de transfert linéaire.....	49
Figure III.7. Fonction de transfert sigmoïde.	50
Figure III.8. Fonction de transfert tangente hyperbolique.....	50
Figure III.9. Architecture d'un réseau de neurones à base de fonctions radiales	52
Figure III.10. Transformer l'espace de départ vers un espace de dimension supérieure.....	53
Figure III.11. Séparation de deux ensembles de points par un Hyperplan H.	54
Figure III.12. Cas linéairement séparable.	54
Figure III.13. Illustration d'un classificateur SVM dans le cadre d'une classification binaire.	55
Figure III.14. Deux cas de la classification (bi-classe et multi-classe).....	56
Figure III.15. One-class SVM.....	58
Figure III.16. Principe des OCSVM.	59
Figure IV.1. La densité spectrale de puissance (PSD) de la bande E5.....	63
Figure IV.2. Amplitude du signal E5aI dans le domaine temporel sans interférence.....	64
Figure IV.3. Amplitude du signal E5aI dans le domaine temporel contaminé par une interférence DMA, ISR varie entre 10dB à 60dB.....	65
Figure IV.4. Amplitude du signal E5aI dans le domaine fréquentiel sans interférence.....	65
Figure IV.5. Amplitude du signal E5aI dans le domaine fréquentiel contaminé par des interférences DMA, ISR varie entre 10dB à 60dB.	66
Figure IV.6. Exemple de classification OCSVM.	68
Figure IV.7. Schéma bloc de la suppression numérique de l'interférence pulsée.....	74

Figure IV.8. Détection et suppression des interférences dans le cas $ISR=50$ dB en utilisant OCSVM.....	76
Figure IV.9. Evolution de fitness et la détection et suppression ($ISR=60$ dB) par GA-RMSE.	77
Figure IV.10. Détection et suppression des interférences en utilisant l'algorithme GA-RMSE dans le cas $ISR=60$ dB.....	78
Figure IV.11. Evolution de la fonction fitness r en fonction de nombre de générations, $ISR=60$ dB.	78
Figure IV.12. Evolution de la fonction fitness ERR en fonction de nombre de générations, $ISR=60$ dB.	79
Figure IV.13. Evolution de la fonction fitness D en fonction de nombre de générations, $ISR=60$ dB.	79
Figure IV.14. Evolution de l'algorithme multi-objectif NSGAI-D-r en fonction de nombre de générations.	81
Figure IV.15. Front de Pareto pour l'algorithme génétique multi-objectif NSGAI-D-RMSE.	82
Figure IV.16. Front de Pareto pour l'algorithme génétique multi-objectif NSGAI-D-ERR.	82
Figure IV.17. Front de Pareto pour l'algorithme génétique multi-objectif NSGAI-ERR-RMSE.	83
Figure IV.18. Front de Pareto pour l'algorithme génétique multi-objectif NSGAI-ERR-r...	83
Figure IV.19. Evolution de RMSE en fonction des différentes valeurs ISR	85
Figure IV.20. Evolution de R en fonction des différentes valeurs ISR	85
Figure IV.21. Application de l'algorithme NSGAI-D-r pour la détection et la suppression des interférences, (a) $ISR=10$ dB, (b) $ISR=20$ dB, (c) $ISR=30$ dB, (d) $ISR=40$ dB, (e) $ISR=50$ dB, (f) $ISR=60$ dB.....	87

Liste des symboles

AWGN	Additive White Gaussian Noise
ADC	Analog to Digital Converter
AI	Artificial Intelligence
AG	Algorithms Génétiques
ATC	Air Traffic Control
ARNS	Aeronautical Radio Navigation Service
CDMA	Code Division Multiple Access
COTS	Commercials Off The Shelf
CS	Commercial Service
CWI	Continuous Wave Interference
C/A	Coarse / Acquisition
DLL	Delay Lock Loop
DME	Distance Measurement Equipment
DVB-T	Digital Video Broadcasting–Terrestrial
DVB-S	Digital Video Broad casting–Satellite
DoD	Direct out ward Dialing
DSSS	Direct-Sequence Spread-Spectrum
DCT	Discrete Cosine Transform
ECEF	Earth-Centered, Earth-Fixed
ESA	European Space Agency
EP	Evolutionary Programming
FFT	Fast Fourier Transform
FLL	Frequency Lock Loop
FM	Modulation de Frequence
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GLONASS	GLObal’<i>n</i>aya NAvigatsionnaya Sputnikovaya Sistema
JTIDS	Joint Tac tical Information Distribution Systè<i>m</i>e
KNN	kNearest Neighbors

MIDS	S ystème de D istribution d' I nformations M ultifonction
MOGA	M ulti O bjective G enetic A lgorithm
MSS	S ervice M obile par S atellite
MCWI	I nterférence à O nde C ontinue M ultiple
ML	M achine L earning
MLP	M ulti- L ayer P erceptron
NAVSTAR	N avigation S atellite T iming A nd R anging
NBI	N arrow B and I nterference
NSGA	N on-dominated S orting G enetic A lgorithm
NSGAI	N on-dominated S orting G enetic A lgorithm 2
NPGA	N iched P areto G enetic A lgorithm
PG	G enetic P rogramming
PLL	P hase L ock L oop
PPS	P recise P ositioning S ervice
PMO	P roblem M ulti- O bjective
PED	P ersonal E lectronic D evelopments
PF	P areto F ront
PVT	P osition, V elocity and T iming
RFI	R adio F rance I nternationale
RTK	R eal T ime K inematic
RMSE	R oot M ean S quare E rror
R	C orrélation
RBF	R adial B asis F unction
SAR	S earch A nd R escue
SE	E volution S trategies
SOL	S afety O f L ife
SPS	S tandard P ositioning S ervice
SSR	S urveillance S econdaire R adar
SPEA	S trength P areto E volutionary A lgorithm
SVM	S upport V ector M achine
SVDD	S upport V ector D ata D escription
OCSVM	O ne C lass S upport V ector M achine

SATCOM	C ommunications par S atellite
SCWI	Interférence à onde continue unique
TACAN	TAC tical A ir N avigation
TF	T ransformée de F ourier
UE	U nion E uropéenne
UHF	U ltra H igh F requency
UWB	U ltra- W ide B and
URSS	U nion des R épubliques S ocialistes S oviétiques
VHF	V ery H igh F requency
VHFCOM	C ommunication en VHF
WBI	W ide B and I nterference

Table des matières

Introduction Générale	1
Chapitre I	
Signaux GNSS	
I.1. Introduction	5
I.2. Principes fondamentaux du GNSS.....	5
I.2.1. La position, la vitesse et le temps	5
I.3. Segments du système de navigation.....	7
I.3.1. Segment spatial	8
I.3.2. Segment de contrôle.....	8
I.3.3. Segment d'utilisateurs	8
I.4. Différents système mondial de navigation	8
I.4.1. Système de positionnement global GPS	8
I.4.2. Système de navigation Russe GLONASS	9
I.4.3. Système de navigation Galileo	9
I.4.4. Système de navigation BEIDOU	10
I.5. Les signaux GNSS	11
I.5.1. Les principales composantes des signaux GNSS	11
I.6. Chaîne de réception GNSS.....	11
I.6.1. Architecture d'un récepteur GPS.....	12
I.6.2. Acquisition du signal GNSS.....	13
I.6.2.1. Mode d'acquisition série	14
I.6.2.2. Acquisition par recherche parallèle sur le retard de code par (FFT).....	15
I.6.2.3. Acquisition par recherche parallèle sur le décalage de fréquences par (FFT)	
.....	15
I.6.3. Poursuite du signal GNSS	16
I.7. Classification de la source interférente	17
I.7.1. Caractéristiques spectrales d'interférences.....	18
I.7.2. Interférence pulsée.....	19

I.8. Les sources potentielles d'interférence.....	19
I.8.1. Signaux hors bande.....	19
I.8.2. Signaux en bande.....	21
I.8.3. Signaux ultra-large bande.....	21
I.9. Les brouilleurs.....	22
I.10. Les impacts des interférences sur les récepteurs GNSS.....	22
I.11. Conclusion.....	23

Chapitre II

Algorithmes d'optimisation

II.1. Introduction.....	25
II.2. Algorithmes génétiques (AGs).....	25
II.3. Population initiale et codage.....	26
II.3.1. Codage binaire.....	27
II.3.2. Codage réel.....	27
II.4. Fonction d'évaluation.....	28
II.5. Opérateurs génétiques.....	28
II.5.1. Sélection.....	28
II.5.2. Croisement.....	29
II.5.3. Mutation.....	30
II.6. Fonctionnement général d'une AG.....	31
II.7. Optimisation multi-objectif (Optimisation multicritères).....	32
II.7.1. Problème multi-objectif.....	33
II.8. Techniques non élitistes.....	34
II.8.1. Algorithmes Génétiques Multi objectifs (MOGA).....	34
II.8.2. Frontière de Pareto.....	36
II.8.2.1. Optimum de Pareto.....	37
II.8.2.2. Domination et front de Pareto.....	37
II.9. NSGA (Non dominated Sorting Genetic Algorithm).....	38
II.10. Techniques élitistes.....	39
II.10.1. NSGAI (Non-dominated Sorting Genetic Algorithm II).....	39
II.10.1.1. Principe de l'algorithme NSGAI.....	40
II.10.1.2. Organigramme de l'algorithme NSGAI.....	41

II.11. Conclusion	42
-------------------------	----

Chapitre III

Intelligence artificielle

III.1. Introduction	45
III.2. Algorithmes d'apprentissage	45
III.2.1. Apprentissage supervisé.....	45
III.2.2. Apprentissage non supervisé.....	46
III.3. Les méthodes de classification traditionnelle.....	47
III.3.1. Réseaux de neurones multicouches MLP	47
III.3.1.1. Fonctions de transfert	49
III.3.1.2. Mise en œuvre des réseaux neuronaux	50
III.3.2. Réseaux de neurones à fonctions radiales de base RBF	51
III.3.3. Machine à vecteurs de support SVM	53
III.3.3.1. Principe de fonctionnement	53
III.3.4. Technique de validation	55
III.3.4.1. Validation simple.....	55
III.3.4.2. Classification binaire (Bi-classe).....	56
III.4. Détection des outliers	56
III.4.1. Machine à vecteur de support à une classe OCSVM.....	57
III.4.1.1. Le principe d'OCSVM	58
III.4.1.2. Comparaison avec OCSVM	59
III.5. Conclusion.....	60

Chapitre IV

Resultats et discussion

IV.1. Introduction	62
IV.2. Présentation du signal Galileo E5	62
IV.3. Méthode proposée	66
IV.3.1. OCSVM.....	66
IV.3.2. Configuration de l'algorithme génétique	69
IV.3.2.1 Structure chromosomique.....	69
IV.3.2.2. Fonctions de Fitness	70
IV.3.2.3. Opérateurs génétiques	72

IV.3.3. Description de l'algorithme proposée	72
IV. 4. Résultats de la simulation.....	74
IV.4.1. Classification OCSVM sans optimisation.....	75
IV.4.2. Optimisation simple objectif GA	76
IV.4.3. Optimisation Multi-objectif NSGAI.....	80
IV. 5. Conclusion.....	88
Conclusion générale.....	90
Bibliographie.....	93

Introduction générale

Introduction générale

Le besoin des systèmes GNSS (Global Navigation Satellite System) augmente rapidement dans plusieurs domaines dans notre vie quotidienne tels que : les systèmes de transport intelligents, les secteurs militaires, les applications de téléphones intelligents, l'agriculture et l'industrie connexe et bien d'autres. Le système de navigation par satellite GNSS a été intégré à la vie sociale et devenu incontournable et extrêmement populaire de part leur accessibilité et leur ouverture [1-3].

La navigation par satellites a commencé dans les années 1970 avec le développement du système américain GPS (Global Positioning System) et a considérablement évolué jusqu'à nos jours avec l'apparition et le développement d'autres systèmes de positionnement, comme par exemple le système russe GLONASS (Global Navigation Satellite System) ou le système européen Galileo. Ces systèmes GNSS ont été conçus essentiellement pour le positionnement d'un utilisateur au terrain équipé d'un récepteur radio fréquence. Le récepteur doit maîtriser et traiter les signaux émis par les satellites de la constellation afin d'estimer le triplet : position, vitesse et temps (PVT) à tout endroit de la terre [2, 4].

Malheureusement, étant donné que les satellites GNSS sont généralement placés entre 20000 et 30000 *kilomètres* au-dessus de la Terre, leurs signaux sont très faibles et ils sont généralement plus de 20dB en dessous du niveau du bruit, ce qui signifie les récepteurs GNSS sont vulnérables face aux interférences non intentionnelles générées par d'autres systèmes de communication et/ou des interférences intentionnelles connues également sous le nom de brouillage, notamment sur les bandes de fréquences de réception. Ainsi, la notion d'interférents aux signaux GNSS est rapidement devenue une problématique majeure, ces types de vulnérabilités peuvent perturber les services basés sur le GNSS se fait par loquer la réception des signaux de navigation à proximité et de dégrader les performances des récepteurs GNSS [1-3].

Dans ce contexte, nombreuses méthodes de détection et de réduire les interférences au niveau de récepteur GNSS ont été proposées dans la littérature scientifique. On peut citer le travail dans [3], dans leur étude ils ont proposé une méthode composé d'une série d'opérations: Premièrement, le signal reçu immergé dans un bruit gaussien blanc additif (AWGN), est multiplié, dans le domaine temporel, par une fenêtre Tukey. Ensuite, il est transformé dans le domaine DCT (Discrete Cosine Transform). Afin de détecter et de localiser les composantes fréquentielles de l'interférence, le signal transformé est partitionné en paquets non superposés. Le seuillage d'un paquet est obtenue grâce à l'utilisation du seuil universel de l'algorithme de

DONOHO. La variance est estimée sur la base de la théorie de l'échantillonnage statistique. La dernière étape consiste en l'application de la DCT inverse pour obtenir une bonne estimation du bruit gaussien.

Dans [5], l'auteur propose une technique d'atténuation des interférences basées sur les distributions Temps-fréquence (TF). Les signaux parasites sont généralement concentrés dans une zone limitée du plan TF bidimensionnel (2D), tandis que le bruit est réparti sur tout le plan. Par conséquent, les méthodes basées sur l'analyse et le traitement TF sont très attrayantes en tant que contre-mesures pour une grande variété de signaux interférents. Les techniques TF permettant de détecter et d'atténuer les brouillages. D'autres travaux, ont été proposées dans la littérature scientifique, telles que : les méthodes basées sur la DWT (Discret Wavelet Transform) [6], les filtres linéaire invariant dans le temps (LTI) [7], le filtrage adaptatif dans le domaine fréquentiel (FDAF) [8].

La résolution de problème des interférences par des méthodes exactes s'avère souvent difficile. La nécessité de trouver des solutions acceptables a entraîné le développement de plusieurs algorithmes d'approximation dont font partie les méta-heuristiques. Parmi les méta-heuristiques, les algorithmes génétiques (AG) s'avèrent être une approche particulièrement intéressante pour résoudre des problèmes d'optimisation. La puissance de la méthode AG à exploiter de vastes espaces de recherche et à générer plusieurs solutions de compromis en une seule étape d'optimisation. Les algorithmes génétiques sont des programmes informatiques qui tentent de simuler le processus de sélection naturelle dans un environnement hostile lié au problème à résoudre. Ce sont à la base des algorithmes d'optimisation stochastiques, mais qui peuvent également servir pour l'apprentissage automatique. Bien que les principes sous-jacents soient simples, ces algorithmes s'avèrent être des mécanismes de recherche généraux, puissants et robustes [9].

En fait, les problèmes d'optimisation rencontrés en pratique sont rarement uni-objectif. La plupart d'entre eux nécessitent l'optimisation simultanée de plusieurs objectifs souvent contradictoires. L'optimisation multi-objectif s'intéresse à la résolution de ce type de problèmes. Contrairement à l'optimisation uni-objectif où la solution optimale est unique et souvent clairement définie, les Problèmes multi-objectif présentent un ensemble de solutions dites optimales. Cet ensemble est généralement nommé ensemble des solutions Pareto optimales (PO). Nombreux algorithmes évolutionnaires multi-objectif (Multi-Objective Evolutionary Algorithms, MOEAs) utilisent la notion de Pareto dominance pour classer "les solutions optimale" qui ont obtenu la meilleure évaluation par rapport de l'objectif défini,

NSGAI (Elitist Non-dominated Sorting Genetic Algorithm) est l'un des algorithmes les plus connus dans cette catégorie. En général, ce sont des algorithmes élitistes. Ils gardent les meilleures solutions dans la population au fil des générations. NSGAI utilise outre une stratégie élitiste, un mécanisme de préservation de la diversité [9-11].

Concernant l'objectif général de ce travail, est de construire un modèle de classification robuste pour la détection et la suppression des différents types d'interférences dans les récepteurs GNSS. Cette stratégie, basée sur les algorithmes génétiques uni-objectif GA et multi-objectif NSGAI, avec l'intermédiaire d'une méthode de classification à base de classifieur OCSVM (One class support vector machine). GA ou NSGAI fonctionnent sur une population de solutions potentielles avec des fonctions d'évaluations (fitness function) pour déterminer automatiquement le paramétrage du classifieur adapté OCSVM et pour l'objectif de détecter et supprimer les valeurs aberrantes (les différents type d'interférence).

Les fonctions d'évaluations jouent un rôle prépondérant dans la possible réussite d'un GA ou NSGAI. Pour cela, nous utiliserons quatre principales fonctions, un critère lié à l'espace de représentation des échantillons qui est la distance de Bhattacharyya, et trois critères qui affectent directement sur les performances de classifieur OCSVM : la racine de l'erreur quadratique moyenne RMSE, le coefficient de corrélation R et l'erreur de classification ERR.

Le contenu de ce mémoire est organisé en quatre chapitres de la façon suivante :

Dans le premier chapitre introductif, nous présentons un ensemble de définitions et de concepts de base liés aux systèmes de navigation par satellites (GNSS). Nous nous intéressons aux signaux GNSS et les perturbations à leurs récepteurs, à cause des interférences.

Dans le deuxième chapitre, nous présentons les fondements nécessaires de la compréhension des algorithmes génétiques.

Dans le troisième chapitre, nous décrivons quelques méthodes de l'intelligence artificielle connue par les méthodes de classification (MLP, RBF, SVM, OSVM).

Le quatrième chapitre est consacré à la discussion des résultats obtenus. Nous présentons une nouvelle méthode d'identification et de suppression des interférences liés aux récepteurs GNSS basé sur deux méthodes : Application de classifieur classique OCSVM, et l'optimisation des coefficients de classifieur OCSVM par les algorithmes génétiques uni-objectif GA ou multi-objectif NSGAI.

Enfin, ce modeste travail est parachevé par une conclusion générale incluant certaines perspectives.

Chapitre I

Signaux GNSS

I.1. Introduction

Avec la globalisation du monde dans lequel nous vivons, nous ne pouvons plus ignorer les grands progrès technologiques et scientifiques que nous avons à notre disposition, même s'ils peuvent comporter quelques inconvénients, mais concentrons notre attention sur les avantages que ces nouvelles technologies offrent. Et depuis quelques années, la localisation est devenue de plus en plus importante dans la vie quotidienne, le concept de positionnement ainsi que la direction de mouvement de toute automobile, tout bateau, tout avion est plus développé grâce à l'apparition de nouvelles technologies de géo localisation. La géo localisation est le fait de localiser quelque chose ou quelqu'un dans un espace selon les systèmes de navigation GNSS [4].

Dans ce première chapitre introductif, la première section nous présentons les généralités principales et tout ce qui est important sur le GNSS (l'architecture du GNSS, les systèmes GNSS), En suit dans deuxième section nous concentrons sur les signaux GNSS et les interférences dans les systèmes GNSS.

I.2. Principes fondamentaux du GNSS

La navigation par satellite prend un virage très important ces dernières années, Les systèmes mondiaux de navigation par satellite (GNSS) désignent les systèmes de navigation à couverture mondiale capables de fournir à l'utilisateur une solution de positionnement et de synchronisation en trois dimensions par des signaux radio allant transmis par des satellites en orbite et avec l'apparence de la science du chronométrage, le niveau de précision de la lecture céleste a augmenté et jusqu'à aujourd'hui, le chronométrage joue l'un des rôles les plus importants dans les systèmes de navigation [12].

La navigation moderne n'évalue plus sa propre position en observant les étoiles, mais, de manière similaire, elle se fait en collectant des signaux radio transmis par des satellites à des positions connues. La position est ensuite récupérée en faisant une trilatération des différentes distances obtenues, en traitant les signaux reçus [2].

I.2.1. La position, la vitesse et le temps

Comme la représente la Figure I.1, la navigation par satellite est basée sur le principe de trilatération qui utilise des mesures de distances par rapport à des balises placées à des positions connues.

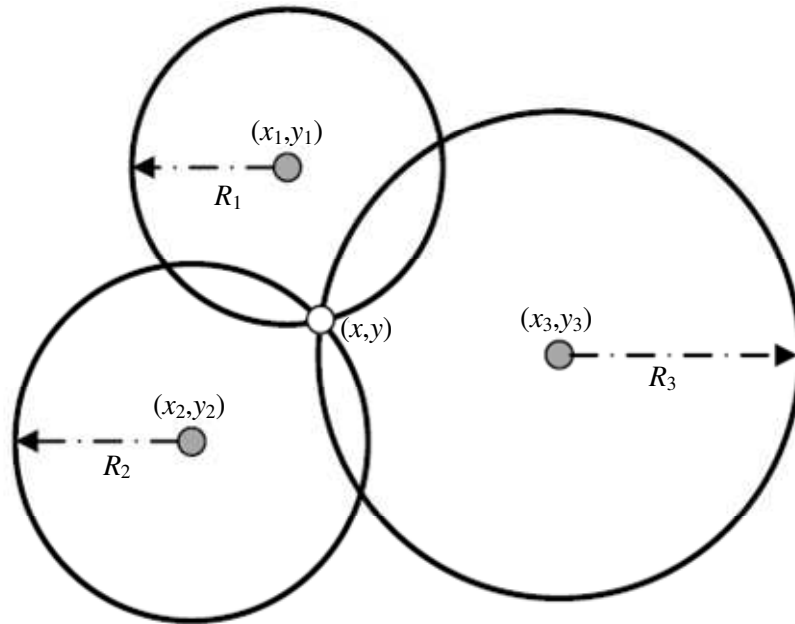


Figure I.1. Principe de trilatération

S'agissant de la navigation par satellite les balises sont des satellites déployés sur des orbites moyennes (30000 Kms) qui transmettent des signaux dans la bande 1.5 GHz. La mesure de distance est basée sur l'estimation du temps de propagation des signaux transmis. La synchronisation des satellites et la forme des signaux permet de déterminer le temps d'émission qui est le même quel que soit le satellite. La mesure du temps de réception est affecté d'un biais qui correspond à l'écart entre le temps satellite et le temps récepteur. La mesure exploitée est en pratique la mesure de pseudo distance notée, pour $i^{\text{ème}}$ satellite P_i :

$$P_i = c \tau_i \quad (\text{I.1})$$

Où τ_i représente le temps de propagation, c est la vitesse de la lumière dans le vide. La mesure de pseudo-distance s'exprime en pratique comme étant la mesure de distance affectée du biais de l'horloge du récepteur ou exprimé en mètre

$$P_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} + b_u \quad (\text{I.2})$$

Où (x_i, y_i, z_i) et (x_u, y_u, z_u) représentent respectivement les coordonnées du i -ème satellite et de l'utilisateur, exprimée habituellement dans un repère centré sur la terre, lié à la terre (repère ECEF). Cette équation fait apparaître les 4 inconnues que constituent les coordonnées de l'utilisateur et le biais d'horloge b_u . Le calcul de la position sur la base de ces mesures nécessite donc de bénéficier d'au moins 4 mesures [12].

Le signal GNSS utilisé pour la mesure du délai de propagation est un signal à spectre étalé obtenu par des méthodes d'étalement de spectre à séquence directe (DSSS). Un code pseudo-

aléatoire (PRN) est attribué à chacun des satellites. Il module le signal transmis, permettant un étalement du signal dans une bande de fréquence dépendant du rythme de ce code. L'estimation des paramètres des signaux reçus nécessite la mise en œuvre d'un filtre adapté qui réalise la corrélation du signal reçu avec un signal réplica dont les paramètres sont estimés par le récepteur. L'alignement du signal réplica avec le signal reçu nécessite d'estimer à la fois le délai de propagation du code transmis, la phase et la fréquence de la porteuse. La fréquence résulte de l'effet Doppler, permettant une observation de la vitesse satellite-récepteur [12].

$$U = \frac{du}{dt} = \frac{u(t_2) - u(t_1)}{t_2 - t_1} \quad (\text{I.3})$$

Par ailleurs une démodulation du signal en sortie du filtre adapté réalise l'extraction du message transmis par le satellite. Ce message contient toutes les informations nécessaires au calcul de la position des satellites, à la synchronisation des satellites, à la correction du temps de propagation. Utilisant la position calculée des satellites et les mesures de retard et fréquence, le récepteur élabore le calcul de la position et de la vitesse du récepteur. En particulier l'estimation de la position nécessite la résolution de l'équation (2). L'inversion de ce système d'équations non linéaires est basée sur la méthode des moindres carrés ou sur un filtre de Kalman [12].

I.3. Segments du système de navigation

Les GNSS se reposent sur trois piliers ou « segments » comme nous présente la figure suivante:

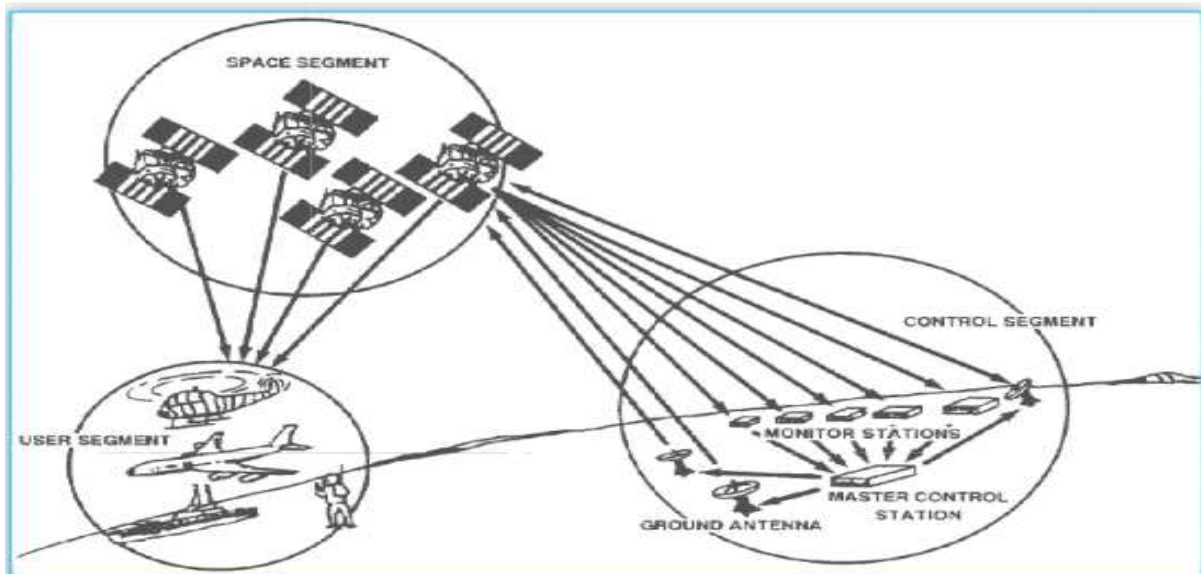


Figure I.2. Segment du système GNSS

I.3.1. Segment spatial

Le segment spatial correspond à la constellation de satellites à partir desquels les utilisateurs génèrent des mesures de télémétrie. Plusieurs constellations GNSS sont déjà déployées ; seules les constellations dont les signaux sont utilisés dans ce travail sont brièvement détaillées dans cette section [13].

I.3.2. Segment de contrôle

Le segment de contrôle est chargé de contrôler les caractéristiques des signaux émis par chaque satellite, de calculer les données d'éphémérides (éléments de Kepler) et les corrections d'horloge satellite, puis de télécharger le message de navigation vers chaque satellite. Il est composé d'un réseau au sol de stations de contrôle principales, de stations de téléchargement et de stations de surveillance. Les stations de surveillance installées dans le monde entier surveillent les signaux et l'état des satellites et transmettent ces 9 informations aux stations de contrôle principales. Le contrôle maître analyse les signaux puis transmet les corrections d'orbite et de temps aux satellites via la station de téléchargement [13].

I.3.3. Segment d'utilisateurs

Le segment utilisateur est composé de tous les récepteurs utilisateurs du GNSS. Le récepteur utilisateur traite le signal transmis depuis le segment spatial pour obtenir une solution de positionnement, de vitesse et de synchronisation généralement appelée solution PVT. Les récepteurs GNSS sont utilisés dans de nombreux équipements de notre vie quotidienne : voiture, Smartphones, appareils liés au sport, etc. De plus, la tendance générale à la miniaturisation des composants conduit à la multiplication des récepteurs à bas coût [13].

I.4. Différents système mondial de navigation

Les applications et les services fournis par cette infrastructure, en les appliquant en géodésie, en cartographie et dans multiples disciplines civiles et militaires, représentent une expédition inédite de vie moderne. Cette nouvelle technologie avait une importance stratégique en raison de son usage militaire; cependant, elle a pris une importance politique et économique surtout après la tendance mondiale de construire un système mondial de navigation par satellites GNSS qui appréhendera tous les systèmes de navigation par satellites existants et prévus. Les sous-sections suivantes vont décrire les systèmes de positionnement par satellites dont nous avons parlé [14].

I.4.1. Système de positionnement global GPS

Appelé officiellement NAVSTAR (Navigation System by Timing And Ranging) est le système militaire américain de radio positionnement par satellite conçu et mis en service par la NGA (National Geospatial Intelligence Agency), C'est le premier système global de

positionnement par satellite opérationnel. La constellation GPS est composée de 31 satellites placés sur des orbites quasi-circulaires d'altitude de l'ordre de 20200 *km* et de période légèrement inférieure à 12 *h*. Elle permet d'assurer la visibilité de 4 à 8 satellites avec une élévation supérieure à 15° en tout point du globe. On distingue plusieurs classes de satellites, en fonction de leur génération (Bloc I en 1978 au Bloc 3, 2014). Elle permet de piloter et de surveiller la constellation GPS. Il est composé d'une quinzaine de stations réparties sur toute la planète et elle regroupe l'ensemble de tous les utilisateurs civils et militaires du système GPS, et elle émet sur deux fréquences L1 (1575.43MHz) et L2 (1227.60 MHz). Les utilisations principales sont le positionnement absolu ou relatif par le code et/ou par la phase et la datation précise d'événements [15].

Il deux types de services, basés sur des mesures de code, sont offert aux utilisateurs :

- Le SPS (Standard Positioning Service) : c'est le service de positionnement qui est accessible à tout utilisateur disposant d'un récepteur GPS. C'est un service gratuit et anonyme.
- Le PPS (Precise Positioning Service) : c'est le service de positionnement précis du GPS; il nécessite d'obtenir des clés de décodage auprès du DoD américain. Il est réservé aux militaires américains et à leurs alliés.

I.4.2. Système de navigation Russe GLONASS

GLOBAL'naya NAVigatsionnaya Sputnikovaya Sistema en russe romanisé est un système de positionnement global par satellite développé par l'URSS durant la guerre froide et maintenu par la Russie depuis la chute du bloc soviétique. Tout comme GPS, ce système a été développé dans un cadre militaire. La constellation Glonass est composée de 27 satellites placés sur des orbites quasi-circulaires d'altitude de l'ordre de 19 000 *km* et de période légèrement inférieure à 12 *h*. Cette constellation permet d'assurer la visibilité d'au moins 5 satellites sur 99% de la surface de la Terre. Le composant sol de Glonass est actuellement constitué d'une dizaine de stations situées principalement sur le territoire russe. L'ensemble de tous les utilisateurs civils et militaires du système Glonass. Les utilisations principales sont le positionnement absolu ou relatif par le code et/ou par la phase et la datation précise d'événements [15].

I.4.3. Système de navigation Galileo

Galileo est le futur système européen de navigation par satellite, lancé par l'UE (Union Européenne) et l'ESA (European Space Agency, l'agence spatiale européenne). L'objectif pour l'Europe est donc de disposer de son propre système global de navigation par satellite et donc d'être indépendant vis à vis des systèmes GPS et GLONASS. La constellation Galileo

est composée de 30 satellites placés sur des orbites quasi-circulaires d'altitude de l'ordre de 23000 km et de période légèrement supérieure à 14 h [15]. Les satellites GALILEO utilisent quatre porteuses nommées (E5a, E5b, E6, E2-L1-E1), pour transmettre un total de 10 signaux, centrés ces bandes fréquences sur les valeurs L1 (1575,42 MHz), E6 (1278,75 MHz) et E5 (1191,795 MHz) [16].

Il sera composé d'une cinquantaine de stations réparties sur toute la planète et il assure 5 niveaux de services sont:

- Le service libre, OS (Open Service) : en libre accès pour tout possesseur d'un récepteur compatible Galileo, il fournira des données de positionnement, de navigation et de datation.
- Le service commercial, CS (Commercial Service) : en échange d'une redevance, il offrira de nombreux services à valeur ajoutée (garantie du service, intégrité et continuité des signaux, données de navigation métrique).
- Le service de sûreté de la vie, SOL (Safety-of-Life) : il délivrera un service sécurisé et intègre pour des applications critiques tels que le transport.
- Le service public réglementé, PRS (Public Regulated Service) : il s'adressera en priorité aux utilisateurs d'applications gouvernementales de sécurité.
- Le service de recherche et secours, SAR (Search And Rescue) : pour la recherche et le sauvetage. il permettra de localiser l'ensemble du parc des balises Cospas-Sarsat utilisées pour les signaux de détresse.

I.4.4. Système de navigation BEIDOU

Egalement nommé COMPASS est un système de navigation et de positionnement par satellites chinois en cours, qui désigne une constellation développée en trois générations par une couverture mondiale par rapport à d'autres systèmes, l'une des caractéristiques mise en évidence du système Beidou est qu'il est capable de communiquer des messages courts. Les bandes de fréquences utilisées par COMPASS sont E1/L1, E5 et E6, ce qui engendre quelques conflits avec le développement de GALILEO sur ces mêmes bandes [2]. La Figure I.3 illustre les différents systèmes de positionnement.

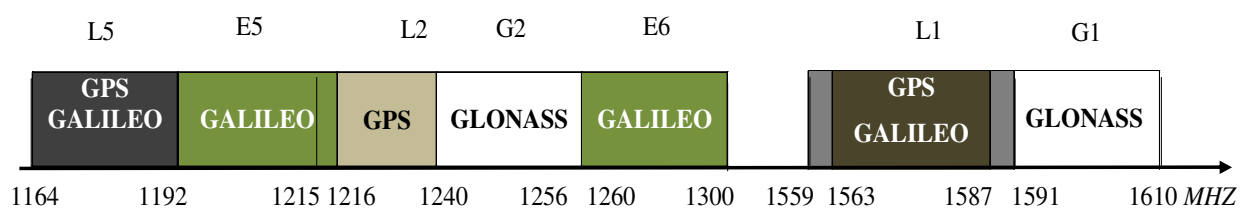


Figure I.3. Phase d'opérationnalité des différents systèmes GNSS existants.

I.5. Les signaux GNSS

Les satellites émettent des signaux radiofréquence contenant des informations qui vont permettre à un utilisateur les recevant de se géo-localiser.

I.5.1. Les principales composantes des signaux GNSS

Comme le montre la figure suivante, les principales composantes des signaux GNSS sont le code PRN propre à chaque satellite, la porteuse sur laquelle le code est modulé, et le message de navigation contient ce qu'on appelle les éphémérides du satellite qui permettent de connaître, en particulier, la position du satellite à un instant donné [1].

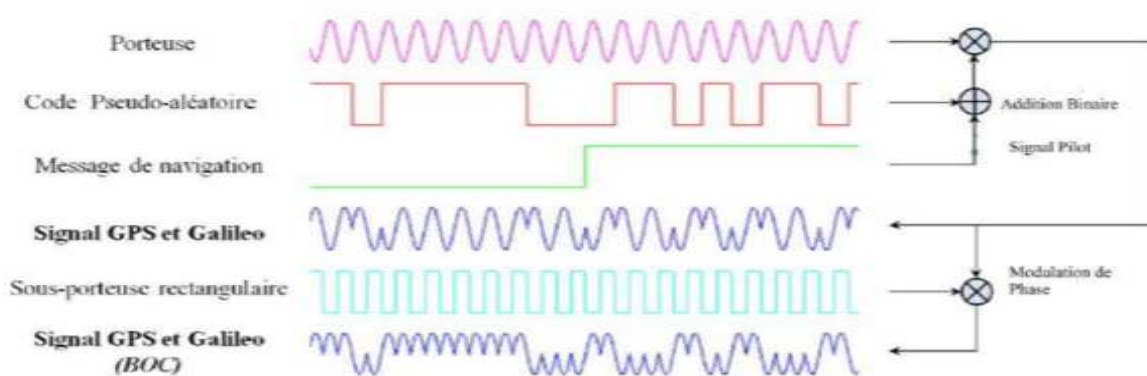


Figure I.4. Structure générales des signaux GNSS.

I.6. Chaîne de réception GNSS

Le suivi dans le temps du délai de propagation et le calcul de la position sont réalisés grâce à une chaîne de réception [17]. Dans la Figure I.5, nous avons représenté la chaîne de récepteur GNSS.

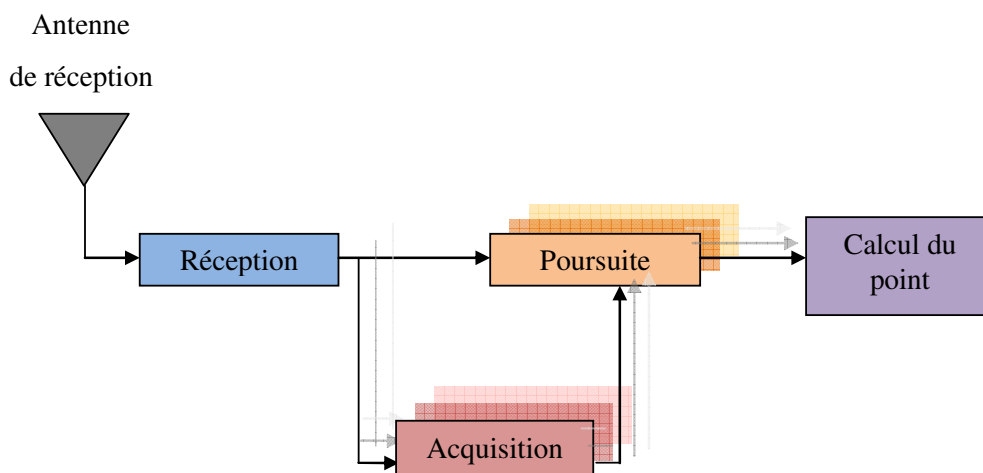


Figure I.5. Chaîne de récepteur GNSS.

Les catégories de récepteurs sont aussi variées que les domaines d'applications, citer [18]:

- Les récepteurs grand public.
- Les récepteurs certifiés pour les transports.
- Les récepteurs de qualité géodésique.
- Les récepteurs militaires.

I.6.1. Architecture d'un récepteur GPS

Un récepteur GPS est de nature numérique constitué de différents blocs permettant d'obtenir les différentes mesures nécessaires au calcul de la position [19]:

- une antenne, permettant de transformer l'onde électromagnétique en signal électrique
- un bloc frontal radiofréquence, composé de hardware permettant la mise en bande de base, l'amplification et la numérisation du signal
- un bloc d'acquisition, permettant d'identifier le satellite dont le signal provient et d'avoir une première estimation des paramètres d'intérêt (temps de propagation, fréquence, phase du signal)
- un bloc de poursuite, permettant de suivre au cours du temps l'évolution des différents paramètres d'intérêt
- un bloc de démodulation des données, permettant de récupérer les éphémérides nécessaires au calcul de la position, de la vitesse du satellite, et des paramètres permettant la correction des erreurs atmosphériques
- un bloc navigateur de Position Vitesse Temps, ou Position Velocity Time (PVT), permettant l'estimation des différents paramètres d'intérêt de l'utilisateur.

Cette configuration est représentée dans la Figure I.6.

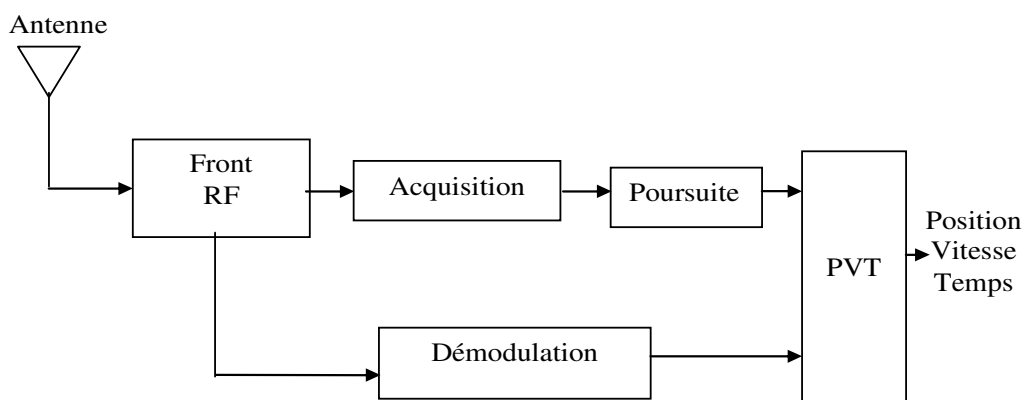


Figure I.6. Schéma global d'un récepteur GPS.

I.6.2. Acquisition du signal GNSS

Afin de lancer le processus de poursuite du signal de navigation, le récepteur effectue une étape d'acquisition. Cette opération primordiale a pour but de détecter la présence du signal d'un satellite dans l'ensemble des signaux reçus par l'antenne du récepteur. L'acquisition permet également de déterminer le décalage du code et l'écart de fréquence provoqués par la propagation du signal et l'effet Doppler. De manière générale, l'acquisition est un processus de détection des signaux en (temps / fréquence), pour fournir une idée sur la position de récepteur grâce à une première estimation approximative des paramètres du signal reçu qui sont modifiés par rapport à ceux du signal émis par le satellite [2, 20].

Son principe consiste à générer une réplique du code transmis qui doit être décalée jusqu'à ce qu'elle soit alignée sur le code reçu. L'alignement correct est identifié par la mesure de la puissance de sortie des corrélateurs (temps/fréquence) [18].

La Figure I.7 montre un exemple d'acquisition du signal restauré, les deux paramètres à estimer prennent les valeurs: Le code de décalage et la fréquence doppler. Le système de navigation par satellite GPS est basé sur la technique DS-CDMA où les signaux transmis par les différents satellites sont identifiés par différents codes pseudo aléatoires PRN. Ces séquences sont utilisées pour étaler les données de navigation à transmettre. Le récepteur exploite donc les propriétés de corrélation des codes d'étalement pour mesurer le temps de propagation de l'onde porteuse.

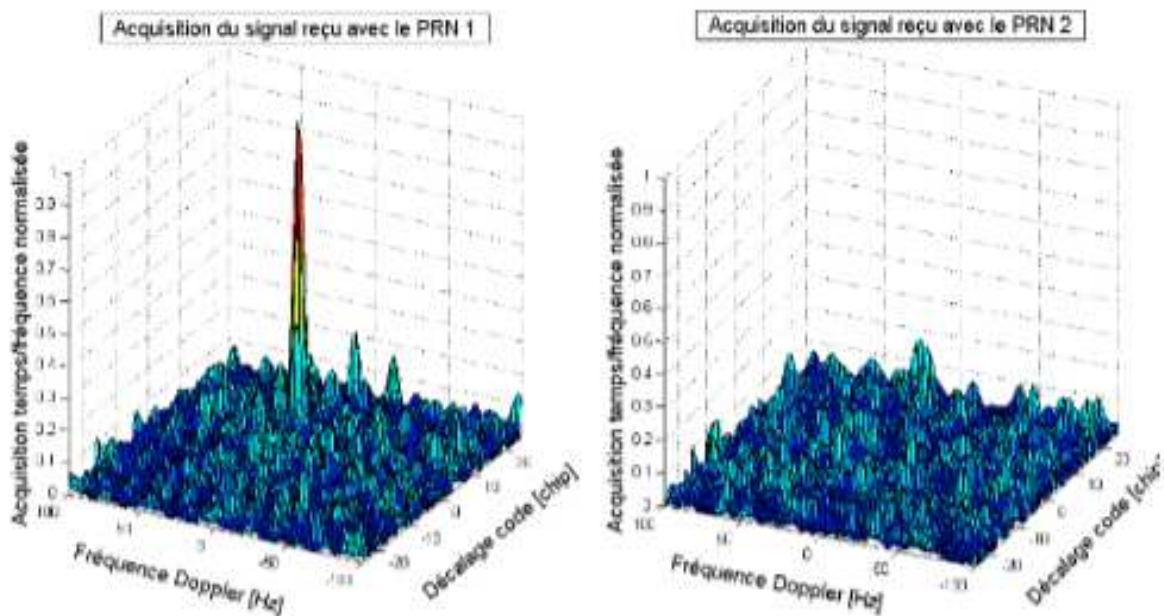


Figure I.7. Exemples de résultat d'acquisition.

Les trois techniques d'acquisition des signaux GNSS sont les suivantes [18]:

- L'acquisition par recherche en série.
- L'acquisition par recherche parallèle sur le retard de code par (FFT)
- L'acquisition par recherche parallèle sur le décalage de fréquences par (FFT).

I.6.2.1. Mode d'acquisition série

L'acquisition par recherche en série est une méthode souvent utilisée pour l'acquisition des systèmes basés sur le "Code Division Multiple Access" (CDMA). Comme le GPS est un système CDMA, les récepteurs hardware sont configurés par ce type d'acquisition. Dans le cas de la recherche en série, les décalages de code et de fréquence sont évalués en série jusqu'à ce que les valeurs correctes soient trouvées [21].

Dans la Figure I.8, nous avons montré le principe d'acquisition par recherche série. Le signal d'entrée est multiplié avec un code pseudo aléatoire PRN. L'étape suivante consiste à générer deux porteuses par un oscillateur local avec une différence de phase 90° entre les deux. On obtient ainsi les deux composantes en phase I et en quadrature Q , puis on passe à une intégration ou une sommation de tous les points qui correspondent à la longueur des données traitées. L'étape finale consiste à faire la somme des deux valeurs de I et Q . Si le code local généré est bien aligné avec le code dans le signal d'entrée et la fréquence générée localement par un oscillateur est égale à la fréquence du signal transmis, la sortie sera plus élevée que les autres décalages (pic de corrélation égal à sa valeur maximale).

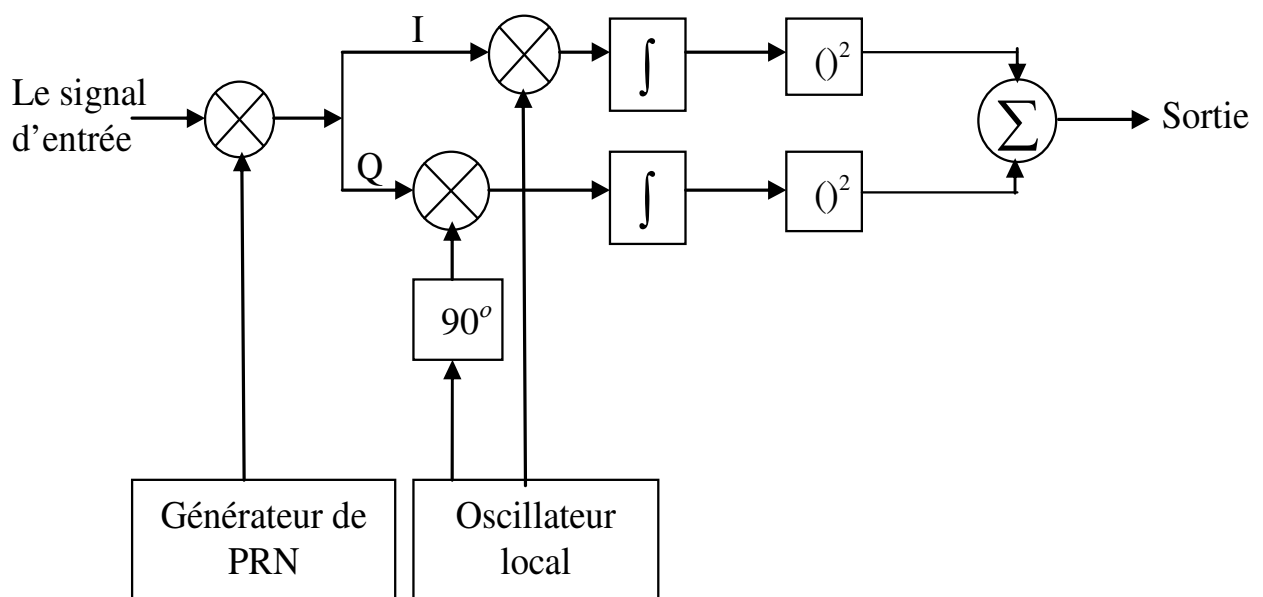


Figure I.8. Principe d'acquisition par recherche série.

I.6.2.2. Acquisition par recherche parallèle sur le retard de code par (FFT)

La Figure I.9, présente la procédure de recherche parallèle dans les espaces code et fréquence. L'objectif de cette méthode consiste à proposer une nouvelle architecture d'acquisition afin de résoudre les défauts des méthodes d'acquisition en série et en parallèle, comme le temps d'exécution (grand nombre d'opérations) pour la première et le manque relatif de précision pour la deuxième. Il s'agit toujours d'une corrélation entre le signal reçu et une réplique du code générée par le récepteur. Afin de gagner en temps d'exécution (diminuer le nombre d'opérations) [18].

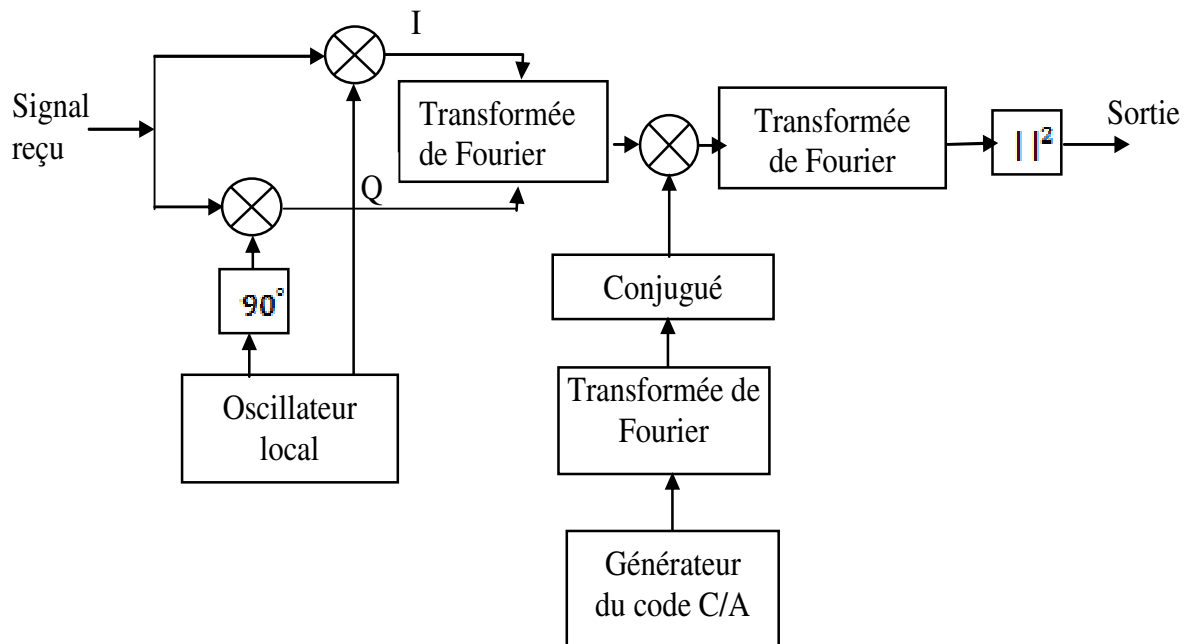


Figure I.9. Acquisition en mode parallèle du code et de la phase.

I.6.2.3. Acquisition par recherche parallèle sur le décalage de fréquences par (FFT)

La méthode de recherche en série est très coûteuse au niveau temps. C'est pourquoi, il est nécessaire d'éliminer un des deux paramètres recherchés, ou le cas échéant le faire en parallèle. La deuxième méthode proposée dans ce contexte consiste à paralléliser la recherche d'un seul paramètre. Ainsi, on n'effectue plus le balayage de l'espace fréquentiel, la première partie de cette méthode est similaire à celle de la méthode de recherche en série, où le signal reçu est multiplié par un code PRN correspondant à un satellite bien défini. Le signal résultant est alors obtenu dans le domaine fréquentiel par la Transformée de Fourier (voir la Figure I.10) [18].

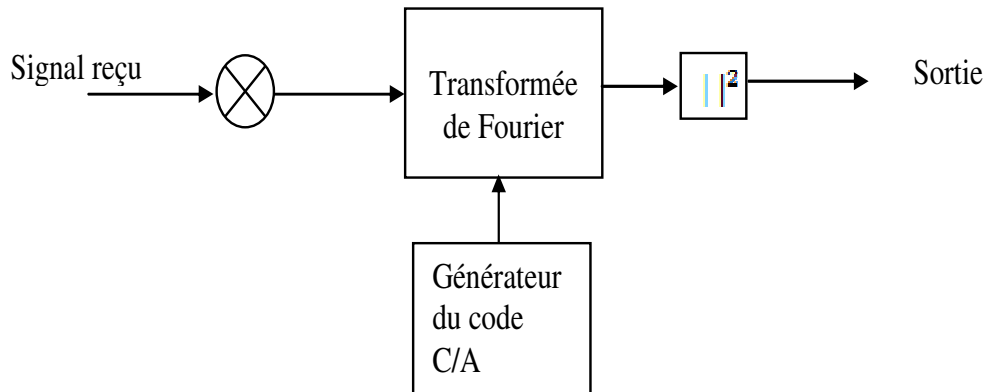


Figure I.10. Recherche en parallèle du Doppler et du retard de code.

Le but de cette opération est double [20]:

- détecter la présence d'un satellite dans l'ensemble des signaux captés par l'antenne du récepteur ;
- déterminer, pour un satellite dont la présence a été détectée, le retard sur le code et la fréquence Doppler à la réception du signal associé afin d'initialiser les boucles de poursuite.

I.6.3. Poursuite du signal GNSS

Après l'étape d'acquisition qui a pour but de détecter la présence d'un satellite et d'estimer le retard et la fréquence Doppler initiaux du signal à la réception, le récepteur bascule en mode poursuite "poursuivre" permet cette étape de poursuite suivre l'évolution du retard sur le code, de la fréquence et de la phase porteuse engendrée par les mouvements relatifs entre le satellite et l'utilisateur et ainsi de mettre à jour le calcul du point. Cette étape est réalisée grâce à des structures qui vont suivre les variations des paramètres nécessaires au calcul de la position grâce à la corrélation du signal reçu par des répliques locales générées par le récepteur. Différentes mesures peuvent être réalisées par le récepteur lors de la poursuite pour estimer les différentes distances utilisateur/satellite [20].

La Figure I.11, représente le système de poursuite, les trois types de boucles sont intégrées (boucle à verrouillage de fréquence FLL, boucle à verrouillage de phase PLL, et la boucle à verrouillage de délai de code DLL). En effet, la sortie de la boucle de poursuite de la porteuse sera utilisée lors de la démodulation de la boucle de poursuite du code. Les boucles de poursuite PLL et FLL utilisent le code synchronisé (code "prompt"), fourni par la boucle de code pour le dés-étalement du signal [2].

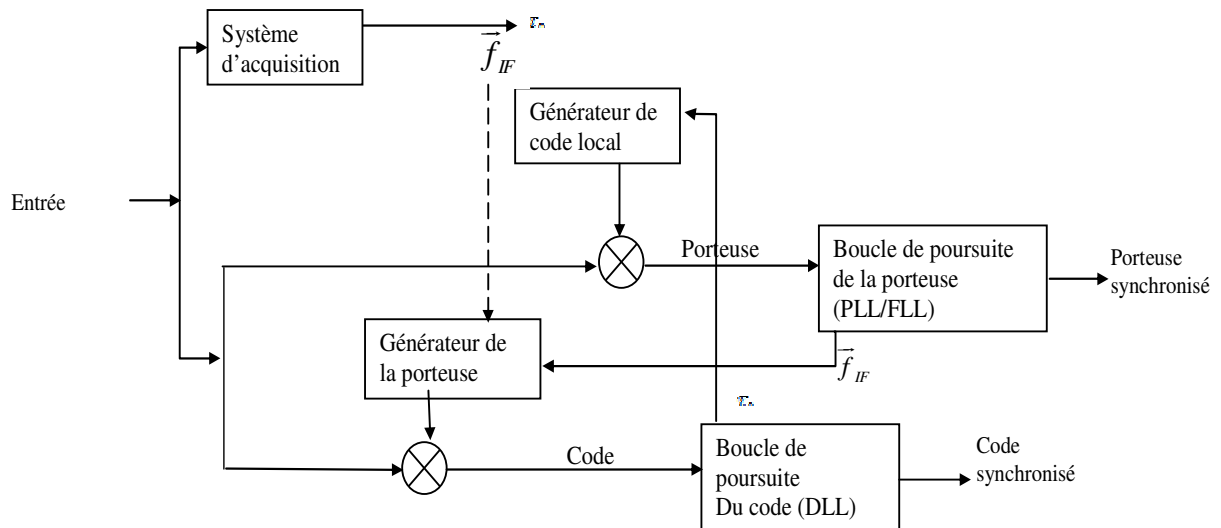


Figure I.11. Schéma fonctionnel d'une boucle de poursuite du code et de la porteuse.

I.7. Classification de la source interférente

La Figure I.12, illustre la classification des principales perturbations des récepteurs GNSS prend en compte les aspects hétérogènes. Les types d'émission peuvent être définis comme étant intentionnels (Jamming) ou non intentionnels. Les premiers sont courants pour les scénarios militaires, même si le brouillage des applications civiles commence à être couramment en raison de la disponibilité des dispositifs de brouillage sur le marché. En outre, un grand nombre de systèmes de communication présents dans notre vie quotidienne émettent une puissance qui pourrait interférer avec la bande L GNSS, en raison des émissions hors bande de ces systèmes électroniques [22].

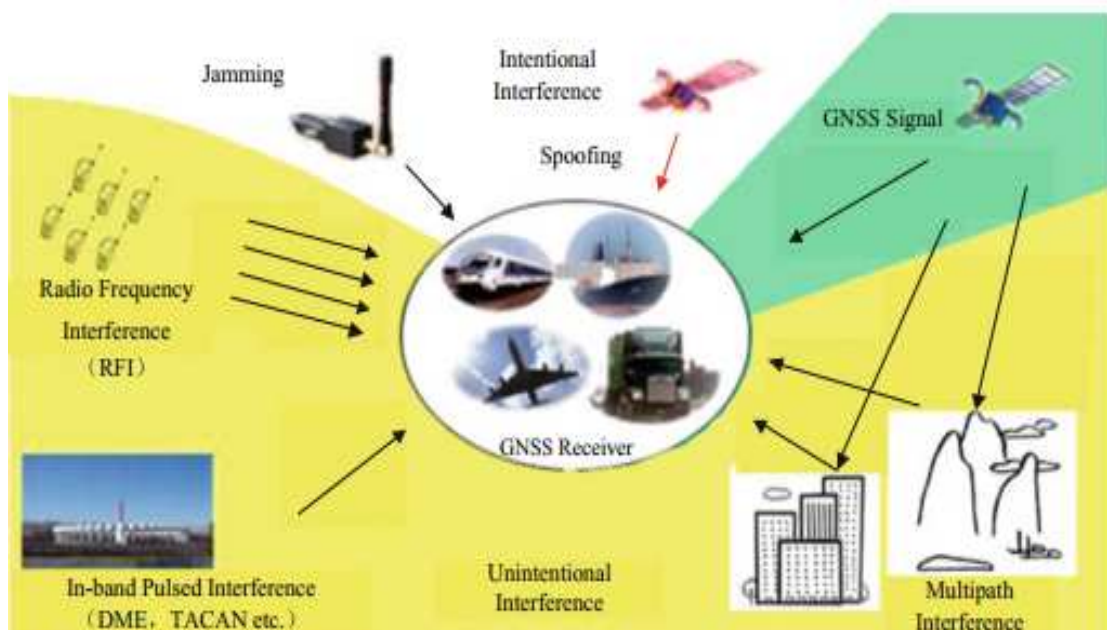


Figure I.12. Différentes type d'interférences GNSS.

I.7.1. Caractéristiques spectrales d'interférences

Les interférences GNSS se présentent sous de nombreuses formes et classifications telles que la bande passante du récepteur Intérieur (dans la bande) ou la bande passante du récepteur extérieur (hors bande).

- **Interférence peut être une onde continue (SCWI) :** Représente la limite ultime de NBI et apparaît comme une seule tonalité dans le domaine fréquentiel [22]. Ce type d'interférence est modélisé par l'équation suivante:

$$J_{SCWI}(t) = J \times \cos(2\pi(f_{IF} \pm \Delta f)t + \Phi) \quad (I.4)$$

où J est l'amplitude du signal d'interférence, f_{IF} est la fréquence de l'interférence reçue, Δf est la différence par rapport à la fréquence centrale du signal GNSS, et Φ est une phase initiale aléatoire uniformément distribuée dans l'intervalle $[-\pi, \pi]$.

- **Interférence a bande étroite (NBI) :** Les méthodes de génération d'interférences à bande étroite sont également différentes, peuvent utiliser des interférences à bande étroite autorégressives et peuvent également être équivalentes à la superposition de plusieurs signaux d'interférence à fréquence unique; aussi, le signal d'interférence est considéré comme un signal modulé par un filtre à bande étroite. L'interférence à fréquence unique est un signal spécial à bande étroite, qui est essentiellement une onde continue à fréquence unique. Ce signal d'interférence peut interférer avec un ou plusieurs points de fréquence dans le système de navigation par satellite [23]. L'expression est donnée par la relation suivante:

$$J(t) = \sqrt{2p_j} \cos(\omega_j t + \phi_j) \quad (I.5)$$

où ω_j est la fréquence angulaire de signal interférence, p_j est la puissance, et ϕ_j est la phase aléatoire répartie sur $[0, 2\pi]$.

- **Interférence à bande large (WBI) :** Dans le domaine du spectre, l'occupation spectrale est comparable par rapport à la largeur de bande du signal GNSS ($B_{int} \approx B_{GNSS}$).

Il peut également s'agir de formes d'onde modulées ou de bruit blanc ainsi que pulsé ou continu dans le temps comme suit :

- **Interférences à ondes continues (CW) et à bruit blanc Gaussien additif (AWGN) :** Les signaux CW et AWGN sont parmi les formes d'interférences les plus connues rencontrées par un récepteur. Les interférences CW peuvent provenir de plusieurs sources. Une interférence indésirable est souvent causée par les harmoniques externes à la bande des amplificateurs ou des émetteurs. Les brouilleurs GNSS COTS, qui sont devenus de

plus en plus répandus, utilisent également des signaux CW pour interrompre la réception du signal GNSS. Le module AWGN peut également être utilisé pour simuler les bruits d'antenne et d'amplificateur ou pour évaluer la performance en présence de brouilleurs de bruit.

➤ **Interférence à onde continue multiple (MCWI) :**

$$J_{MCWI}(t) = \sum_{i=1}^N J_i \times \cos(2\pi(f_{IF} \pm \Delta f_i)t + \Phi_i) \quad (I.6)$$

où: J_i , Δf_i et Φ_i sont respectivement l'amplitude, la différence de fréquence et la phase aléatoire de la $i^{\text{ème}}$ tonalité.

I.7.2. Interférence pulsée

Les signaux parasites pulsés sont caractérisés par un état marche-arrêt de courte durée (de l'ordre de la microseconde), qui alterne dans le domaine temporel. Ce type de signal d'interférence est typique des scénarios aéronautiques, où plusieurs services de radionavigation aéronautique (ARNS) diffusent des signaux pulsés puissants dans une bande passante partagée avec certains des systèmes de navigation par satellite [2, 22].

La modalisation de l'interférence pulsée (DME) est donnée par l'équation suivante :

$$J_{DME}(t) = \sqrt{P} \sum_{i=1}^M \left(e^{-\frac{\alpha(t-t_k)^2}{2}} - e^{-\frac{\alpha(t-\Delta t-t_k)^2}{2}} \right) \times \cos(2\pi(f_{IF} \pm \Delta f)t + \Phi) \quad (I.7)$$

où : P est la puissance du pic du DME/TACAN à l'antenne, t_k est l'ensemble des temps d'arrivée des paires d'impulsions, f_{IF} est la fréquence de l'interférence reçue, Φ est la phase de la porteuse du signal d'interférence.

I.8. Les sources potentielles d'interférence

L'interférence potentielle peut partager les fréquences GNSS (RFI dans la bande) ou être loin de la porteuse GNSS (RFI hors bande). Il n'y a presque pas d'émissions autorisées dans les largeurs de bande GNSS. Cependant, les interférences viennent principalement des émissions non essentielles de systèmes hors bande, qui génèrent l'harmonique qui entre en collision avec les largeurs de bande GNSS [2].

I.8.1. Signaux hors bande

Dans les sections suivantes, quelques-unes des principales interférences potentielles hors bande les sources sont analysées [2, 22].

➤ **Chaînes TV analogiques :** Les émissions TV sont de véritables sources d'interférences pour un récepteur GNSS. Ils peuvent se manifester par des interférences à large bande et à bande étroite: la voiture vidéo les rieurs sont considérés comme des signaux à bande

moyenne / large, tandis que le son. Les rieurs sont considérés comme CWI. Dans le signal de diffusion TV, VHF et UHF des bandes sont utilisées. Les harmoniques de ces bandes générées par la télévision au sol les émetteurs peuvent générer des interférences potentiellement dangereuses pour le GNSS récepteurs

- **Signaux DVB-T:** La norme DVB a été définie (depuis 1993) dans le cadre d'une initiative impliquant plus de 300 membres européens et extra-européens. Le projet DVB etc. a harmonisé les stratégies d'introduction de la télévision numérique et des nouveaux services interactifs multimédias sur des réseaux de transmission. Il a également défini les spécifications techniques.
- **VHFCOM:** D'autres systèmes de communication VHF peuvent être considérés comme dangereux pour un GNSS récepteur. La bande VHF (118–137 MHz) contient 760 canaux espacés Par 25 kHz, et il est couramment utilisé par les communications du contrôle de la circulation aérienne (ATC). Les harmoniques sont considérées comme NBI avec une bande passante d'environ 25kHz. Les canaux VHF, centrés sur 121,150, 121,175 et 121,200 MHz, Ont la 13e harmonique dans la bande passante GPS, alors que les canaux centrés sur 131,200, 131,250 et 131,300 MHz sont dangereux pour leur 12e harmonique.
- **Harmoniques FM:** De plus, les petites bandes de fréquences à l'intérieur de la bande FM (87,5–108 MHz) ont des harmoniques qui tombent dans les bandes GNSS. Les canaux à 104,9 et 105,1 MHz Ont leurs 15 harmoniques près des largeurs de bande GPS et Galileo.
- **Appareils électroniques personnels:** Les appareils électroniques personnels (PED) à proximité d'un récepteur GNSS peuvent la perturbation de la réception du signal GNSS. Les PED comprennent les téléphones portables, les téléavertisseurs, radios bidirectionnelles, jouets télécommandés, ordinateurs portables et bien d'autres. Un plus grand nombre On s'attend à ce que plusieurs PED incluent, à l'avenir, l'ultra-large bande (UWB) transmission qui permet le développement d'appareils personnels à haut débit.
- **SATCOM:** Les communications par satellite (SATCOM) fonctionnent dans les bandes de fréquences 1626–1660,5 MHz avec canaux espacés de 0,75 MHz et bande passante de 20 kHz. La transmission multi-porteuse dans un service SATCOM produit de l'intermodulation produits qui peuvent tomber dans la bande GNSS.
- **Service mobile par satellite (MSS):** Le système du service mobile par satellite (MSS) peut générer deux inter-Menaces à un récepteur GNSS. Les stations terriennes mobiles du SMS utilisent le Bande 1610-1660,5 MHz, introduisant potentiellement une puissance large bande dans le Bande GNSS Interférence de téléphone portable En général, aucune conséquence directe des téléphones mobiles sur un récepteur GNSS ont été rapportés dans

la littérature jusqu'à présent. Certaines informations sont disponibles pour équipement de navigation pour avions, où un récepteur GPS est couramment utilisé une enquête sur les rayonnements non essentiels de six technologies de téléphonie sans filages est décrit, analysant les effets sur les systèmes de l'avion, parmi lesquels un GPS. Les tests sont effectués à partir d'une chambre semi-échoïque et de réverbération. Bers utilisant des technologies de téléphonie sans fil avec différentes fréquences de transmission ciels et différentes antennes de réception.

I.8.2. Signaux en bande

Certaines sources d'interférences diffusent des signaux dont la fréquence porteuse est attribuée dans les bandes GNSS, et donc ils génèrent des interférences dans la bande. Cependant, le niveau d'acceptable l'interférence est définie pendant la phase de conception des systèmes et, jusqu'à présent, le niveau acceptable a été le résultat de négociations internationales, de discussions et accords (considérez, par exemple, l'accord d'interopérabilité GPS/Galileo). Dans cette section, l'accent est mis sur les systèmes terrestres non GNSS. Une courte description des systèmes émettant de la puissance les plus pertinents dans un ou plusieurs des bandes GNSS sont fournies dans les sous-sections suivantes. Systèmes de communication aéronautique militaire / civile. Les systèmes de communications militaires peuvent être considérés comme des brouilleurs dans la bande raison de la bande de signal utilisée par les systèmes. Les bandes Galileo E5a et E5b, situés entre 1164 et 1214 MHz, occupent les fréquences déjà utilisées pour les Services de radionavigation nautique (ARNS) tels que la navigation aérienne tactique (TACAN), équipement de mesure de distance (DME) et surveillance secondaire radar (SSR), ainsi que par le DoD Joint Tactical Information Distribution System (JTIDS) et le système de distribution d'informations multifonction (MIDS) [2, 22].

I.8.3. Signaux ultra-large bande

La définition d'UWB inclut tout signal occupant plus de 500 MHz entre 3,1 et 10,6 GHz. Les signaux UWB sont apparus en tant qu'une solution potentielle pour une consommation à faible complexité, à faible coût et à faible consommation d'énergie, et connectivité sans fil à haut débit de données. Les technologies basées sur UWB offrent une communication simultanée à haut débit, avec des débits de transmission de données de 100 à 500 Mbps à des distances de 2 à 10 m en même temps une puissance rayonnée moyenne de quelques centaines de microwatts Les systèmes UWB utilisent souvent les modulations : modulation en position d'impulsions (PPM) et la modulation d'impulsions en amplitude (PAM).

Le signal UWB est généré en utilisant des impulsions sous-nanosecondes qui répartissent l'énergie du signal sur une large bande de fréquences. Ainsi, ces systèmes utilisent des signaux de faible puissance mais avec une bande passante extrêmement large. Cet aspect est critique pour les systèmes tels que les GNSS dont la puissance du signal est très inférieure au bruit de fond [2].

I.9. Les brouilleurs

Sont également appelés dispositifs de confidentialité personnels (PPD) est capable de brouiller ou de bloquer les signaux GNSS, ils sont destinés à de faire perdre la poursuite et d'empêcher la réacquisition du signal dans une zone limitée. Le brouillage fait référence à la transmission intentionnelle d'énergie RF pour empêcher le service de navigation en masquant les signaux GNSS avec du bruit. Les brouilleurs peuvent causer de graves dommages si leurs signaux ne sont pas correctement détectés et les effets atténués dans les récepteurs des utilisateurs [2].



Figure I.13. Les brouilleurs.

I.10. Les impacts des interférences sur les récepteurs GNSS

Étant donné que GNSS utilise généralement la technologie multi-accès CDMA, un code de propagation périodique de pseudo-accès est utilisé pour la modulation de spectre étalé sur l'émetteur et le même code d'épandage de pseudo-random est utilisé pour la dés-marquage en fonction de la corrélation sur le récepteur. Après la modulation du spectre de propagation, l'énergie de l'information est répartie uniformément sur une largeur de bande plus large et la puissance de densité spectrale diminue. Après la propagation du signal de spectre étalé, le signal à large bande est reconstitué dans un signal de bande étroite et la densité spectrale de puissance augmente. Par conséquent, la technologie Spread Spectrum transporte elle-même une certaine capacité d'atténuation des interférences. Néanmoins, la distance entre le satellite

et le récepteur est très loin et la puissance émettrice du satellite est plutôt faible. Lorsque le récepteur est sous les impacts de diverses intentionnelles et Interférences RF involontaires, le signal GNSS ne peut pas être extrait du bruit, entraînant enfin une diminution de la précision de positionnement et de la perte de verrouillage de la boucle de suivi. Parfois, même le satellite ne peut pas être acquis normalement. La direction entrante de l'interférence liée au récepteur connaît des changements rapides. Etant donné que le processus d'adaptation réel nécessite des signaux stables, seuls quelques instantanés de données peuvent être utilisés pour estimer les valeurs de poids adaptatives. Sous la condition, si une méthode de formulation adaptative régulière est utilisée, la nulle formée autour de la direction d'interférence est trop étroite. L'interférence pourrait ensuite se déplacer facilement de la zone nulle et éviter la suppression, entraînant la défaillance de la méthode d'adaptation régulière [22].

I.11. Conclusion

Dans ce premier chapitre, nous avons parcourus les notions fondamentales sur la navigation par satellites notamment les systèmes de navigation et présenté brièvement les principes de base du GNSS telles que le principe de trilatération. Ensuite, nous avons exposé l'architecture d'un récepteur GNSS. L'accent a été mis sur la description précise des signaux, les performances d'acquisition et poursuite des signaux de navigation GNSS. En fin, nous avons vu une description sur les différents types d'interférences, les caractéristiques et les impacts d'interférence au récepteurs GNSS.

Dans le prochaine chapitre nous allons parlerons sur les algorithmes évolutifs multi-objectif qui trouve un ensemble des solutions possibles efficaces (appelé front de Pareto) pour notre problème pour faire la détection et la suppression des signaux interférences.

Chapitre II

Algorithmes d'optimisation

II.1. Introduction

Dans l'industrie mécanique, chimique, des télécommunications etc. Ont des problèmes complexes ce genre de problèmes ont généralement des plusieurs objectifs, La plupart d'entre eux nécessitent l'optimisation simultanée de plusieurs objectifs souvent contradictoires. L'optimisation multi-objective vise donc à optimiser simultanément plusieurs objectifs. Contrairement à l'optimisation uni-objectif où la solution optimale est unique et souvent clairement définie, en raison de la capacité des algorithmes d'optimisation multi-objectifs évolutionnaires MOGA à trouver un ensemble de solutions compromis qui se appel front de Pareto [9, 11, 24, 25].

Dans ce chapitre, nous avons établi les fondements nécessaires de la compréhension des algorithmes génétiques (les opérateur , le principe de fonctionne) , ensuite nous présentons les principes et les concepts mathématiques relatifs à l'optimisation multi objectif, nous concentrons sur un algorithme génétique multi-objectif MOGA puis nous décrivons quelques algorithmes de résolution tout en accordant plus d'importance à l'algorithme NSGAI (Non-dominated Sorting Genetic Algorithm II) qui est utilisé dans notre travail.

II.2. Algorithmes génétiques (AGs)

Dans les années 90, les algorithmes évolutionnaires ont été développé isolément et à peu près simultanément, par différents scientifiques on quatre type : la programmation évolutionniste (PE) et les stratégies d'évolution (SE) et les programmations génétique (PG) et les algorithmes génétiques (AG) [26]. John Holland en 1975 [27], il introduit le premier modèle formel des algorithmes génétiques dans son livre «*Adaptation in Natural and Artificial Systems*». L'auteur Goldberg [28], définit les algorithmes génétiques comme des algorithmes d'exploration fondés sur les mécanismes de la sélection naturelle de Darwin, il combine une stratégie de struggle for life (survance du plus adapté) avec un échange d'information aléatoire mais structure [29].

Ces algorithmes font partie de la classe des algorithmes dits stochastiques, donc sont des algorithmes d'optimisation stochastique populaire et de recherche, fourni des solutions aux problèmes n'ayant pas de solutions calculables en temps raisonnable de façon analytique ou algorithmique, des milliers de solutions plus ou moins bonnes sont créés au hasard puis sont soumises à un procédé d'évaluation de la pertinence de la solution les plus adaptés [29, 30].

Les algorithmes génétiques sont des approches d'optimisation qui utilisent des techniques dérivées de la science génétique et de l'évolution naturelle : la sélection, la mutation et le croisement. Pour utiliser ces approches, on doit disposer des éléments suivants [31]:

Population : Départ de l'algorithme sont ensemble initial de solutions aléatoires d'un problème (nombre d'individus, de générations ou de probabilités ...). Chaque solution candidate possède un ensemble de propriétés (ses chromosomes ou génotype) qui peuvent être mutées et altérées.

Croisement, mutation et sélection: D'après les opérateurs génétiques, certaines solutions de la première population sont utilisées pour former et évaluation des nouveaux individus soit meilleure que la précédente.

Nouvelle génération: Le processus aléatoire étant orienté par la sélection. Les individus de la nouvelle génération tendent à être de meilleure qualité que ceux de la génération précédente.

Individus plus qualifiés: Le processus est répété jusqu'à ce qu'un critère de convergence soit satisfait (meilleure solution).

On peut résumer ce qui différencie les AGs des autres approches d'optimisation en quatre points [32]:

- Utilisent un codage des paramètres (et non les paramètres eux-mêmes);
- Travaillent sur une population de points (et non un point unique);
- Utilisent les valeurs de la fonction étudiée (et non la dérivée par exemple...);
- Utilisent des règles de transition probabilistes (et non déterministes).

Dans les sections suivantes, avant de passer à l'explication des processus d'optimisation des algorithmes multi-objectif, il faut tout d'abord expliquer: le choix de population initiale, les différents types de codage, les opérateurs génétiques, la fonction d'évaluation.

II.3. Population initiale et codage

La première étape de l'algorithme génétique est la formation de la population initiale c'est-à-dire le choix des individus (solutions) de départ que nous allons faire évoluer. Cette population est constituée de " N " individus, ceux-ci représentent un ensemble de solutions de départ. On pourrait prendre des individus régulièrement répartis dans l'espace. Néanmoins, une initialisation aléatoire est plus simple à réaliser. Les individus sont tirés au hasard selon une distribution uniforme. Le choix de cette population peut conditionner fortement la rapidité de notre algorithme. Tout d'abord, il faut codées les solutions initiales, ensuite l'AG démarre avec une population composée de N individus dans le codage retenu [33].

Le codage est une partie très importante des algorithmes génétiques, est une fonction qui permet de passer de la donnée réelle du problème traité à la donnée utilisée par l'algorithme génétique. Il permet de représenter l'individu sous la forme d'un chromosome. Chaque

paramètre de chromosome est assimilé à un gène qui prend des valeurs dans un alphabet binaire ou non, toutes les valeurs qu'il peut prendre sont les allèles de ce gène.

Le principe de codage est permet d'établir une connexion entre la valeur de la variable et les individus de la population. Et associe à chacun des points de l'espace d'état une structure de données.

Le choix du codage est délicat est l'élément le plus important dans la conception de l'algorithme puisqu'il doit permettre de coder toutes les solutions et permettre la mise en œuvre des opérateurs de reproduction. C'est ainsi que le bon déroulement des algorithmes génétiques sera assuré. Il existe plusieurs type de codages (binaire, réel, forme arbre etc.) [31, 34-37]. Les plus utilisé sont détaillés dans les deux sous sections suivantes :

II.3.1. Codage binaire

Ce type de codage a pour intérêt de permettre la création d'opérateurs de croisement et de mutation simple, dans le codage binaire le gène est codé par un caractère binaire 0 ou 1. Ce codage est certainement le plus utilisé. Permis certes de résoudre beaucoup de problèmes, mais il s'est avéré obsolète pour certains problèmes d'optimisation numérique, malgré toutes les critiques qui lui ont été adressée, le codage binaire reste le plus utilisé [26, 30, 31, 35]. Dans la Figure II.1, nous avons représenté un exemple de codage binaire d'une deux chromosome.

1	0	1	1	0	0	1	1
1	0	0	0	1	0	0	1

Figure II.1. Exemple de codage binaire d'une deux chromosome.

II.3.2. Codage réel

Comme l'exemple de la Figure II.2, contrairement au codage binaire, un gène est représenté par une suite de bits (un bit dans le code binaire) qui est associé à un réel. Cela peut-être utile notamment dans le cas où l'on recherche le maximum ou le minimum d'une fonction réelle. La procédure d'évaluation des chromosomes est plus rapide vu l'absence de l'étape de transcoding (du binaire vers le réel) [29, 31, 38, 39].

1	3	6	2	5	8	4	7
6	1	3	2	7	4	5	8

Figure II.2. Exemple de codage réel d'une deux chromosome.

II.4. Fonction d'évaluation

Après avoir généré aléatoirement la population initiale sur l'espace de recherche, l'évaluation des individus s'effectue à l'aide d'une fonction, la fitness ou valeur sélective, pour mesurer les performances de chaque individu qui correspond à une solution donnée du problème à résoudre, permet de quantifier la capacité d'un individu à survivre en lui affectant un poids couramment appelé fitness, qui permet d'associer une valeur à chaque individu ce qui offre la possibilité de le comparer avec d'autres individus est par de calculée la force de chaque chromosome de la population (étape de sélection) puis modifiés (croisement et mutation). Enfin ces valeurs calculées serviront au processus de sélection si sera sélectionné pour être reproduit ou pour déterminer s'il sera remplacé [40].

II.5. Opérateurs génétiques

Les opérateurs d'un algorithme génétique jouent un rôle prépondérant dans la possible réussite d'un AG qui est la diversité de la population doit être entretenue aux cours des générations afin d'explorer le plus largement possible l'espace de recherche [41].

II.5.1. Sélection

La sélection est pour l'objectif de "favoriser" les meilleurs individus qui participent à une reproduction (croisement ou mutation) généralement les individus les plus forts (meilleurs scores de fitness) pour produire les enfants les plus performants. Cet opérateur tend à augmenter l'importance de l'utilisation des solutions de bonne qualité pour la génération de descendants [29, 30, 42, 43]. Pratiquement, la probabilité p_i d'évolution d'un individu c_i est définie par :

$$p_i = \frac{f_e(c_i)}{\sum_{j=1}^N f_e(c_j)} \quad (\text{II.1})$$

où : f_e la valeur de la fonction d'aptitude associée au $i^{\text{ème}}$ individu, et N est la taille de la population.

Dans la littérature scientifique, il existe plusieurs techniques proposées par différents auteurs (la roue de loterie biaisée (roulette Wheel), la Sélection par rang, uniforme, par tournois ...), les plus utilisés sont :

- **La sélection uniforme:** C'est une technique très simple qui consiste à sélectionner un individu aléatoirement de la population P et ne fait appelle à aucune valeur d'adaptation.
- **La sélection par tournois:** Elle consiste à choisir aléatoirement deux ou plusieurs individus et à sélectionner le plus fort. Ce processus est répété plusieurs fois jusqu'à l'obtention de N individus qui serviront des parents.

II.5.2. Croisement

L'opération de croisement (crossover) est une opération de reproduction qui permet l'échange d'information entre les chromosomes (individus). Il utilise deux parents pour former un ou deux enfants. Les deux parents sont choisis par l'opération de sélection. Le croisement permet l'innovation (les enfants sont différents de leurs parents) et repose sur l'idée que deux parents performants produiront des enfants plus performants [29].

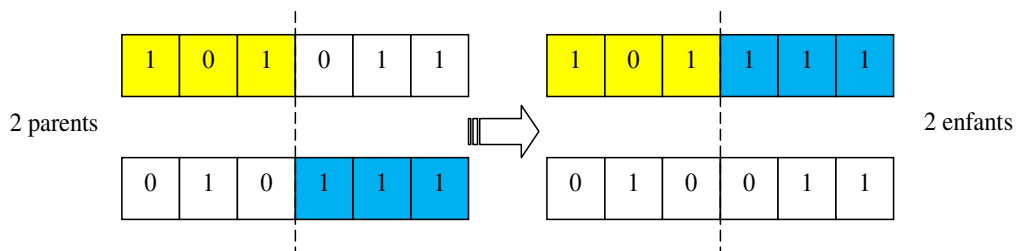


Figure II. 3. Croisement avec un point de crossover.

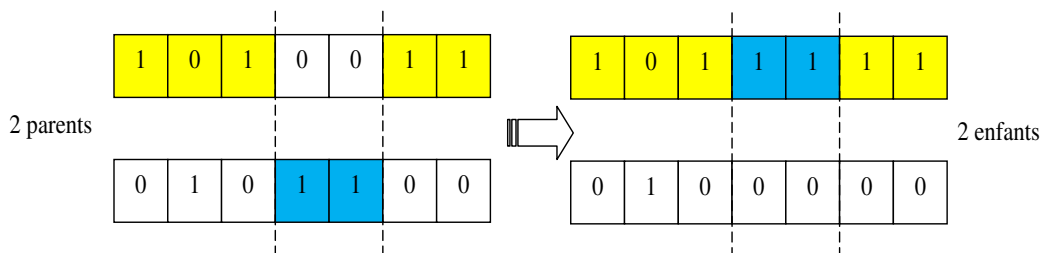


Figure II.4. Croisement avec 2 points de crossover.

Dans notre travail, nous avons utilisé la méthode de croisement intermédiaire (intermediate crossover), représenté dans l'équation suivante :

$$\begin{cases} \text{enfant}_1 = \text{parent}_1 + R_1 \times R_2 \times (\text{parent}_2 - \text{parent}_1) \\ \text{enfant}_2 = \text{parent}_2 - R_1 \times R_2 \times (\text{parent}_2 - \text{parent}_1) \end{cases} \quad (\text{II.2})$$

où R_1 est une valeur aléatoire entre 0 et 1. R_2 peut être spécifié par l'utilisateur, dans notre cas est fixé 1.

II.5.3. Mutation

La mutation permet aux chromosomes des fils biologiques d'être différents de ceux des pères biologiques. Le rôle de cet opérateur est de modifier aléatoirement la valeur d'un gène dans un chromosome (serait de modifier un ou plusieurs gènes) [31, 44].

Dans le cas du codage binaire, chaque bit $a_i \in [0,1]$ est remplacé par son complémentaire et dans le cas de codage réel [31], les opérateurs de mutation les plus connus et les plus utilisés sont :

- **Mutation par inversion binaire:** est souvent utilisée avec la représentation binaire. Elle consiste simplement à inverser de manière aléatoire un gène d'un individu (voir la Figure II.5) [9].

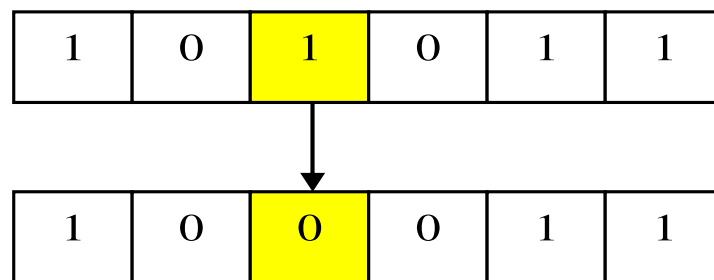


Figure II.5. Illustration de la mutation par inversion.

- **Mutation par échange :** consiste à sélectionner de manière aléatoire deux gènes d'un individu et d'échanger les positions respectives des deux éléments choisis [9]. La figure suivante représente un exemple de ce type de mutation.

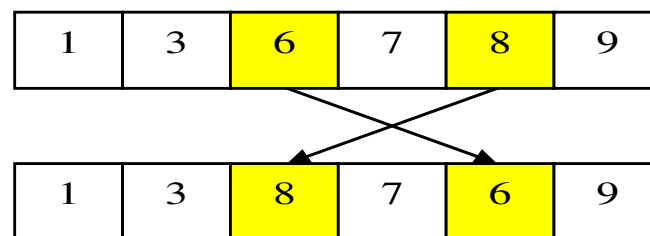


Figure II.6. Illustration de la mutation par échange.

- **Mutation par insertion réel:** Comme représente la Figure II.7, l'opérateur d'insertion choisit un gène aléatoirement et l'insère dans une autre position du chromosome [45].

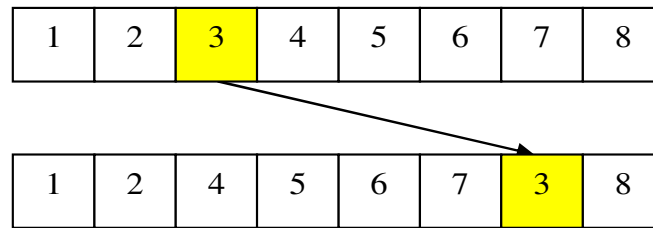


Figure II.7. Illustration de la mutation insertion.

- **Mutation uniforme :** Pour la mutation uniforme, qui est une simple extension de la mutation binaire, on remplace le gène x_k , $k = \{1, 2, \dots, N\}$ sélectionné par une valeur quelconque x'_k qui est une valeur aléatoire qui appartienne à l'intervalle $[x_k^{\min}, x_k^{\max}]$. Le résultat est alors le chromosome: $S_x^{t+1} = (x_1, \dots, x'_k, \dots, x_N)$
- **Mutation multi non-uniforme :** Le calcul de la nouvelle valeur d'un gène est un peu plus complexe. Le gène x_k subit des modifications importantes durant les premières générations, puis graduellement décroissantes au fur et à mesure que l'on progresse dans le processus d'optimisation. Pour une génération t , on tire au sort une valeur binaire b qui décidera si le nouvel élément x'_k doit être supérieur ou inférieur à l'élément initial x_k . La nouvelle valeur est alors définie par :

$$x'_k = \begin{cases} x_k + \Delta(t, x_k^{\max} - x_k) & \text{si } b \text{ est égal à } 1 \\ x_k - \Delta(t, x_k - x_k^{\min}) & \text{si } b \text{ est nul} \end{cases} \quad (\text{II.3})$$

où : $\Delta(t, x)$ est une fonction qui définit l'écart entre la nouvelle valeur et la valeur initiale à la génération t et b est un nombre aléatoire qui prend les valeurs 0 ou 1. Dans la littérature scientifique, les chercheurs proposent d'utiliser une fonction $\Delta(t, x)$ correspondante à une décroissance exponentielle de l'écart à travers les générations. Cette fonction est définie par : $\Delta(t, y) = y \cdot (1 - r^{(1-t/T)^\beta})$

où T est l'indice de génération pour laquelle l'amplitude de la mutation s'annule, β est un paramètre de l'opérateur de mutation, r est un nombre produit aléatoirement dans l'intervalle $[0, 1]$ et t le numéro de la génération.

Dans notre travail, nous avons utilisé pour tous les chromosomes l'équation (II.3) [46].

II.6. Fonctionnement général d'une AG

Pour pouvoir exploiter efficacement le potentiel d'une AG, indépendamment du problème traité, il est demandé de prendre en compte les principes fondamentaux suivants:

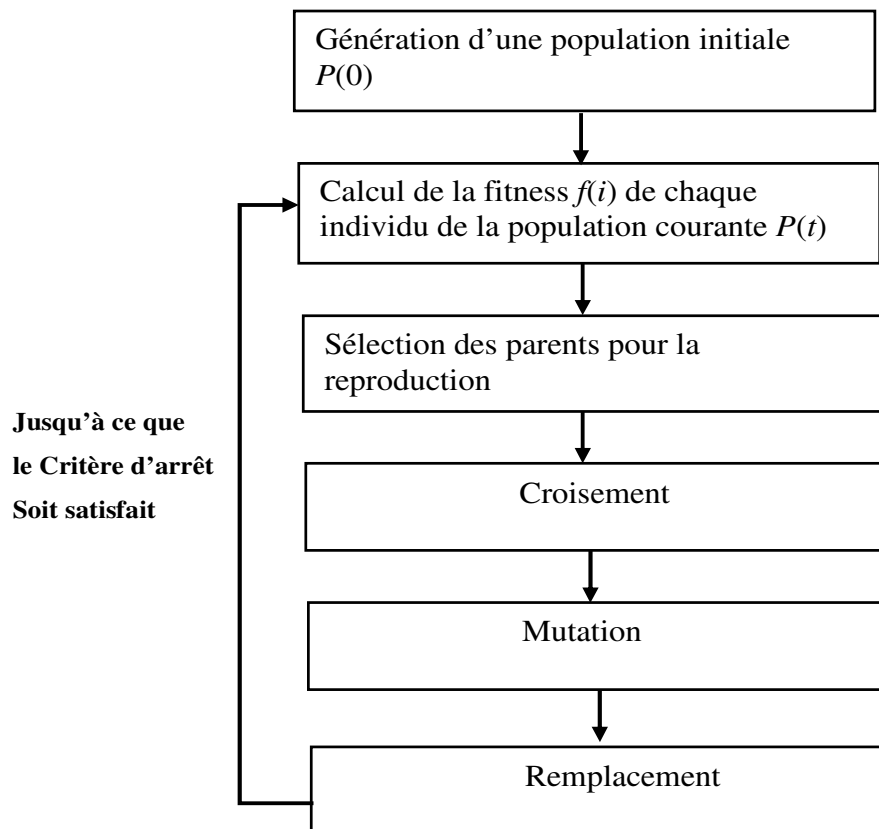


Figure II.8. Schéma général d'AG.

- Un principe de codage des individus de la population (la qualité du codage conditionne le succès d'AG), après une phase de modélisation mathématique du problème traité.
- Un mécanisme de génération de la population initiale $P(0)$. Le choix de la population initiale est important car il influe sur la rapidité de la convergence vers un optimum global.
- Définir une fonction d'évaluation f pour chaque individu i et évaluer une solution et la comparer aux autres.
- Choisir les solutions par un mécanisme de sélection pour un éventuel couplage.
- Générer de nouvelles solutions $P(t)$ à l'aide des opérateurs de croisement et de mutation.
- Etablir un compromis entre les solutions produites (progénitures) et les solutions productrices (les parents) en utilisant un mécanisme d'insertion.

II.7. Optimisation multi-objectif (Optimisation multicritères)

Lorsqu'un problème d'optimisation implique plus d'une fonction objectif, la tâche consistant à rechercher une ou plusieurs solution(s) optimale(s) est appelée optimisation multi-objectif, la plupart des problèmes réels sont en fait des problèmes multi-objectif, c'est-à-dire que l'on cherche à optimiser simultanément plusieurs critères qui sont généralement conflictuels

l'optimisation multi-objectif (dite aussi multicritères), donc se réfère à la recherche de valeurs de variables (de décision) qui correspondent à l'optimum (minimum ou maximum) de plus d'une équation objectif [47, 48].

On peut formaliser un problème d'optimisation multi-objectif comme suit:

$$F(X) = (f_1(X), f_2(X), \dots, f_m(X)) \quad (\text{II.4})$$

avec les contraintes

$$\begin{aligned} H(x) &= 0 \\ G(X) &\leq 0 \end{aligned} \quad (\text{II.5})$$

où: F est la fonction (scalaire) à minimiser, appelée fonction coût ou fonction objectif, X est un ensemble de variables indépendantes, m est le nombre de fonctions objectif, G et H sont respectivement des contraintes d'inégalités et d'égalités.

L'optimisation multi-objectif vise donc à optimiser simultanément plusieurs objectifs. Contrairement à l'optimisation uni-objectif où la solution optimale est unique et souvent clairement définie, de nombreux algorithmes génétiques ont été développés pour résoudre ce type de problème. Parmi les plus significatifs d'entre eux figurent: l'algorithme Multi Objective Genetic Algorithm (MOGA), l'algorithme Niched Pareto Genetic Algorithm (NPGA), utilisant une sélection par tournoi, basée principalement sur la dominance de Pareto, l'algorithme Non Dominated Sorting Genetic Algorithm (NSGAI) et enfin l'algorithme NSGAII qui est sans doute l'algorithme le plus populaire [9, 11, 25].

Dans le contexte de l'optimisation multi-objectif qui se base sur le Pareto, on vise en général:

- à trouver l'ensemble des solutions Pareto optimales, c'est-à-dire celles qui couvrent tout le front de Pareto;
- à s'assurer que les solutions soient suffisamment différentes les unes des autres et qu'elles ne soient pas biaisées en favorisant un objectif particulier [45].

II.7.1. Problème multi-objectif

Un problème multi-objectif ou multicritères peut être défini comme un problème dont on recherche l'action qui satisfait un ensemble de contraintes et optimise un vecteur de fonctions objectives. Les problèmes d'optimisation ont en général plusieurs solutions car la définition d'un optimum ne peut pas être établie dans les problèmes multi-objectif.

Comme le montre l'équation suivante (II.5), un problème d'optimisation multi-objectif contraint implique de minimiser (ou de maximiser) k différentes fonctions objectives f basées sur m différentes contraintes d'inégalité g [49].

$$\left. \begin{array}{l} f_i(x_1, x_2, \dots, x_n) (i = 1, 2, \dots, k) \\ g_j(x_1, x_2, \dots, x_n) \leq 0 (j = 1, 2, \dots, m) \end{array} \right\} \quad (\text{II.6})$$

Les fonctions objectif du problème d'optimisation forment un espace multidimensionnel appelé espace des fonctions objectif, en plus du traditionnel espace des variables de décision. Le schéma de la figure suivante illustre les deux espaces des solutions $X = (x_1, \dots, x_n)$, pour chaque solution dans l'espace des variables de décision, il existe un point dans l'espace des fonctions objectif tel que [50].

$$F(X) = (f_1(X), f_2(X), \dots, f_m(X)) \quad (\text{II.7})$$

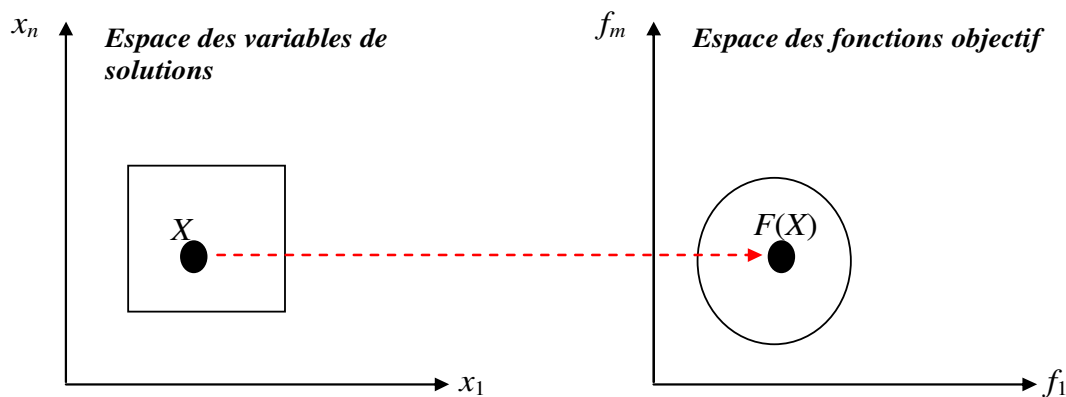


Figure II.9. Représentation d'un problème multi-objectif.

II.8. Techniques non élitistes

Ces techniques n'utilisent pas le principe d'archivage pour conserver les solutions Pareto optimales durant le processus d'optimisation. Elles mettent en œuvre le concept d'optimalité de Pareto pour classer les solutions en différents groupes de solutions non dominées, en vue d'aider à la convergence des algorithmes. Pour assurer la diversité de la population dans la région de Pareto, ces algorithmes utilisent la technique de la fonction de partage. Ces algorithmes comprennent le MOGA, le NPGA et le NSGAI. Les inconvénients majeurs de ces techniques sont leur convergence lente et la difficulté de maintenir une bonne diversité dans la population [11].

II.8.1. Algorithmes Génétiques Multi objectifs (MOGA)

MOGA proposé par Fonseca et Fleming En 1993 [51], est une méthode dans laquelle chaque individu de la population est rangé en fonction du nombre d'individus qui le domine, utilise la relation de dominance pour déterminer l'efficacité d'un individu. Ensuite, l'algorithme utilise

une fonction de calcul de la performance permettant de prendre en compte le rang de l'individu et le nombre d'individus ayant le même rang.

Soit un individu x à la génération t , le rang de cet individu est donné par la relation suivante:

$$\text{rang}(x,t) = 1 + p(t) \quad (\text{II.8})$$

Tous les individus non dominés on affecte le rang 1, les individus dominés se voient donc affectés à un rang important. Pour le calcul de l'efficacité, on peut suivre les étapes suivantes:

1. Classer les individus en fonction de leur rang.
2. Affecter une efficacité à un individu en interpolant à partir du meilleur (rang 1) jusqu'au plus mauvais (rang n).

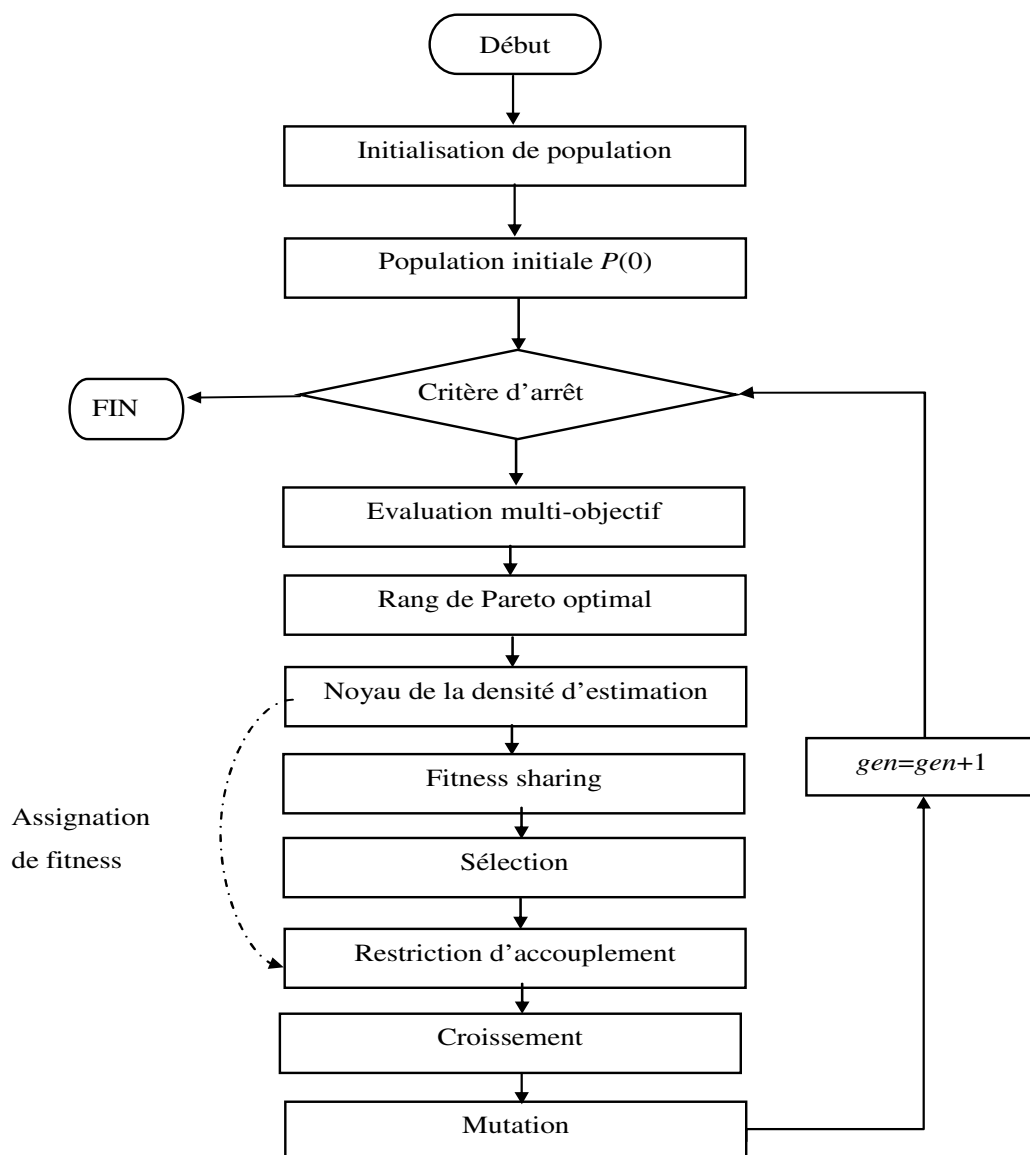


Figure II.10. Algorithme Génétique Multi objectifs.

MOGA est un algorithme évolutionnaire emploie les opérateurs suivantes:

- Le rang de Pareto optimal.
- Fitness Sharing.
- Restriction d'accouplement.
- Sélection, croisement et mutation.

Dans MOGA, la population initiale est choisie aléatoirement dans une marge définie et alors décodée (en cas de chromosome non réel). Et puis évalué chacun individu de la population initiale par un ensemble de valeurs de fonction objectif. Alors la séquence des opérateurs génétiques est appliquée, ayant un résultat pour la génération suivante avec la fitness est un processus plus raffiné et importante l'explication comme la représente dans la Figure II.10.

II.8.2. Frontière de Pareto

Dans un problème multicritères dans lequel on cherche à optimiser plusieurs objectifs contradictoires, on appellera *Front de Pareto* du problème, l'ensemble des points de l'espace de recherche tels qu'il n'existe aucun point qui est strictement meilleur qu'eux sur tous les critères simultanément. Il s'agit de l'ensemble des meilleurs compromis réalisables entre les objectifs contradictoires [49]. La figure suivante présente un problème à deux objectifs f_1 et f_2 , comme le montre la figure on a quatre possibilités des frontières de Pareto en fonction de notre problème à minimiser ou maximiser. La question qui doit être posée est : *Comment trouver l'optimum de Pareto ?*

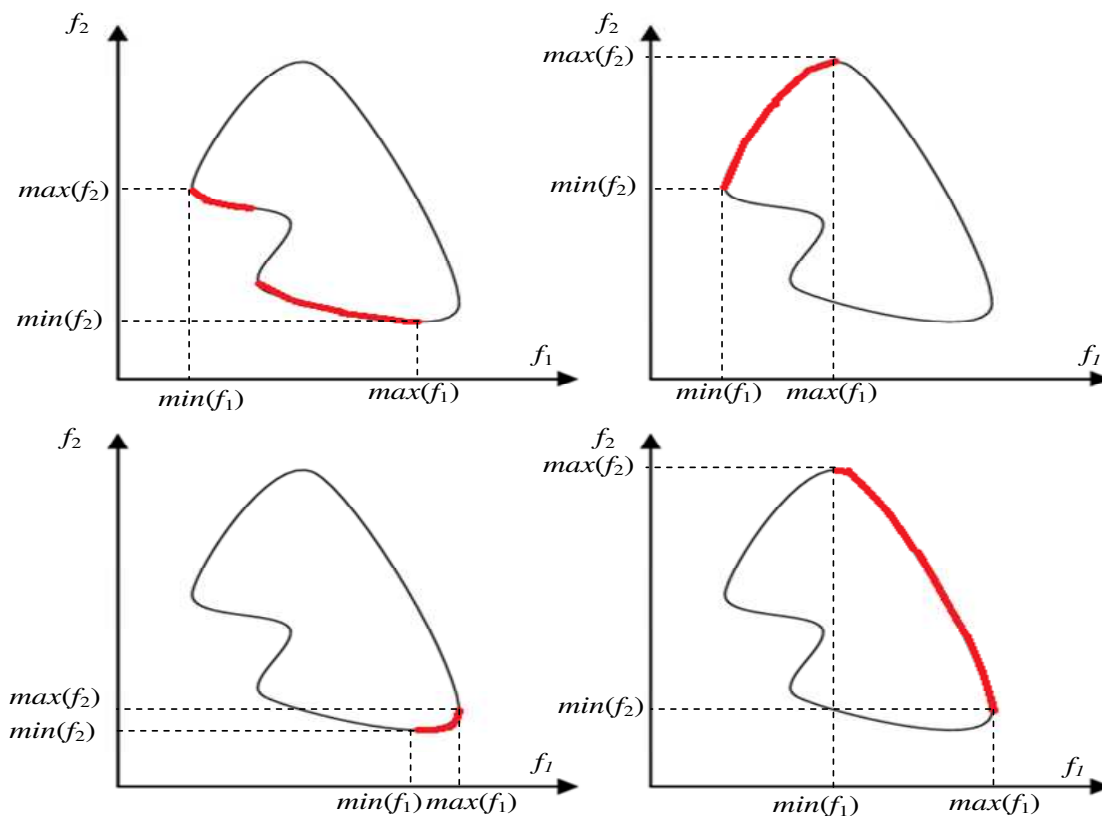


Figure II.11. Exemples de frontière de Pareto.

II.8.2.1. Optimum de Pareto

L'auteur Vilfredo Pareto [52, 53], un mathématicien italien, forme le concept suivant : dans un problème multi-objectif, il existe un équilibre tel que l'on ne peut pas améliorer un critère sans détériorer au moins un des autres critères.

Cet équilibre a été appelé optimum de *Pareto*. Un point x est dit *Pareto-optimal* s'il n'est dominé par aucun autre point appartenant à X . Ces points sont également appelés solutions non inférieures ou non dominées.

II.8.2.2. Domination et front de Pareto

Soit $u=(u_1, u_2, \dots, u_m)$ et $v=(v_1, v_2, \dots, v_m)$ deux vecteurs de décision. v domine u ($u < v$) pour un problème de minimisation si et seulement si :

$$\begin{cases} \forall_i \in 1, \dots, n, f_i(v) \leq f_i(u) \\ \exists_i \in 1, \dots, n, f_i(v) < f_i(u) \end{cases} \quad (\text{II.9})$$

Le front de Pareto fp^* est l'ensemble des vecteurs de décision qui ne sont pas dominés [54] :

$$fp^* = \left\{ x \in X \mid \nexists x' \in X, x < x' \right\} \quad (\text{II.10})$$

Le front de Pareto est l'ensemble des solutions de compromis. Le but de l'optimisation multi-objectif va être d'identifier cet ensemble de compromis optimaux entre les critères (le front de Pareto pour un problème donné).

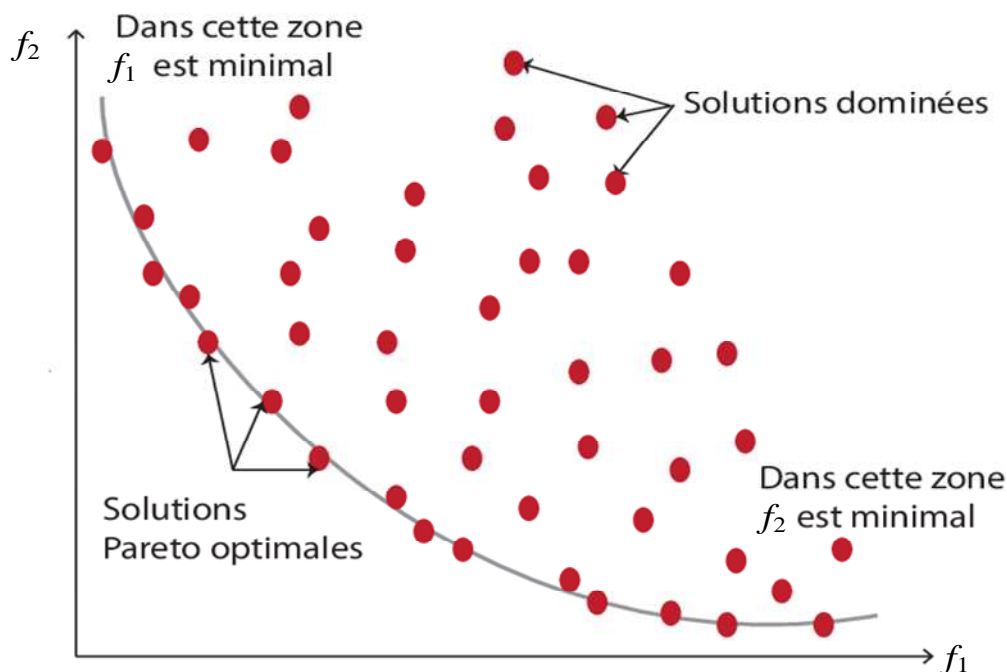


Figure II.12. Front de Pareto et rangs de Pareto pour un problème de minimisation de deux objectifs.

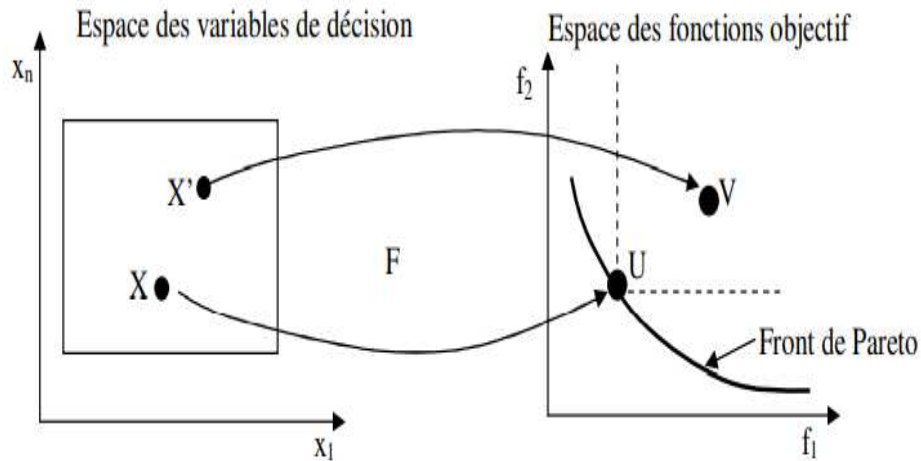


Figure II.13. Dominance de Pareto et optimalité de Pareto.

Comme la représente les figures II.12 et II.13, l'ensemble des solutions optimales est appelé "ensemble de Pareto optimal", et l'ensemble des valeurs des fonctions objectif correspondantes dans l'espace des fonctions objectif est appelé "front de Pareto". L'utilisation d'une sélection basée sur la notion de dominance de Pareto va faire converger la population vers un ensemble de solutions très efficaces.

II.9. NSGA (Non dominated Sorting Genetic Algorithm)

Cette méthode a été proposée par les auteurs Srinivas et Deb en 1994 [55]. Comme le MOGA, l'approche NSGA est basée sur le classement non dominé. Elle affecte à chaque individu une fitness selon le front auquel il appartient. Chaque front correspond à un groupe d'individus (solutions) ayant le même degré de dominance au sens Pareto. Les individus du Front 1 auront une meilleure fitness que ceux du Front 2 qui eux, auront une meilleure fitness que ceux du Front 3 et ainsi de suite (voir la Figure II.14). Pour la sélection, elle applique ensuite le partage de fitness au niveau de chaque front pour maintenir la diversité.

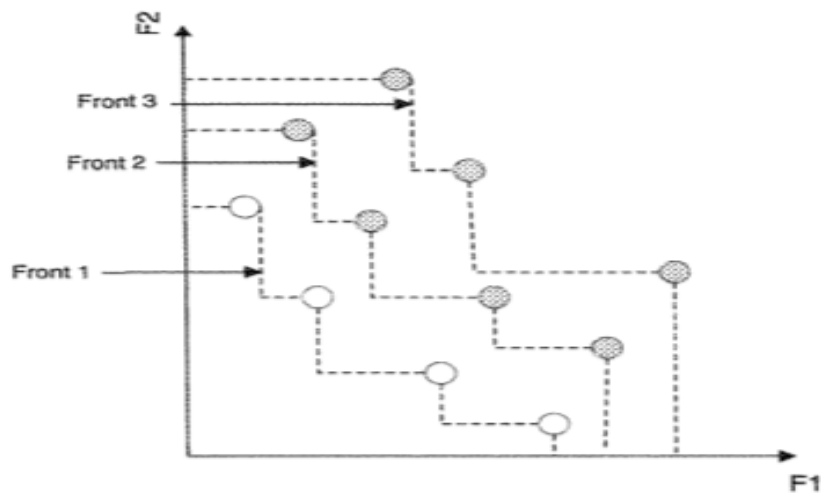


Figure II.14. Illustration du classement par front dans le NSGA.

Cette méthode paraît moins efficace en temps de calcul que la méthode MOGA car le temps de calcul nécessaire à la construction des fronts et au partage de la performance est important. Pour cela, dans notre travail on utilise les techniques élitistes basées sur la méthode NSGAI.

II.10. Techniques élitistes

C'est une génération d'algorithmes évolutifs multi-objectif qui intègrent le concept d'archivage dans leur fonctionnement. Il s'agit entre autres de l'algorithme SPEA1 et SPEA2 (Strength Pareto Evolutionary Algorithm 1,2) NSGA.

II.10.1. NSGAI (Non-dominated Sorting Genetic Algorithm II)

Un algorithme d'optimisation multi-objectif évolutif typique, qui a été proposé par Deb et al. en 2000 [11], comme une version améliorée du NSGA dotée des nouveaux mécanismes. Il recherche des solutions en utilisant une combinaison de tri rapide non dominé, de tri groupé et de sélection de tournois bondés. L'auteur tente de résoudre toutes les critiques faites à la première version de l'algorithme: complexité, non-élitisme et utilisation du partage de performance. Aussi, établissant les rapports de dominance entre les individus et offrant une méthode de tri particulièrement rapide des chromosomes. Cet algorithme utilise la mesure du surpeuplement autour des individus pour assurer la diversité dans la population. Cette nouvelle version de NSGA a permis de réduire la complexité de l'algorithme et de créer une méthode plus élitiste et de supprimer les paramètres de la fonction de partage de performance [9].

Dans le NSGAI, l'auteur Deb remplace la fonction de partage de performance par une fonction de remplacement. Pour cela, il attribue deux caractéristiques à chaque individu:

- ❖ i_{rank} qui représente le rang de non-dominance de l'individu. Cette caractéristique dépend de la frontière à laquelle appartient l'individu. Les individus du Front 1 auront un i_{rank} de 0 car ils sont non dominés les individus du front 2 un i_{rank} de 1 car ils ne sont dominés que par des individus du premier front et ainsi de suite.
- ❖ $i_{distance}$ qui représente la distance de remplacement de l'individu et permet d'estimer la densité de la population autour de lui.

Pour estimer la densité au voisinage d'une solution i , calcule la distance moyenne sur chaque objectif, entre les deux points les plus proches situés de part et d'autre de la solution, comme la représente la Figure II.15 [11].

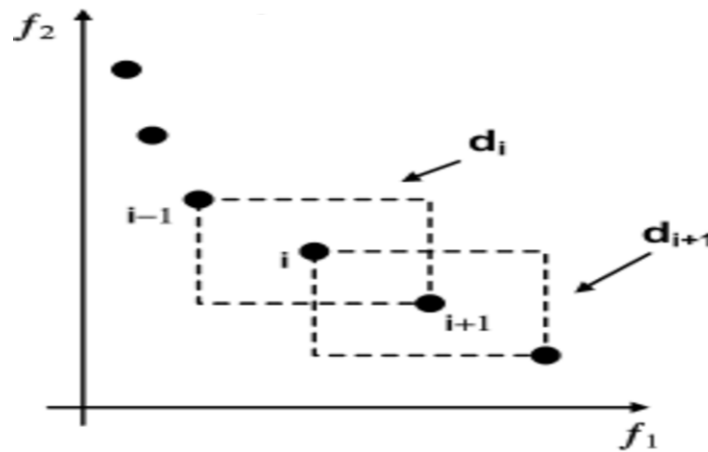


Figure II.15. Illustration du calcul de l' i distance du NSGA II.

II.10.1.1. Principe de l'algorithme NSGAII

La figure dessous montre une illustration conceptuelle de la façon dont la population est mise à jour dans une sélection de tournoi bondée. D'abord, générez un groupe $R_t = P_t \cup Q_t$ qui combine la population parent P_t et la population enfant Q_t (d'après les opérateurs de croisement et de mutation). Ce groupe R_t est d'abord soumis à un tri non dominé pour classer chaque solution candidate (la taille de la population totale R_t est $2N$). Cette population R_t ensuite classée en différents fronts de non-dominance (F_1, F_2, \dots, F_n), en utilisant le concept de dominance de Pareto [24].

L'accent est alors porté sur les meilleurs individus, les individus faisant partie du premier front sont évidemment meilleurs que les individus du second front ainsi de suite. Si le nombre d'individus de ce front est inférieur à la taille fixée N , alors tous les individus de ce front feront partie de la prochaine population P_{t+1} . Pour compléter la taille de la population P_{t+1} à N , les nouveaux membres doivent être choisis dans les fronts restants, suivant leur rang. Ainsi les membres du second front F_2 sont choisis, puis les membres de F_3 et ainsi de suite jusqu'à obtenir une nouvelle population de taille N . Les individus qui ne peuvent entrer dans la nouvelle population sont rejetés. Ensuite les individus parents sélectionnés à partir de P_{t+1} sont ensuite croisés et manipulés génétiquement par mutation pour générer une nouvelle population enfant Q_{t+1} . Ces opérations constituent une étape de génération, et ces étapes de génération sont répétées le nombre de fois spécifié [11, 24].

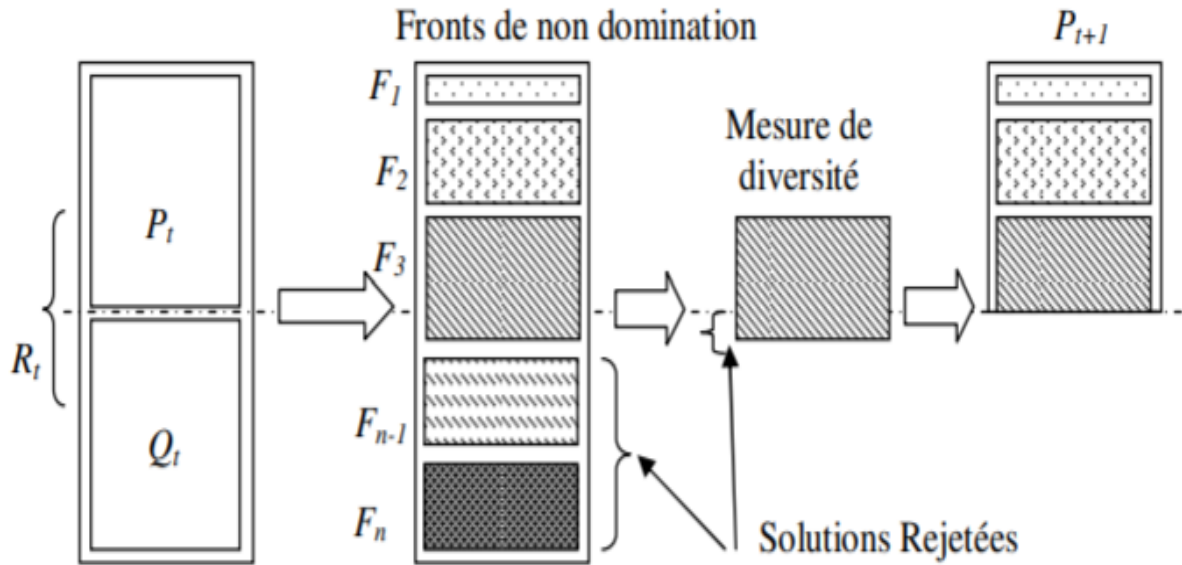


Figure II.16. Algorithme NSGAII.

II.10.1.2. Organigramme de l'algorithme NSGAII

La Figure II.17 montre l'organigramme général d'un algorithme NSGAII qui comporte trois étapes importantes : le tri non dominé, le calcul de la distance de crowding et la stratégie élitiste.

Après initialisation aléatoire de la population initiale P_0 , chaque itération de NSGAII se déroule comme suit [56]:

- Créer Q_t à partir de P_t en utilisant le tournoi et en appliquant des opérateurs de variation génétique aux individus gagnants
- Réunir les populations des parents et des enfants $R_t = P_t \cup Q_t$. Trier l'ensemble résultant R_t en sous-ensemble F_i .
- Soit une nouvelle population p_{t+1} (formée par ajoutant les fronts au complet). Soit le compteur des sous-ensembles non-dominés $i=1$.
- Tant que $|p_{t+1}| + |F_i| < N$, $P_{t+1} \leftarrow P_{t+1} \cup F_i$ et $i \leftarrow (i+1)$.

Si le nombre d'individus présents dans (P_{t+1}) est inférieur à (N) , une procédure de crowding est appliquée sur le premier front suivant, (F_i) , non inclus dans (P_{t+1}) pour insérer les $(N - |P_{t+1}|)$ meilleurs individus de (F_i) qui manquent dans la population (P_{t+1}) . Les individus de ce front sont utilisés pour calculer la distance de crowding entre deux solutions voisines. Une fois que les individus de la population (P_{t+1}) sont identifiés, une nouvelle population enfant (Q_{t+1}) est

créée par sélection, croisement et mutation. Répétées ces étapes, jusqu'à obtenu le résultat qui contient des individus du front de Pareto (individus non dominés).

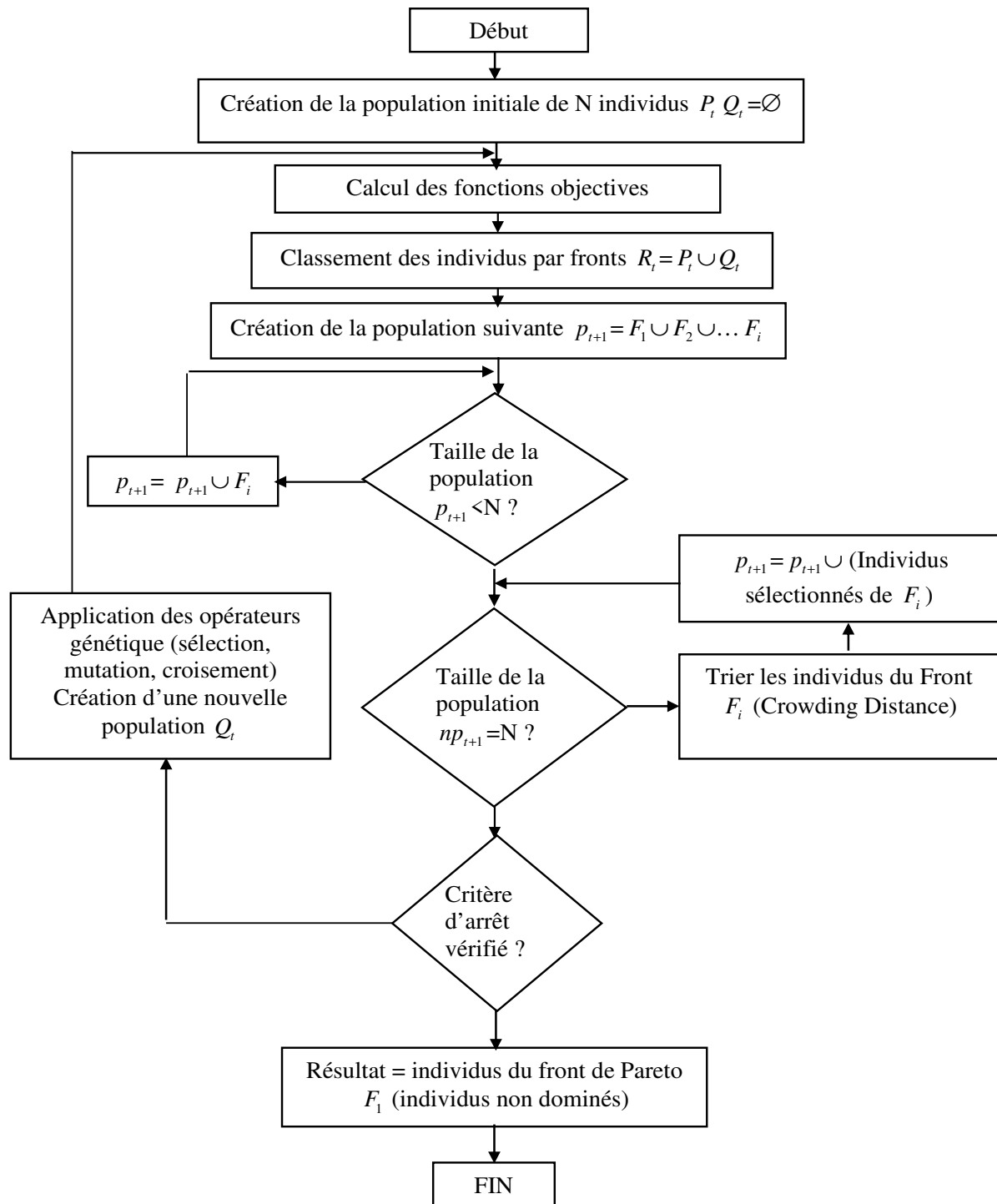


Figure II.17. Organigramme de l'algorithme NSGAI.

II.11. Conclusion

Dans ce chapitre nous avons présenté le problème multi-objectif et l'état de l'art des différentes techniques d'optimisation multi-objectif avec les algorithmes génétiques. Nous

avons expliqué c'est pourquoi ils conviennent est parfaitement pour traiter les problèmes d'optimisation multi-objectif et aborder la majorité de ces problèmes qui sont caractérisés, non pas par une solution unique, mais par un ensemble de solutions optimales (Pareto optimales, solutions efficaces). Ainsi nous avons présenté le principe de base et le fonctionnement de l'algorithme NSGAI, ce dernier est utilisé dans notre travail pour l'objectif de la détection et la suppression des signaux interférences, on utilise NSGAI comme une méthode d'optimisation à cause des caractéristiques qu'elles présentent en particulier la rapidité de tri et l'élitisme. La mise en œuvre de l'algorithme NSGAI et ainsi que leurs applications font l'objet des prochains chapitres.

Chapitre III

Intelligence artificielle

III.1. Introduction

Avec l'évolution de la technologie, l'intelligence artificielle a apparue et ne cesse de progresser dans plusieurs domaines déjà trouvés de multiples usages au sein de la société (les moteurs de recherche, la traduction automatique, les systèmes de navigation GPS, traitement du signal, les signaux GNSS, ...etc.), cela contribue à changer le monde entièrement. C'est une discipline scientifique recherchant des méthodes de solution de problèmes à forte complexité logique ou algorithmique, ensemble des théories et des techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence [57, 58].

On expose dans ce chapitre une petite revue des algorithmes d'apprentissage et leur types, ensuite on présente quelques méthodes de classification traditionnelle (perceptron multicouche MLP, Réseaux de neurones à fonctions radiales de base RBF et machine à vecteurs de support SVM), et après nous passons en revue de la détection des outliers par machine à vecteur de support à une classe OCSVM.

III.2. Algorithmes d'apprentissage

L'intelligence artificielle (IA) englobe l'apprentissage automatique (Machine Learning), qui consiste à programmer des algorithmes permettant d'apprendre automatiquement à partir des données et d'expériences passées ou par interaction avec l'environnement. Ce qui rend l'apprentissage machine vraiment utile est le fait que l'algorithme peut "apprendre" et adapter ses résultats en fonction de nouvelles données sans aucune programmation a priori. L'algorithme d'apprentissage est la méthode mathématique qui va modifier les poids de connexions afin de converger vers une solution qui permettra au modèle d'accomplir la tâche désirée de façon efficace [59, 60].

L'apprentissage automatique est une méthode d'identification paramétrique qui permet d'optimiser les valeurs des poids du modèle, représente beaucoup de méthodes telles que la méthode des k-moyennes, les SVM (séparatrices à vaste marge), les arbres de décision, les k plus proches voisins (KNN), et les réseaux de neurones [58, 61]. Il y a deux algorithmes d'apprentissage principaux : l'apprentissage supervisé, et l'apprentissage non supervisé. Ces deux méthodes sont détaillées dans les sections suivantes.

III.2.1. Apprentissage supervisé

L'apprentissage supervisé (supervised learning) s'intéresse aux données étiquetées, matérialisé par les méthodes de classification et de prédiction ou estimation, les classes des exemples sont connues et l'objectif est de trouver un modèle qui permet de prédire les classes des exemples X (voir la Figure III.1). L'apprentissage est supervisé si les différentes familles

des exemples sont connues a priori et si la tâche d'apprentissage est guidée par un superviseur ou un expert. Cette approche a pour l'objectif de la conception d'un modèle reliant des données d'apprentissage à un ensemble de valeurs de sortie y [58, 61]. Dans ce type d'apprentissage on cherche à :

- Estimer une fonction $f(x)$ qui est la relation entre les échantillons et leurs classes.
- Les échantillons utilisés comme données d'apprentissages sont accompagnés par la classe à laquelle ils appartiennent y .

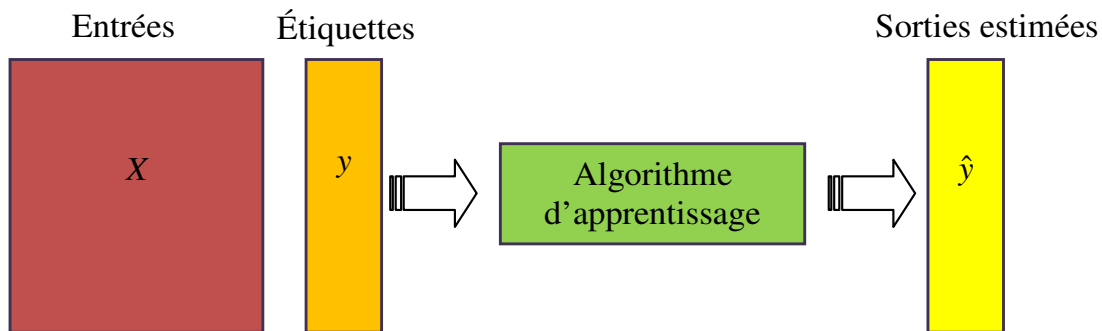


Figure III.1. Schéma d'un modèle supervisé.

III.2.2. Apprentissage non supervisé

L'apprentissage non-supervisé (unsupervised learning) traite des données non-étiquetées. La différence ici est que les comportements (ou catégories ou les classes) des données d'apprentissage ne sont pas connus, c'est ce que l'on cherche à trouver \hat{y} , c-à-d ne fournit pas au modèle les sorties que l'on désire obtenir, l'algorithme doit découvrir par lui-même la structure en fonction des données d'entrées. Donc on laisse évoluer librement jusqu'à ce qu'il se stabilise, veut dire que l'apprentissage non supervisé consiste à doter le système d'un mécanisme automatique qui s'appuie sur des règles précises de regroupement pour trouver les classes de référence avec une assistance minimale (sans l'intervention de l'expert). La figure II. 2, illustre le mécanisme de l'apprentissage non supervisé, dans ce cas les échantillons sont introduits en un grand nombre par l'utilisateur sans indiquer leur classe [59].

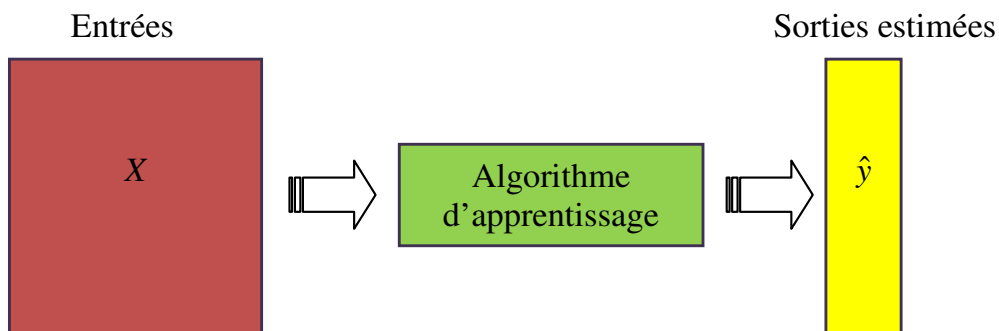


Figure III.2. Schéma d'un modèle non supervisé.

III.3. Les méthodes de classification traditionnelle

Sont nombreuses, telle que les séparateurs à Vastes Marges, les réseaux de neurones, arbres de décision, ...etc. Les méthodes de classification sont des méthodes d'apprentissage supervisé ont pour but d'identifier les classes auxquelles appartiennent des échantillons à partir de certains paramètres descriptifs. Pour choisir une méthode il faut relève de la bonne compréhension des fondements des méthodes et de ce qui permet de les distinguer afin de déterminer les modèles qui traiteraient au mieux un cas d'étude donné.

III.3.1. Réseaux de neurones multicouches MLP

Le perceptron multicouche (multi-layer perceptron MLP) sont des réseaux de neurones les plus connus, développé par les auteurs M. Minsky et S. Papert [62] en 1969, est un classifieur linéaire de type réseau de neurones artificiels (feed-forward), c'est à dire à propagation directe constitué de deux couches de poids adaptatifs avec une connectivité complète entre les entrées et les unités cachées, et entre les unités cachées et des sorties (voir la Figure III.3). L'information circule dans un seul sens, de la couche d'entrée vers la couche de sortie uniquement, entre ces deux couches il peut exister une ou plusieurs couches appelées couches cachées (sont pas visibles à l'extérieur du réseau) et constitue le cœur du réseau de neurones, ses fonctions d'activation sont du type sigmoïde. Chaque couche est constituée d'un nombre variable de neurones, les neurones de la couche de sortie correspondant toujours aux sorties du système [62, 63].

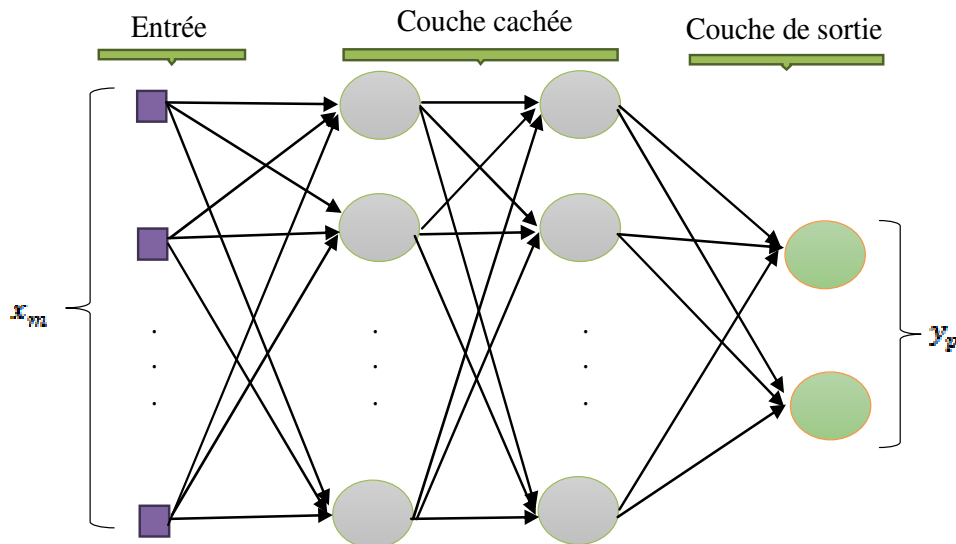


Figure III.3. L'architecture d'un réseau de neurones de type MLP.

L'idée principale est de grouper des neurones dans une couche. La première couche est reliée aux entrées, puis ensuite chaque couche est reliée à la couche précédente. Les neurones de la

première couche sont reliés au monde extérieur et reçoivent tous le même vecteur d'entrée. Ils calculent alors leurs sorties qui sont transmises aux neurones de la deuxième couche, etc.

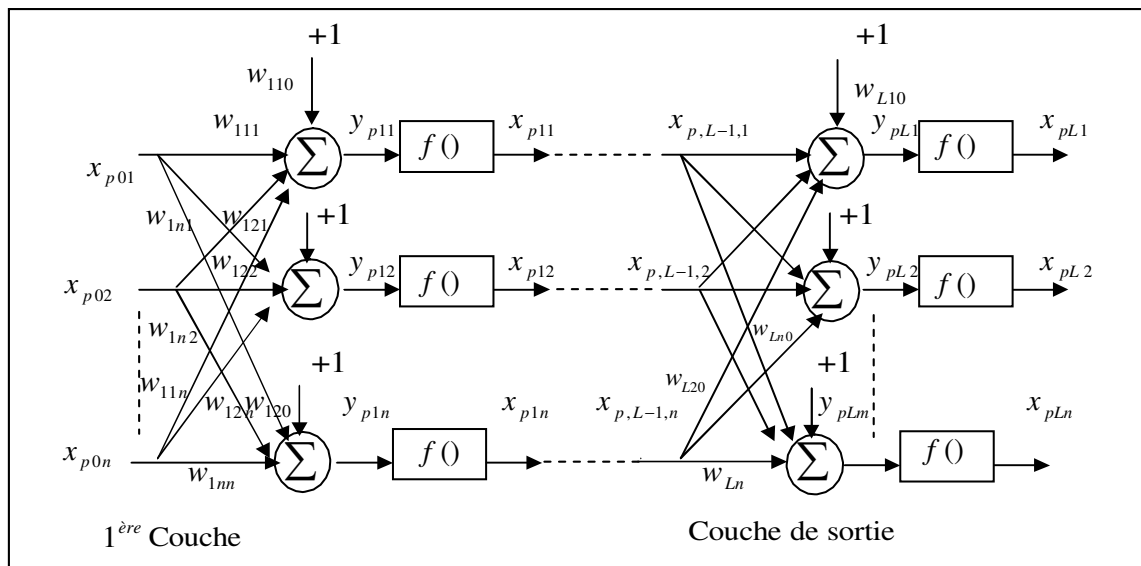


Figure III.4. Structure d'un réseau de neurone multicouche.

L'MLP utilise généralement une technique de descente de gradient connu sous le nom de rétropropagation des erreurs comme un algorithme d'apprentissage. L'algorithme de rétropropagation du gradient de l'erreur, est une technique de calcul des dérivées d'une entité par rapport à l'ensemble de ses entrées, utilisé pour résoudre le problème de calcul des erreurs de sorties des neurones des couches cachées. Cet algorithme change les poids d'un réseau dont l'architecture est fixée par l'utilisateur, repose sur la minimisation de l'erreur quadratique entre les sorties calculées et celles souhaitées.

L'erreur commise sur le $k^{\text{ème}}$ nœud de sortie est :

$$\delta_{pk} = O_{pk} - x_{pk} \quad (\text{III.1})$$

Par conséquent l'erreur totale (pour tous les nœuds) est :

$$E_p = \frac{1}{2} \sum_{k=1}^m \delta_{p,k}^2 = \frac{1}{2} \sum_{k=1}^m (O_{p,k} - x_{p,l,k})^2 \quad (\text{III.2})$$

Pour minimiser E_p (erreur quadratique), on calcule son gradient par rapport à chaque poids w , puis on modifie les poids dans le sens inverse du gradient [64].

III.3.1.1. Fonctions de transfert

Dans la littérature scientifique, il se trouve que plusieurs possibilités existent. Les fonctions de transfert pouvant être utilisées comme fonction d'activation [65, 66], les plus utilisées dans le contexte d'un neurone sont les fonctions : seuil, linéaire, sigmoïde, et tangente hyperbolique.

- **La fonction seuil:** applique un seuil sur son entrée, plus précisément une entrée négative ne passe pas le seuil, la fonction retourne alors la valeur 0 (on peut interpréter ce 0 comme signifiant faux), alors qu'une entrée positive ou nulle dépasse le seuil, et la fonction retourne 1 (vrai). La fonction seuil est donnée par la relation suivante:

$$f(x) = \begin{cases} 1 & \text{si } x > b \\ 0 & \text{si } x \leq b \end{cases} \quad (\text{III.3})$$

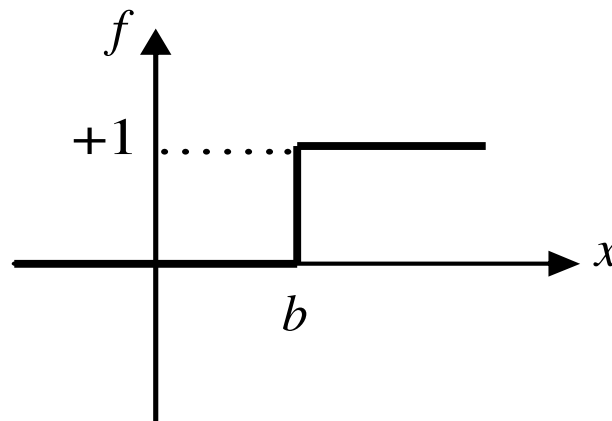


Figure III.5. Fonction de transfert seuil.

- **La fonction linéaire :** est très simple, elle affecte directement son entrée à sa sortie (voir la Figure III.6), leur équation est donnée par :

$$f(x) = x - b \quad (\text{III.4})$$

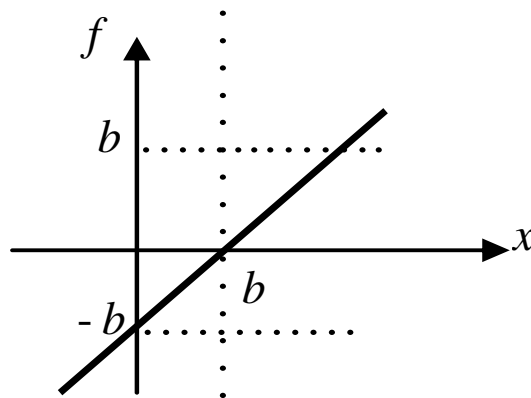


Figure III.6. Fonction de transfert linéaire.

- **La fonction sigmoïde** : la représentation graphique de cette fonction est illustrée dans la figure suivante, alors leur équation est donnée par la relation suivante:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (\text{III.5})$$

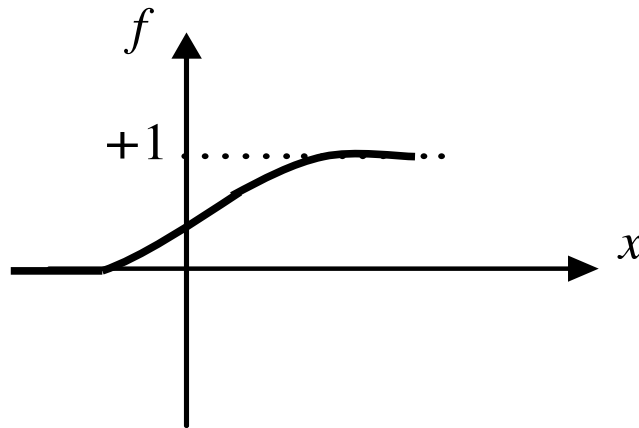


Figure III.7. Fonction de transfert sigmoïde.

- **La fonction tangente hyperbolique** : est une version symétrique de la sigmoïde, son équation est donnée par:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (\text{III.6})$$

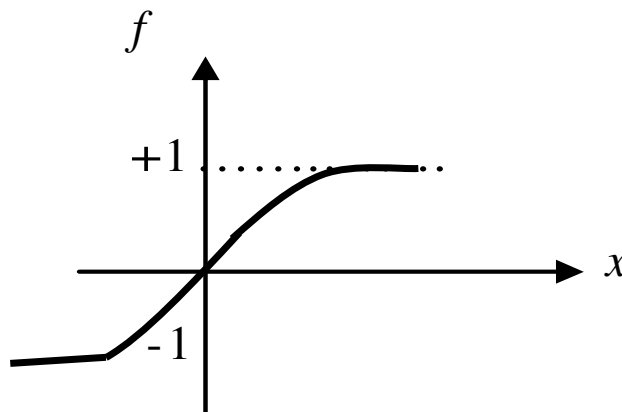


Figure III.8. Fonction de transfert tangente hyperbolique.

III.3.1.2. Mise en œuvre des réseaux neuronaux

La mise en œuvre des réseaux de neurones multicouches comporte à la fois une partie de conception, dont l'objectif est de permettre de choisir la meilleur architecture possible, et une partie de calcul numérique, pour réaliser l'apprentissage d'un réseau de neurones. Nous

allons suivre une démarche reprise par Wierenga et Kluytmans [67] qui est composée de quatre étapes principales :

Etape 01: Fixer le nombre de couches cachées

Mis à part les couches d'entrée et de sortie, l'analyse doit décider du nombre de couches intermédiaires ou cachées. Sans couche cachée, le réseau n'offre que de faibles possibilités d'adaptation, avec une couche cachée, il est capable avec un nombre suffisant de neurones, d'approximer toute fonction continue. Une seconde couche cachée prend en compte les discontinuités éventuelles.

Etape 02: Déterminer le nombre de neurones par couches cachées

Chaque neurone supplémentaire permet de prendre en compte des profils spécifiques des neurones d'entrée. Un nombre plus important permet donc de mieux coller aux données présentées mais diminue la capacité de généralisation du réseau.

Etape 03 : Choisir la fonction d'activation

Nous considérerons la fonction logistique pour le passage de la couche d'entrée à la couche cachée. Le passage de cette dernière à la couche de sortie sera selon nos types de variables: soit linéaire, soit sigmoïde, soit tangente hyperbolique.

Etape 04 : Choisir l'apprentissage

A partir d'une architecture de réseau de neurones donnée et des exemples disponibles (la base d'apprentissage) on détermine les poids optimaux, par l'algorithme de la rétropropagation des erreurs, pour que la sortie du modèle s'approche le plus possible du fonctionnement désiré.

L'apprentissage par rétropropagation nécessite la détermination du paramètre d'ajustement des poids synaptiques à chaque itération. La détermination du critère d'arrêt est aussi cruciale dans la mesure où la convergence peut passer par des minima locaux.

III.3.2. Réseaux de neurones à fonctions radiales de base RBF

Les réseaux RBF (Radial Basis Function) sont également considérés comme des réseaux de neurones, en termes d'architectures et d'algorithme d'apprentissage du réseau de neurones RBF est identique à un réseau de neurones MLP à une seule couche cachée. À l'exception de sa fonction d'activation qui est une fonction à base radiale, de type Gaussienne qui joue le rôle de fonction noyau de modèle. Dans les neurones de la couche d'entrée et la couche de sortie, le plus fréquent est d'utiliser les fonctions de transfert linéaires.

La structure du réseau RBF se compose de trois couches uniquement chaque couche se compose d'un certain nombre de neurones (nœuds), à savoir : une couche d'entrée (elle seulement retransmet les inputs à la couche cachée sans distorsion), une couche cachée (contient les neurones RBF) et une couche de sortie (ou de rendement qui n'est qu'une simple couche

qui contient une fonction linéaire). Comme le montre à la figure III.9, la sortie du réseau est une combinaison linéaire des sorties des neurones RBF multipliés par le poids de leur connexion respective [68].

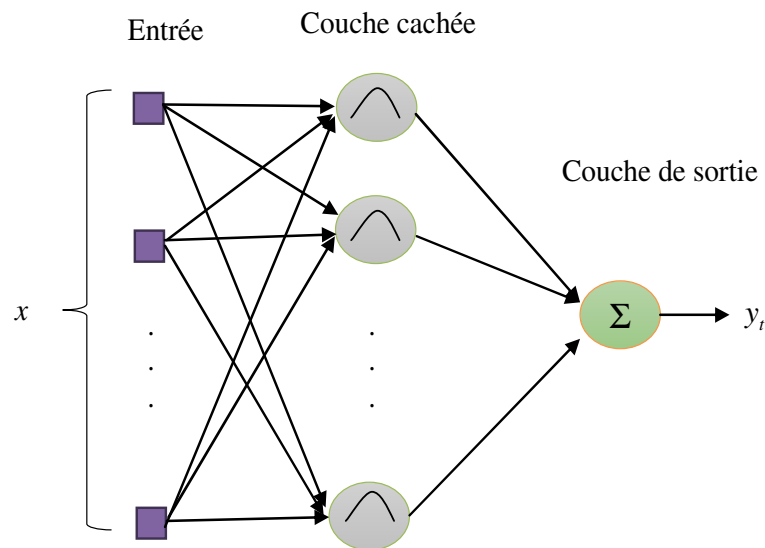


Figure III.9. Architecture d'un réseau de neurones à base de fonctions radiales

Le réseau RBF utilise le plus souvent la règle l'apprentissage de correction d'erreur et la règle par apprentissage compétitif. Il peut avoir un apprentissage qui combine en même temps l'apprentissage supervisé et l'apprentissage non supervisé [68].

Chaque neurone RBF contient une Gaussienne qui est centrée sur un point de l'espace d'entrée. Pour une entrée donnée, la sortie du neurone RBF est la hauteur de la gaussienne en ce point. La fonction Gaussienne permet aux neurones de ne répondre qu'à une petite région de l'espace d'entrée, région sur laquelle la gaussienne est centrée.

Donc il y a quatre paramètres principaux à régler dans un réseau RBF :

- Le nombre de neurones RBF (nombre de neurones dans l'unique couche cachée).
- La position des centres des Gaussiennes de chacun des neurones.
- La largeur de ces gaussiennes.
- Le poids des connexions entre les neurones RBF et le(s) neurone(s) de sortie.

Toute modification d'un de ces paramètres entraîne directement un changement du comportement du réseau RBF.

Nous disposons d'un ensemble d'entrées x_i et un ensemble de sorties y_i . L'approximation de y_i par un RBF sera notée \hat{y}_i . Cette approximation sera la somme pondérée de m fonctions gaussiennes Φ [69].

$$\hat{y}_t = \sum_{i=1}^m \lambda_i \Phi(x_i, c_i, \sigma_i) \quad (\text{III.7})$$

où

$$\Phi(x_i, c_i, \sigma_i) = e^{-\left(\frac{\|x_i - c_i\|}{\sqrt{2}\sigma_i}\right)^2} \quad (\text{III.8})$$

avec σ_i est l'écart type de la fonction d'activation, et c_i sont les centres des Gaussiennes de chacun des neurones.

La complexité du RBF est déterminée par le nombre de noyaux Gaussiens. Les différents paramètres à déterminer sont la position des noyaux Gaussiens (c_i), leurs variances (σ_i) et les facteurs multiplicatifs (λ_i) [68]. Les poids synaptiques qui existent entre la couche cachée et la couche de sortie sont calculés par l'algorithme de rétropropagation du gradient.

III.3.3. Machine à vecteurs de support SVM

Appelés aussi séparateurs à vaste marge développée en grande partie par Vladimir Vapnik en 1995 dans son livre "*The nature of Statistical Learning Theory*" peuvent être définies comme étant une famille d'algorithmes d'apprentissage permettant de résoudre efficacement des problèmes dans plusieurs domaines statistiques (classement, régression, fusion,...etc.) [70].

SVM sont l'un des classificateurs les plus répandus est développé comme un algorithme de classification binaire supervisée. L'idée essentielle de cette approche est appliquée à des données de l'espace d'entrée (appartenant à des classes différentes) séparables de façon non linéaire en transformant les données dans un espace de dimension supérieure où elles deviennent linéairement séparables (groupes : G1 et G2) comme montré dans la figure suivante.

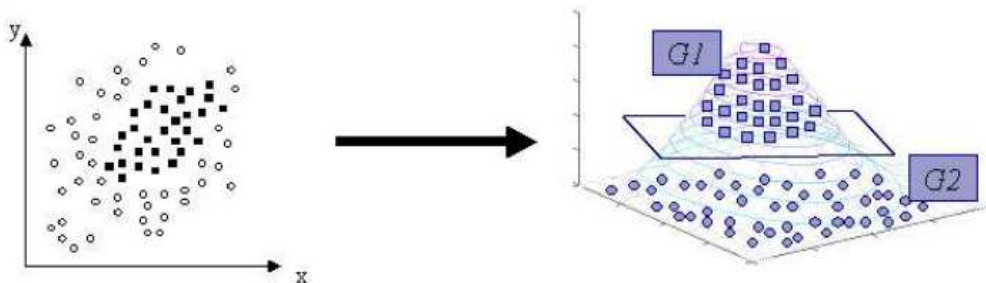


Figure III.10. Transformer l'espace de départ vers un espace de dimension supérieure.

III.3.3.1. Principe de fonctionnement

Le principe des SVM est de trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes, avec l'SVM, ce classificateur est linéaire appelé

«hyperplan» qui séparer les individus représentés dans un espace de dimension égale au nombre de caractéristiques, les individus étant alors séparés en deux classes possible quand les données à classer sont linéairement séparables, si cas contraire les données seront projetées sur un espace de plus grande dimension afin qu'elles deviennent linéairement séparables [71]. Dans la Figure suivante, on détermine un hyperplan qui sépare deux ensembles de points.

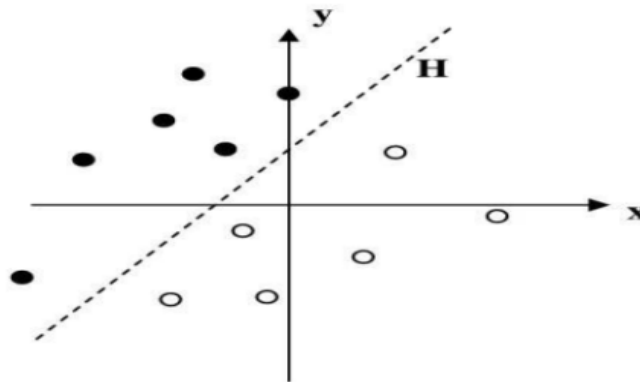


Figure III.11. Séparation de deux ensembles de points par un Hyperplan H.

Les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan, appelés vecteurs de support. Il existe une infinité d'hyperplans capable de séparer parfaitement les deux classes d'exemples. Le principe des SVM est de choisir celui qui va maximiser la distance minimale entre l'hyperplan et les exemples d'apprentissage (la distance entre l'hyperplan et les vecteurs supports), appelée la marge le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple (voir la Figure III.12).

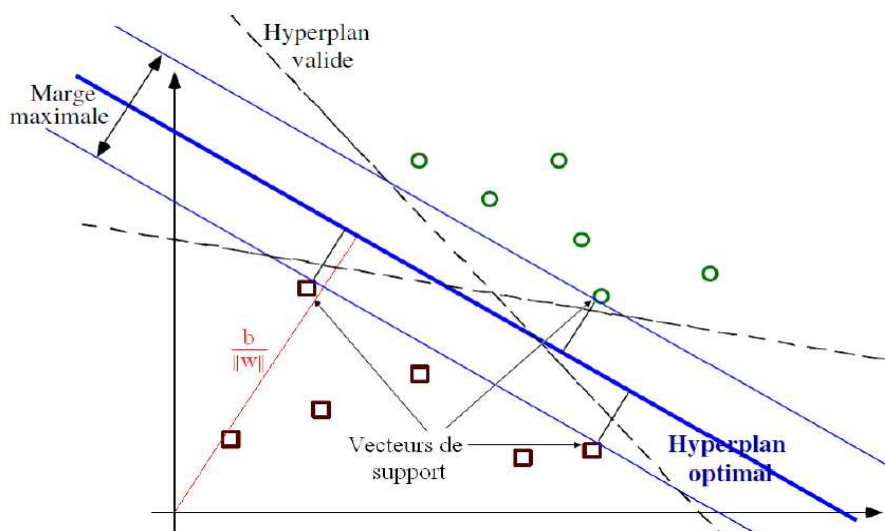


Figure III.12. Cas linéairement séparable.

S'il existe un hyperplan qui sépare les deux classes, les points de l'hyperplan sont décrits par l'équation suivante:

$$x_i \omega + b = 0$$

(III.9)

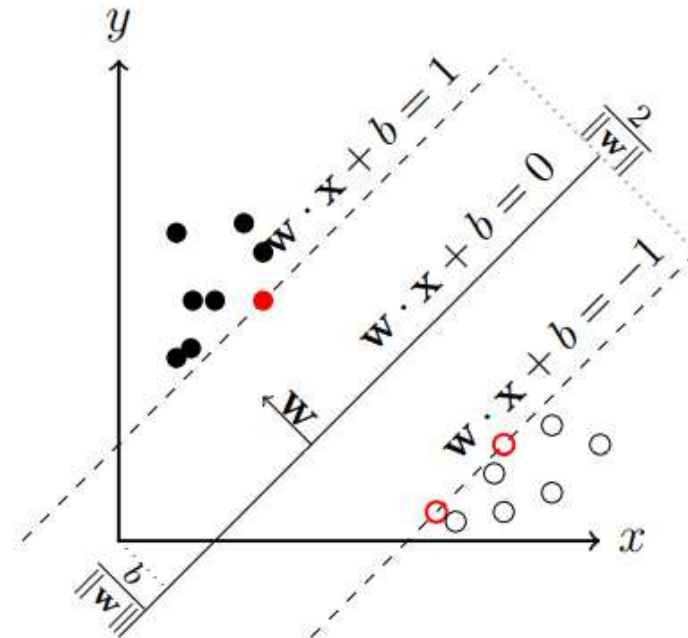


Figure III.13. Illustration d'un classificateur SVM dans le cadre d'une classification binaire. où : ω est la normale au plan et $b/\|\omega\|$ la distance entre l'hyperplan et l'origine et $2/\|\omega\|$ est la marge maximal, les points en rouge sont les vecteurs de support [72]. Pour estimer la classe d'un nouveau point x , on calcule [72]:

$$f(x) = \text{sign}(x \cdot \omega^0 + b^0) = \text{sign}\left(\sum_{vs} \alpha_i^0 y_i(x_i, x) + b^0\right) \quad (\text{III.10})$$

où : ω^0 , b^0 sont des constantes qui vérifiant l'hyperplan optimal et les α_i^0 sont les multiplicateurs de Lagrange.

III.3.4. Technique de validation

Dans la littérature scientifique, il existe plusieurs techniques pour l'optimisation des meilleurs paramètres de l'algorithme de classification (MLP, SVM, RBF): Validation simple (nous sommes intéressés à cette technique), Monte-Carlo validation Croisé, k -fold cross-validation

III.3.4.1. Validation simple

La technique de validation consiste tout simplement à diviser les données disponibles en deux ensembles (les ensembles d'apprentissage et de validation), sans qu'une donnée ne soit

commune. Un nombre non négligeable de données est bien entendu nécessaire dans l'ensemble de validation pour estimer correctement le critère de généralisation, donc pour optimiser les paramètres de classifieur (exemple : SVM). Souvent on garde 2/3 des données dans l'ensemble d'apprentissage et en réserver 1/3 pour la validation. Les ensembles d'apprentissage et de validation sont tirés aléatoirement parmi les N données disponibles [65].

III.3.4.2. Classification binaire (Bi-classe)

Les réseaux de neurones (MLP, RBF) ou machines à vecteur support SVM sont dans leur origine bi-classes (classification binaire) dans notre mémoire le problème est un problème binaire qui est contient sauf 2 classes qui sont signal interférence et signal utile. On parle donc de la classification binaire quand où les données appartiennent à 2 classes (le nombre de classes Y est égal à 2). Dans tous les cas, il s'agit d'un attribut qualitatif pouvant prendre un nombre fini de valeurs, autre cas une donnée peut appartenir à plusieurs classes dans ce cas est un problème multi-classes. En effet, plusieurs travaux abordent le problème multi-classe sous la forme de multiples problèmes de classification bi-classe. Dans le cas de la classification bi-classe, l'objectif est de retrouver tous les éléments qui appartiennent à une classe ou non [73].

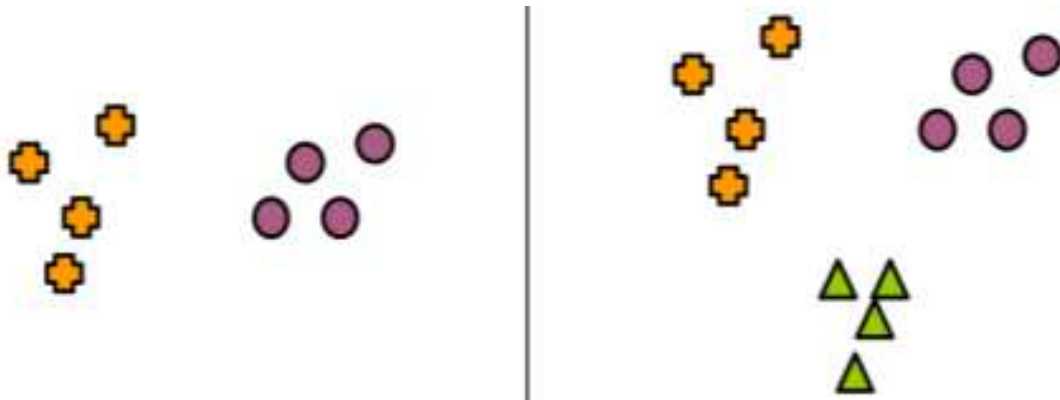


Figure III.14. Deux cas de la classification (bi-classe et multi-classe).

III.4. Détection des outliers

Selon Hawkins (1980), «une observation aberrante est une observation qui s'écarte tellement des autres observations qu'elle éveille la suspicion qu'elle a été générée par un mécanisme différent». Presque toutes les études qui considèrent l'identification des valeurs aberrantes comme leur objectif principal sont dans le domaine des statistiques.

La détection des outliers est utilisée depuis des siècles dans le domaine de la détection de *l'analyse de la robustesse du réseau ou modèle*, pour détecter les observations anormales des données et supprimer si nécessaire. Les outliers surviennent en raison de changements dans le

comportement du système, d'un comportement frauduleux, d'une erreur humaine, d'une erreur de périphérique, les interférences, bruit extérieur dans le cas des signaux GNSS... etc. Leur découverte peut identifier les interférences dans les récepteur GNSS avant qu'elles ne dégénèrent avec des conséquences potentiellement. Il permet d'identifier les erreurs et de supprimer leur effet polluant sur le jeu de données et ainsi d'affiner les données à traiter. Les méthodes originales de détection des outliers étaient stochastiques, mais maintenant, des techniques fondées sur des principes et systématiques sont utilisées, tirées de la gamme complète de l'informatique et des statistiques.

La détection des outliers est une étape primordiale dans de nombreuses applications d'exploration de données, à des applications importantes, comme le domaine des signaux GNSS. De nombreux algorithmes récents utilisent des concepts de proximité afin de trouver des valeurs aberrantes en fonction de leur relation avec le reste des données.

Dans cette section, nous retrouvons des approches qui contournent les problèmes mentionnés précédemment concernant la mise en pratique d'approches liées à la détection des interférences dans les récepteur GNSS. Dans les sous sections suivantes, nous avons proposé d'utiliser une méthode basée sur l'SVM à une seule classe, cette méthode est très sensible aux différents types de bruit ou de valeurs aberrantes (outliers).

III.4.1. Machine à vecteur de support à une classe OCSVM

La machine à vecteur de support à classe unique (souvent appelée 1SVM ou OSVM ou OCSVM) [74, 75] étendant la méthodologie SVM qui synthétise les données de la classe non représentée (un unique exemple de cette classe est placé à l'origine de l'espace de travail). OCSVM est souvent mentionné comme un classifieur de référence dans la classification one-class comme alternative aux approches par estimation de densité comme le Parzen ou la mixture de gaussiennes (les problèmes de classification à une classe sont résolus en estimant la densité cible).

Les SVM à une classe ont été conçus pour les cas dans lesquels une seule classe est connue et le problème est de détecter quoi que ce soit en dehors de cette classe sont utilisées pour séparer les données d'une classe spécifique, la classe cible, des autres données. Ils sont formés uniquement avec des exemples positifs, c'est-à-dire des points de données de la classe cible (ou le signal sans interférence). Ceci est également connu sous le nom de détection de nouveauté (novelty detection) et fait référence à l'identification automatique de phénomènes imprévus ou anormaux c'est-à-dire des outliers (les valeurs aberrantes, dans notre cas sont les interférences ou bruit), intégrées dans une grande quantité de données normales. Les SVM à une classe (OCSVM) apprennent une frontière de décision qui réalise une séparation

maximale entre les échantillons de la classe connue et l'origine (voir la Figure III.15). Seule une petite fraction des points de données est autorisée à se trouver de l'autre côté de la frontière de décision : ces points de données sont considérés comme des valeurs aberrantes (outlier) [76].

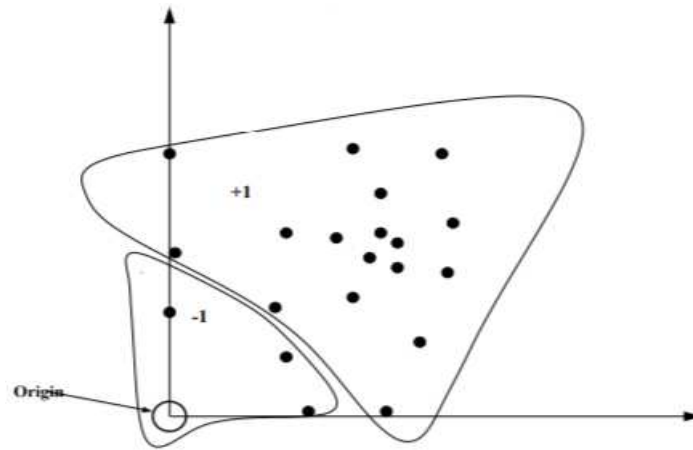


Figure III.15. One-class SVM.

III.4.1.1. Le principe d'OCSVM

OCSVM est une méthode prometteuse pour le problème de la classification à une classe, qui vise à détecter des échantillons qui ne ressemblent pas à la majorité de l'ensemble de données. Il utilise deux idées de l'algorithme original de la machine à vecteurs de support pour assurer une bonne généralisation: la maximisation de la marge et le mappage des données à un espace de caractéristiques de dimension supérieure induit par une fonction noyau [77].

Son principe de base consiste à séparer à l'aide d'un hyperplan avec une marge maximale les données (Target) du point origine, unique représentant des données de l'out-of-class. Cette approche est intéressante et originale car elle permet de simuler une approche de classification binaire standard et donc de résoudre la tâche de classification one-class à l'aide d'une formulation identique à celle du SVM standard.

Les données sont donc projetées sur une hypersphère de rayon unitaire centrée à l'origine de H . La phase d'apprentissage d'un OCSVM consiste à définir l'hyperplan séparateur $W = \{h \in H : \langle h, w \rangle_H - b = 0\}$ dans H tel que la marge $\frac{b}{\|w\|_H}$ soit maximum. La

Figure III.16 explique le principe d'OCSVM.

Les paramètres w et b sont les solutions du problème d'optimisation suivant :

$$\min_{w, \xi, b} \frac{1}{2} \|w\|_H^2 - b + \frac{1}{\nu l} \sum_{i=1}^l \xi_i \quad (\text{III.11})$$

délimitée pour les échantillons de la classe cible plutôt que sur l'estimation de la densité de probabilité. L'estimation de la zone délimitée est obtenue en séparant les échantillons cibles (dans un espace de caractéristiques de dimension supérieure pour les cas séparables non linéairement) de l'origine par un hyperplan à marge maximale qui est aussi éloigné que possible de l'origine [76].

III.5. Conclusion

Dans ce chapitre nous avons présenté les méthodes de l'intelligence artificielle connue par les méthodes de classification (MLP, RBF, SVM, OCSVM), et nous avons concentré sur les fondamentales de la méthode de machine à vecteurs de support qui une méthode de séparation entre les ensembles de données ont deux classes par un hyperplan dans une dimension supérieure. Ensuite on a exposé la méthode utilisée dans notre travail (OCSVM) qui est similaire de la méthode SVM à unique classe, leur principe est détecter les outliers ainsi faire la séparation de données basée uniquement sur l'estimation d'une zone délimitée pour les échantillons de la classe des données sans interférences.

Dans le prochaine chapitre et le dernier on va tester notre méthode qui est utilisé dans ce travail avec l'intermédiaire d'une méthode d'optimisation multi-objectif pour trouver une meilleure séparation entre les deux signaux utile et les signaux avec interférences dans les récepteur GNSS pour effectue les taches de traitement dans la chaîne de réception.

Chapitre IV

Résultats et discussion

IV. 1. Introduction

On a vu dans le premier chapitre que les systèmes de navigation jouent un rôle important et le besoin de tels systèmes augmente rapidement dans plusieurs domaines dans notre vie, mais malheureusement ces systèmes ont des vulnérabilités qui perturbent les services de communication, la faiblesse de ces systèmes est due à l'interférence dans les récepteurs GNSS.

Dans ce contexte, une solution intégrée pour l'atténuation des interférences employée au niveau du récepteur est requise afin de la détection automatiquement d'interférence GNSS, on peut recourir à des approches de l'apprentissage automatique (classification).

Dans ce travail, nous proposons une méthode pour l'identification et la suppression des interférences GNSS, d'où la nécessité d'une détection automatique. Ce travail combine un système de classification qui basé sur le classifieur OCSVM (one class support vector machine) avec une méthode heuristique basé sur l'optimisation uni-objectif GA et multi-objectif NSGAI, ces derniers fonctionnent sur une population de solutions potentielles avec quatre fonctions d'évaluations (fitness functions) pour déterminer automatiquement le paramétrage du classifieur adopté OCSVM et pour l'objectif de détecter et supprimer les valeurs aberrantes (les différents type d'interférence).

Dans ce dernier chapitre on a expliqué les bases de données utilisé dans notre travail, ensuite on a exposé la méthode proposée avec plusieurs exemples, enfin on a présenté les résultats de la simulation après la sélection des paramètres optimaux de OCSVM avec l'intermédiaire de GA ou NSGAI.

IV.2. Présentation du signal Galileo E5

Les performances des techniques d'atténuation des interférences ont été obtenues à l'aide d'un simulateur 'GE5-TUT' open-source dans le MATLAB. Dans tous nos scénarios de simulation, nous avons choisi le signal Galileo E5a-I, ce simulateur du système Galileo est dédié à la bande E5 [2].

Le Galileo E5 est constitué de deux bandes, la première est centrée sur 1176.45 MHz et la seconde est centrée sur 1207.140 MHz. En conséquence, le signal Galileo E5 est un signal modulé AltBOC (15,10) avec un débit de découpage de 10.23 Mbps. La Figure IV.1 illustre les densités spectrales de puissance (PSD) simulées de la bande E5 dans lesquelles nous observons au niveau du récepteur que le signal est submergé dans l'AWGN en raison de sa faiblesse comme les autres systèmes GNSS. Les détails concernant les paramètres du signal Galileo sont illustré dans le tableau IV.1 [3, 78]:

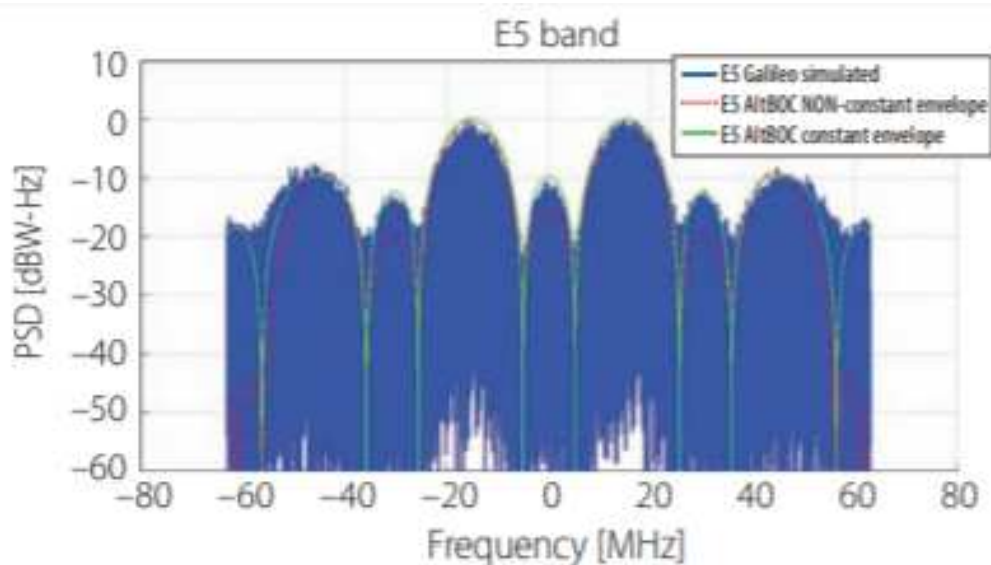
Tableau IV.1. Paramètres du signal E5aI.

Paramètres	valeur
Signal désiré	Galileo E5a-I
Fréquence d'échantillonnage (F_e)	31.500 MHz
Fréquence intermédiaire (F_i)	4.655 MHz
Intégration cohérente	1 ms
CNR 49 dB-Hz	CNR 49 dB-Hz

Le test de simulation est composé du signal Galileo E5aI et du signal d'interférence à ondes continues ajouté. Le rapport d'interférence au signal ISR du SCWI (l'interférence à onde continue unique) et du MCWI (l'interférence à onde continue multiple) et DME (l'interférence pulsé) est compris entre 10 à 60 dB. Il est a noté que le rapport est décrit par l'équation suivante:

$$ISR = 10 \times \log\left(\frac{I}{S}\right) \quad (\text{IV.1})$$

où : I est la puissance d'interférence et S est la puissance du signal GNSS.

**Figure IV.1.** La densité spectrale de puissance (PSD) de la bande E5.

Notez que le CWI utilisé est un signal sinusoïdal pur dans le cas de SCWI, situé au centre du lobe principal à la fréquence intermédiaire $F_i=4.655 \text{ MHz}$ de E5a bande qui correspond à une fréquence porteuse de 1176.45 MHz . Ce scénario présente l'attaque la plus dangereuse sur le système, et l'interférence MCWI simulée est la somme de plusieurs signaux sinusoïdes avec des fréquences différentes répartis sur la bande E5aI, avec celle de la fréquence intermédiaire $F_i=4.655 \text{ MHz}$ [2].

Pour plus d'illustration, des exemples de l'amplitude du signal E5aI dans le domaine temporel sans interférence et avec interférence, sont montrés dans les Figures IV.2 et IV.3. Dans le cas de signal contaminé par une interférence DME, le rapport interférence sur signal ISR est varié entre 10dB à 60dB avec un pas de 10dB . La représentation des signaux contaminé et non contaminé dans le domaine fréquentiel sont illustrés dans les Figure IV.4 et IV.5.

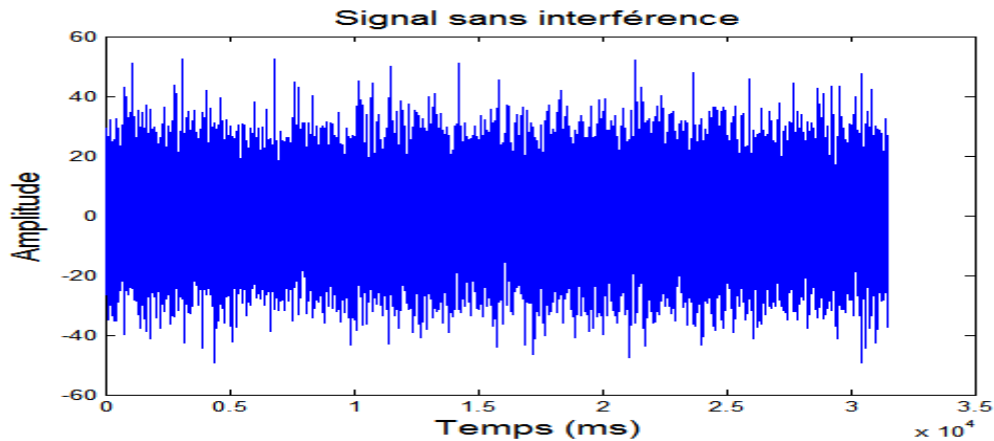
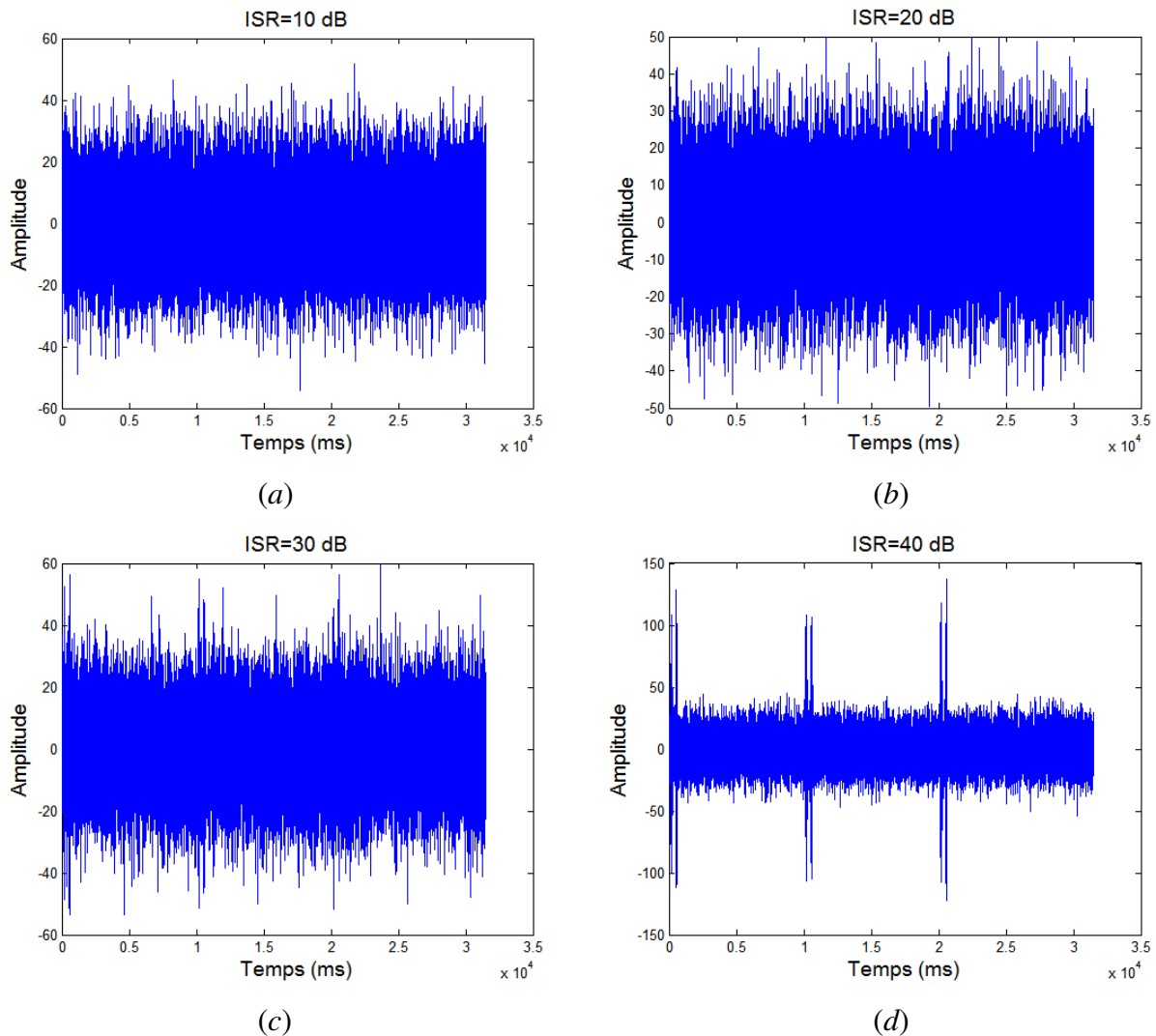


Figure IV.2. Amplitude du signal E5aI dans le domaine temporel sans interférence.



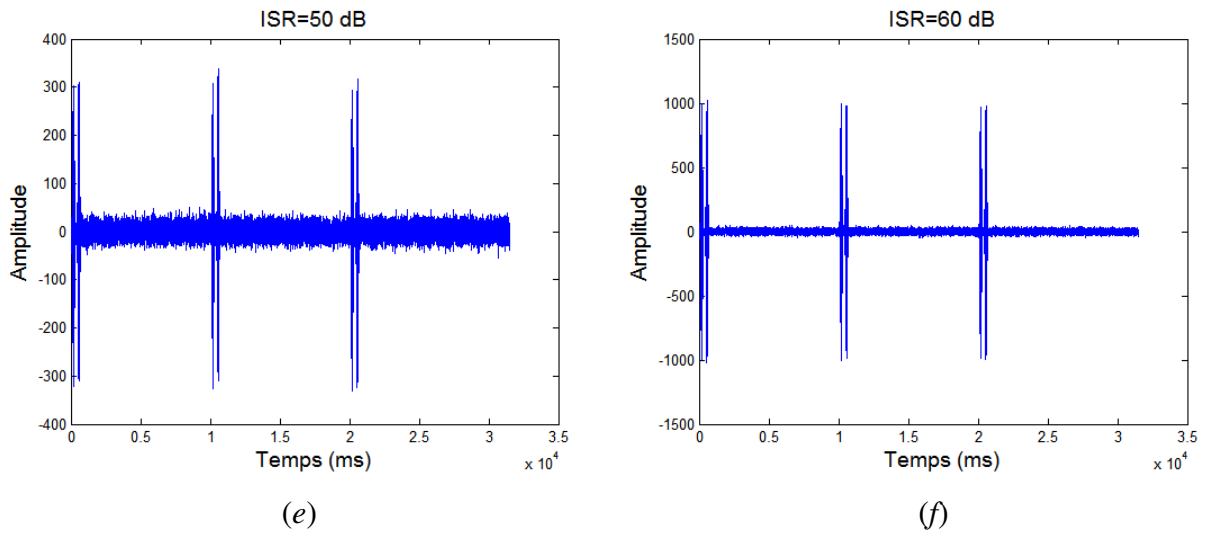


Figure IV.3. Amplitude du signal E5aI dans le domaine temporel contaminé par une interférence DMA, ISR varie entre 10dB à 60dB.

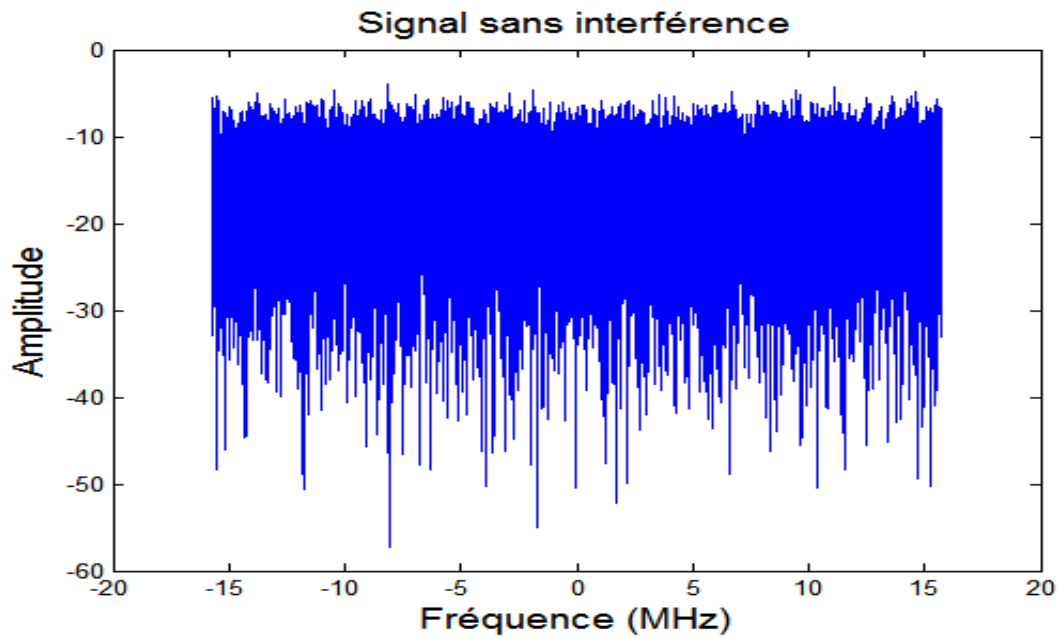


Figure IV.4. Amplitude du signal E5aI dans le domaine fréquentiel sans interférence.

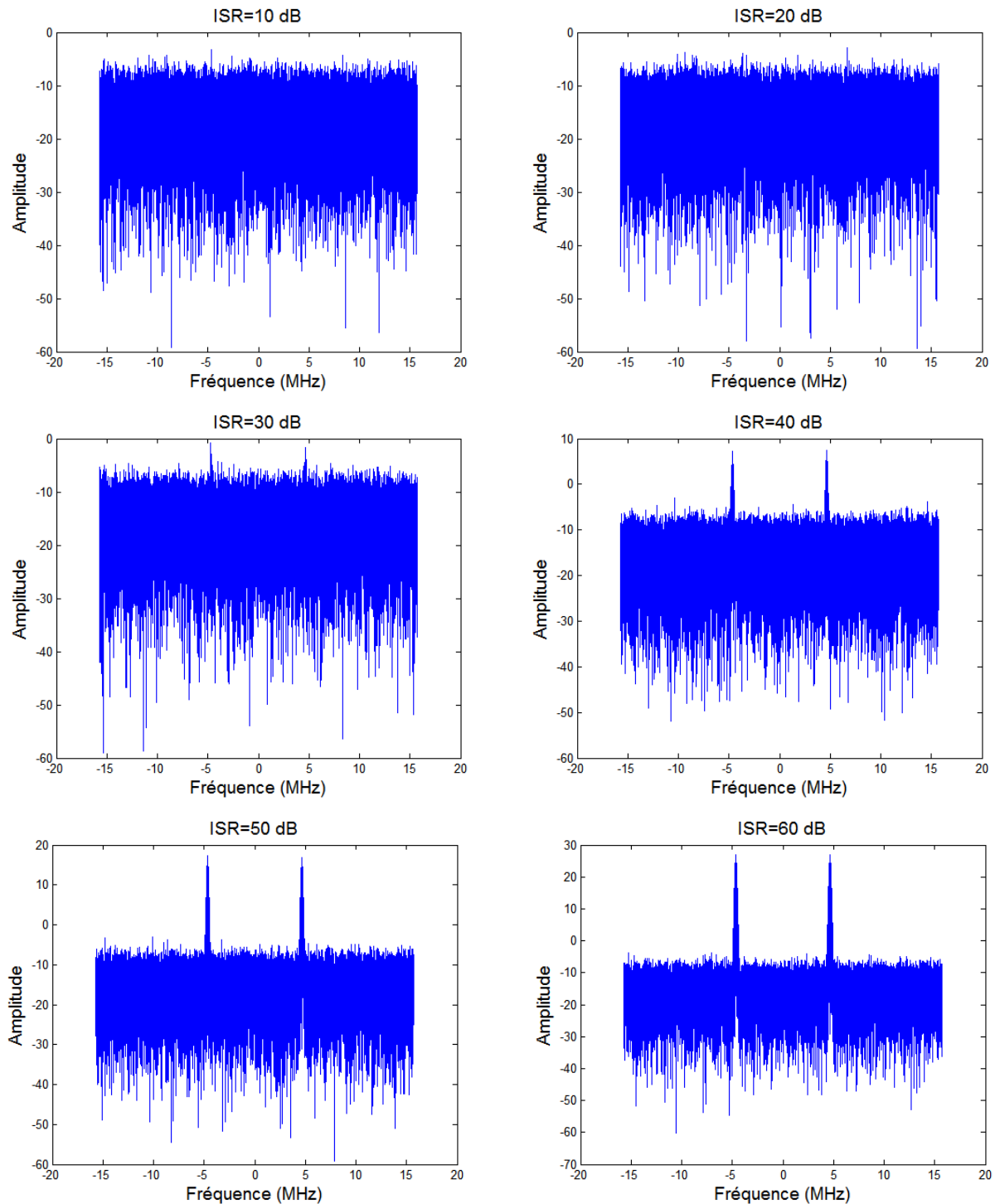


Figure IV.5. Amplitude du signal E5aI dans le domaine fréquentiel contaminé par des interférences DMA, ISR varie entre 10dB à 60dB.

IV.3. Méthode proposée

IV.3.1. OCSVM

Comme mentionné précédemment, le classifieur OCSVM permet de classer uniquement les échantillons d'une seule classe et de les distinguer de tous les autres échantillons possibles classés comme aberrants (outliers), il consiste à trouver une hypersphère dans laquelle la

plupart des échantillons d'apprentissage sont inclus dans un volume minimum. Dans OCSVM, la théorie de la machine à vecteurs de support (SVM) est utilisée pour estimer un hyperplan qui sépare au maximum les classes uniques et les valeurs aberrantes. En raison de la séparabilité non linéaire entre la classe unique et les valeurs aberrantes, le noyau gaussien est généralement utilisé dans OCSVM:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \quad (\text{IV.2})$$

où : $k(x_i, x_j)$ est le noyau gaussien de classifieur OCSVM qui permet de projeter des échantillons de l'espace d'origine vers l'espace des caractéristiques, x_i et x_j , ($i, j=1,2,\dots, n$), sont respectivement les $i^{\text{ème}}$ et $j^{\text{ème}}$ cellules unitaires, et σ est la largeur de gaussienne d'OCSVM.

La projection $k(x_i, x_j)$ des échantillons de l'espace d'origine vers l'espace des caractéristiques permet de séparer un échantillon de l'origine au moyen de la fonction de décision. Donc un échantillon x est accepté lorsque $f(x) > 0$, sinon, il est rejeté.

Le jeu de données $\{x_1, x_2, \dots, x_n\}$ est utilisé pour entraîner le modèle OCSVM initialisé avec les paramètres μ et σ . Selon le modèle OCSVM formé, le score d'anomalie de chaque échantillon x est calculé par l'équation suivante:

$$f(x) = \sum_{i=1}^n \alpha_i [k(x_i, x_j) - k(x_i, x)], \quad j \in (1, 2, \dots, n) \quad (\text{IV.3})$$

où : $f(x)$ est le score d'anomalie de l'échantillon x , et α_i ($i = 1, 2, \dots, n$) est le paramètre de Lagrange.

OCSVM est un processus de classification binaire qui classe les n échantillons en classe unique et en valeurs aberrantes. Deux paramètres principales sont : μ et σ affectent directement les performances du modèle OCSVM pour la séparation parfait entre le signal utile et les valeurs aberrantes qui sont des interférences.

- Le paramètre μ ($0.1 \leq \mu \leq 0.9$) est utilisé pour contrôler le pourcentage des échantillons considérés comme valeurs aberrantes contenant dans l'ensemble de donnée, c'est-à-dire que les n échantillons ne contiennent pas plus de $\mu \times n$ valeurs aberrantes.
- Le paramètre σ ($10^{-3} \leq \sigma \leq 10$) est la largeur de la gaussienne k , veut dire est l'échelle du noyau qui est un autre paramètre d'OCSVM qui permet de déterminer le rayon de l'hypersphère.

Pour montrer les performances de classifieur OCSVM, nous l'avons testée premièrement sur un ensemble de donnée choisi d'une façon aléatoire. La figure suivante représente un exemple de l'application de OCSVM, dans chaque cas en utilise des paramètres μ et σ différentes.

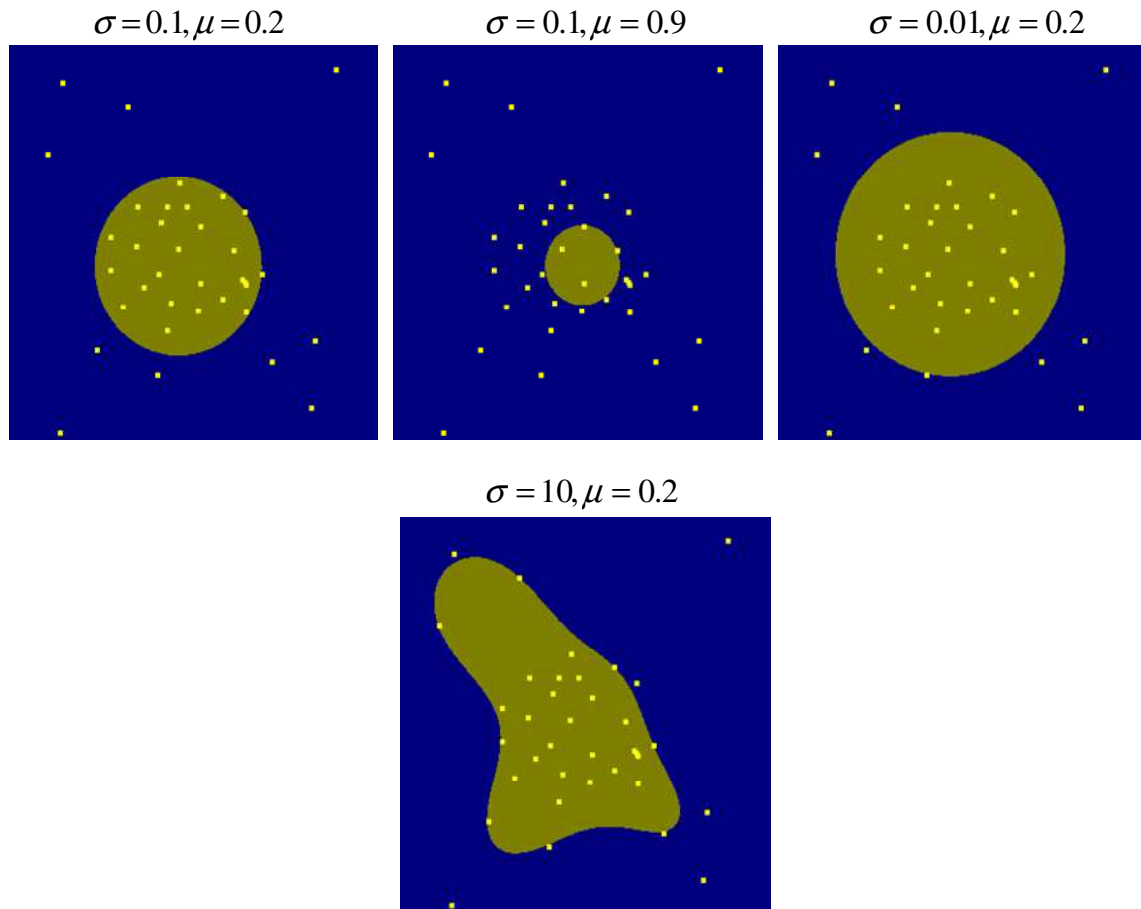


Figure IV.6. Exemple de classification OCSVM.

D'après la figure IV.6, les paramètres μ et σ joue un rôle très important dans le processus de classification par d'OCSVM, les points situés dans le cercle définit comme un signal utile (sans interférence) et les points en dehors du cercle sont les valeurs aberrantes, les paramètres qui ont un effet directement les performances du modèle OCSVM sont μ et σ .

- Le paramètre μ : Il exprime le pourcentage des outliers restant dans l'interface, veut dire si sa valeur est proche à 0 le pourcentage des valeurs aberrantes restant est petit, si sa valeur proche à 1 le pourcentage des valeurs aberrantes restant est plus grand. si sa valeur égale 1 le pourcentage des valeurs aberrantes est 100% toutes les données considérées comme des valeurs aberrantes.
- Le paramètre σ : exprime le nombre d'échantillons qui entouré par le noyau d'OCSVM, contrôler ce noyau par valeur de σ est exprimé l'échelle du noyau (plus σ est faible, plus le noyau est large), si σ est petite leur forme est cercle entoure où sont concentrés le plus

d'échantillons, si σ est grand le cercle transforme en une forme qui ne ressemble pas à un cercle et essayez d'entourer presque tous les échantillons d'interface.

Donc la détermination des valeurs optimales des paramètres μ et σ peut être améliorée les performances d'OCSVM. Pour cela, dans la suite nous avons utilisé les algorithmes génétiques comme une méthode d'optimisation.

Dans les sections suivantes, pour la construction des bases de données pour le classificateur OCSVM, on a appliqué deux étapes : (i) le signal dans le domaine temporel est divisé en L paquets non superposés. Nous avons trouvé, empiriquement, que la longueur appropriée de chaque paquet peut être $L=100$ coefficients. (ii) la deuxième étape est la division des données trouvées dans l'étape précédente en 3 ensembles : apprentissage, validation, et test :

- ❖ Le premier ensemble est utilisé pour estimer les paramètres du modèle OCSVM.
- ❖ L'ensemble de validation est utilisé pour trouver le meilleur modèle, c'est-à-dire la meilleure architecture.
- ❖ L'ensemble de test est utilisé pour estimer les performances du meilleur modèle sélectionné.

Notre base de données est constituée de deux classes (classe 1 : non interférence, classe 2 : interférence), la répartition de cette dernière en trois ensembles est donnée par le tableau suivant :

Tableau IV.2. Nombre total des échantillons apprentissage, validation et test.

	Classe 1	Classe 2
Apprentissage	100×100	100×100
validation	100×100	100×100
Test	630×100	630×100

IV.3.2. Configuration de l'algorithme génétique

Les algorithmes génétiques (AG) comme nous l'avons mentionné précédemment sont des techniques d'optimisation inspirées des principes biologiques. L'utilisation de l'AG nécessite la détermination de quatre paramètres fondamentaux, à savoir la structure chromosomique, les fonctions de fitness, l'initialisation et les opérateurs génétiques. Le succès d'un processus d'optimisation génétique dépend principalement de leur mécanisme qui traduit le problème d'optimisation considéré et guider la recherche vers la meilleure solution.

IV.3.2.1 Structure chromosomique

Nous allons considérer une population de N chromosomes C_m ($m=1,2,\dots,N$). Puisque l'objectif est de concevoir un classifieur robuste, donc chaque structure C_m est représentée par un

vecteur de deux dimensions ($m=2$), où chaque chromosome C_m se compose de deux paramètres de OCSVM : μ et σ . C_m peut donc s'écrire :

$$C = (\sigma, \mu) \quad (IV.4)$$

Concernant le problème de la détection et de la suppression des interférences consiste à trouver la meilleure sélection d'échantillons et le meilleur paramétrage de classifieur adoptés. En particulier pour le classifieur OCSVM, l'intervalle des valeurs de paramétrage de σ est varié entre 10^{-3} et 10, pour μ l'intervalle est entre 0.1 à 0.9.

IV.3.2.2. Fonctions de Fitness

La qualité de chaque chromosome peut être identifiée par l'évaluation des fonctions de fitness, le choix de la fonction de fitness est l'aspect le plus important de toute procédure d'optimisation génétique. Dans notre cas, nous avons utilisés quatre fonctions comme suit:

➤ La racine carrée de l'erreur quadratique moyenne $RMSE$:

$RMSE$ est une mesure fréquemment utilisée des différences entre les vraies valeurs (valeurs d'échantillon du signal original sans interférence) et les valeurs de signal récupéré après la suppression des interférences. La première fonction fitness est l' $RMSE$ donnée par l'équation suivante :

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (IV.5)$$

où : y_i est le signal original, \hat{y}_i est le signal récupéré, et N est le nombre d'échantillons.

➤ Le coefficient de corrélation R :

R indique le degré de similitude entre les deux signaux : le signal récupéré et le signal propre origine, ce critère donné par l'équation suivante :

$$R = \frac{\sum_{i=1}^N (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}}_i)^2}} \quad (IV.6)$$

où : \hat{y}_i est la moyenne du signal original et $\bar{\hat{y}}_i$ est la moyenne du signal récupéré.

Le coefficient de corrélation aura toujours une valeur qui se situe dans l'intervalle $[-1, 1]$. Plus la valeur du coefficient de corrélation linéaire est près de 1 ou -1, plus le lien linéaire entre deux signaux est fort. À l'inverse, plus sa valeur est près de 0, plus le lien linéaire entre les deux signaux est faible.

Dans un processus d'optimisation pour trouver une meilleure similarité entre deux signaux, il faut maximiser la valeur de R^2 ou minimiser la valeur de $1-R^2$. Dans notre cas nous avons minimisé les fonctions coûts. Pour cela, la deuxième fonction à minimiser utilisé dans notre travail est donnée comme suite :

$$r = 1 - R^2 \quad (IV.7)$$

➤ L'erreur de classification ERR :

Dans le cas de l'identification des interférences, nous avons besoin d'un critère de précision. Dans notre cas nous avons choisi de calculer le taux de classification ou d'identification, ce critère il est donné par la formule suivante:

$$Acc(\%) = \frac{P}{N} \times 100 \quad (IV.8)$$

où : P est Echantillons bien classées et N est le nombre d'échantillons des classes cibles.

Un "bon" classifieur doit présenter un accuracy ou précision élevé (proche de 100%). Pour cela, on utilise comme troisième fonction fitness à minimiser l'erreur de classification :

$$ERR(\%) = 100 - Acc(\%) \quad (IV.9)$$

➤ La distance Bhattacharyya :

Nous utiliserons une distance statistique *inter-classes* basée sur la mesure de distance $Dist$. $Dist(i, j)$ est la distance entre deux ensembles des données (dans notre cas est entre l'ensemble de donnée sans interférence et l'ensemble avec interférences i et j), définie comme suit :

$$Dist(i, j) = \frac{(\mu_i - \mu_j)^2}{4(\sigma_i^2 + \sigma_j^2)} + \frac{1}{2} \ln \left(\frac{\sigma_i^2 + \sigma_j^2}{2\sigma_i\sigma_j} \right) \quad (IV.10)$$

où : μ_i , μ_j et σ_i^2 , σ_j^2 sont respectivement la moyenne et la variance des données i et j .

La dernière fonction fitness à minimiser est :

$$D = -Dist(i, j) \quad (IV.11)$$

Les fonctions fitness utilisé dans notre travail sont : $RMSE$, r , ERR , D .

Lorsque le processus de sélection est terminé, l'étape suivante est consacrée à la reproduction de la population. Ceci est fait par des opérateurs génétiques qui sont: le croisement et la mutation comme nous avons mentionné déjà dans le deuxième chapitre. L'ensemble du processus est itéré jusqu'à ce qu'un critère de convergence défini par l'utilisateur soit atteint (minimisation des fonctions fitness).

IV.3.2.3. Opérateurs génétiques

Dans notre cas, nous avons utilisé la technique d'optimisation NSGAI, qui est basé sur le concept de dominance de *Pareto*. L'algorithme commence par générer une population aléatoire. Les individus (chromosomes) sélectionnés lors d'un tournoi de sélection surpeuplé subissent des opérations de croisement et de mutation pour former une population de descendants. Les populations descendantes sont ensuite combinées et triées en fronts de dominance décroissante. Après le processus de tri, la nouvelle population est remplie de solutions de différents fronts en commençant par la meilleure. Si un front ne peut effectuer que partiellement la génération suivante, une sélection de tournois est à nouveau utilisée pour assurer la diversité. Une fois que la population de la génération suivante a été effectuée, l'algorithme crée une nouvelle population de descendants et le processus se poursuit jusqu'à la convergence.

IV.3.3. Description de l'algorithme proposée

Dans cette section, nous décrivons les différentes phases sur lesquelles notre méthode de détection et suppression des interférences est proposée.

Dans un premier temps, on a quatre algorithmes qui surviennent lors de la combinaison des fonctions de fitness avec l'algorithme GA à objectif unique (uni-objectif): GA-RMSE et GA-r et GA-D et GA-ERR :

- 1) **Initialisation** : générer une population initiale de M chromosomes d'une manière aléatoire.
- 2) **Sélection** : évaluer la fonction de fitness de chaque chromosome de la population actuelle. Cette étape d'évaluation de la fonction de coût permet de trouver les meilleurs chromosomes, basée sur le principe que plus la fonction d'évaluation est meilleure, plus les chances d'être sélectionné sont élevées.
- 3) **Opérateurs génétiques** : la création d'une nouvelle génération avec l'utilisation d'opérateurs de croisement et de mutation pour obtenir une nouvelle population.
- 4) **Évaluation de la fonction fitness** : minimisez les quatre fonctions de fitness : $RMSE$ et r , ERR et D .
- 5) **Meilleure solution** : l'ensemble du processus est itéré jusqu'à le critère de convergence défini par l'utilisateur soit atteint, sinon Retour à l'étape 2.

Deuxièmement, on a cinq combinaisons entre les différents critères d'évaluations : D-r, D-RMSE, D-ERR, RMSE-ERR et r-ERR. Les différentes phases d'optimisations multi-objectif sont données par la suite :

Phase 1 Initialisation :

- Etape 1: générer une population initiale de chromosomes M de manière aléatoire.

Phase 2 Processus de recherche :

- Étape 2: construire un nouvel ensemble d'entraînement et de validation, ensuite calculer ses fonctions de fitness.
- Étape 3: effectuer des opérations aléatoires de sélection de tournois binaires, de croisement et de mutation afin de créer une population de descendants $Q(t)$ ayant la même taille N que la population des parents $P(t)$.

Phase 3 Optimisation :

- Étape 4: Fusionner les deux populations, c'est-à-dire $R(t)=P(t)\cup Q(t)$, pour garantir l'élitisme (mécanisme qui assure que tous les meilleurs chromosomes sont transmis à la génération suivante) et donc stabilité et convergence rapide du processus d'optimisation. Trier la population fusionnée $R(t)$ en différents fronts de rang de domination décroissant selon la méthode de tri non dominé. Notons qu'une solution est dite en dominer une autre si et seulement si les valeurs de ses fonctions de fitness sont partiellement inférieures à celles de l'autre solution.
- Étape 5: Créer une nouvelle génération $P(t+1)$ de taille N en choisissant les premières meilleures de N solutions de $R(t)$.
- Étape 6 : Si le critère d'arrêt (par exemple, le nombre maximum de générations et/ou une vérification de la variation de la racine de l'erreur quadratique moyenne RMSE au cours de la génération actuelle et de la dernière génération) n'est pas satisfait, définir $t \leftarrow t+1$ et retour à la phase 3.

Phase 4 Sélection de la meilleure solution :

- Étape 7: Le front de Pareto pour la génération finale contient les meilleures solutions (la sélection optimale des échantillons et le meilleur paramétrage du classifieur OCSVM). Concernant le traitement automatique, dans toutes les expériences de l'optimisation multi-objectif NSGAI, le choix final de la solution optimale entre plusieurs solutions dominantes du meilleur front de Pareto peut être identifié par le meilleur compromis entre deux critères d'évaluation (le plus proche de l'origine).

Pour le GA et NSGAI nous avons utilisé les paramètres standards dans le tableau suivant :

Tableau IV.3. Paramètres du GA et NSGAI.

les paramètres de AG et NSGAI	valeurs
La taille de la population	200
Le nombre maximum de générations	100
Le croisement	intermédiaire
La mutation	multi non-uniforme
Paramètre de l'opérateur de mutation (β)	5

IV. 4. Résultats de la simulation

La méthode de suppression d'impulsions dans le domaine temporel dans la partie numérique de la chaîne radiofréquence du récepteur GNSS qui a illustré dans la Figure IV.7, est la méthode la plus courante pour la contre-mesure de l'interférence pulsée. Ce type de circuit numérique fournit une excision d'interférence pulsée par le moyen d'une opération de seuillage sur les échantillons à la sortie du convertisseur ADC.

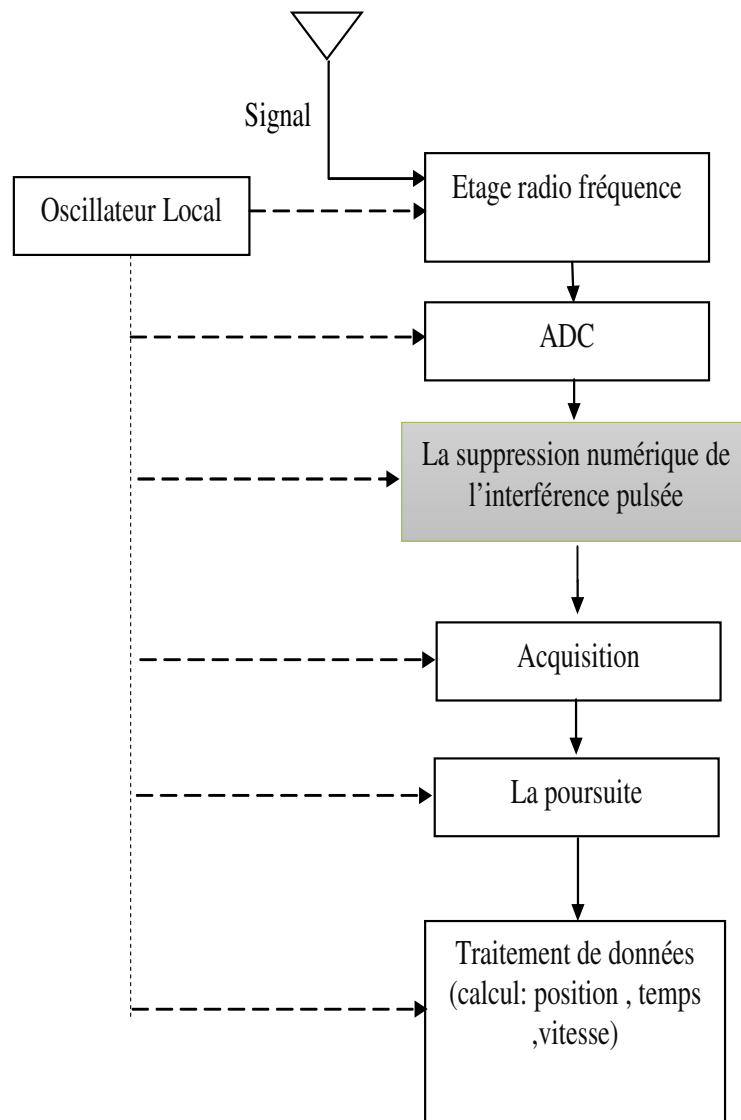


Figure IV. 7. Schéma bloc de la suppression numérique de l'interférence pulsée.

Fondamentalement, chaque échantillon est comparé à un niveau de seuil numérique, qui est établi en fonction d'une estimation de la puissance de bruit thermique uniquement «les impulsions sont courtes et ont une très grande amplitude par rapport au niveau de bruit». Afin

d'éviter le problème de seuillage, nous avons proposée une méthode d'identification et de suppression automatique des interférences basée sur le classifieur OCSVM.

Dans les sections suivantes, nous avons effectué plusieurs expériences pour évaluer les performances de la méthode proposée et leur efficacité pour séparer le signal utile (sans interférence) et le signal interférence. La série d'expérience consistait à appliquer premièrement une classification OCSVM sans optimisation, après nous avons appliqué une classification OCSVM avec l'intermédiaire d'une méthode d'optimisation basée sur les algorithmes génétiques, dans cette cas nous avons utilisé deux scénarios : uni-objectif (GA) et multi-objectif (NSGAI).

Il convient de noter que les performances de la méthode suggérée sont évaluées en termes de:

- La qualité du signal en terme de RMSE avant le bloc d'acquisition dans laquelle la comparaison des DSP renseigne sur l'état du signal reçu sans et avec le bloc d'atténuation des interférences.
- Une autre métrique utilisée dans de nombreux domaines tels que : applications GNSS, domaine biomédical et autres. Ce critère est bien connu sous le nom de coefficient de corrélation R et il indique le degré de similarité entre deux signaux (récupéré et l'origine).

IV.4.1. Classification OCSVM sans optimisation

Dans le premier scénario, on utilise le classifieur OCSVM sans optimisation comme une méthode d'identification et de suppression des interférences. Les résultats montrés dans le tableau IV.4 présente RMSE et R obtenues pour les six cas d'ISR (ISR varie entre 10dB à 60dB). Nous avons vu que le meilleur cas par rapport aux critères : le minimum RMSE et le maximum R est obtenu dans le cas ISR=50dB. La Figure IV.8 représente l'identification (en couleur rouge) et la suppression des interférences dans le domaine fréquentiel (densité spectrale de puissance DSP).

Tableau IV.4. Les résultats obtenus pour les différents cas d'ISR (10 à 60 dB).

ISR (dB)	RMSE	R
10	7.7615	0.7882
20	7.6333	0.7918
30	8.2708	0.7515
40	6.7413	0.8427
50	5.8380	0.8861
60	6.2233	0.8689

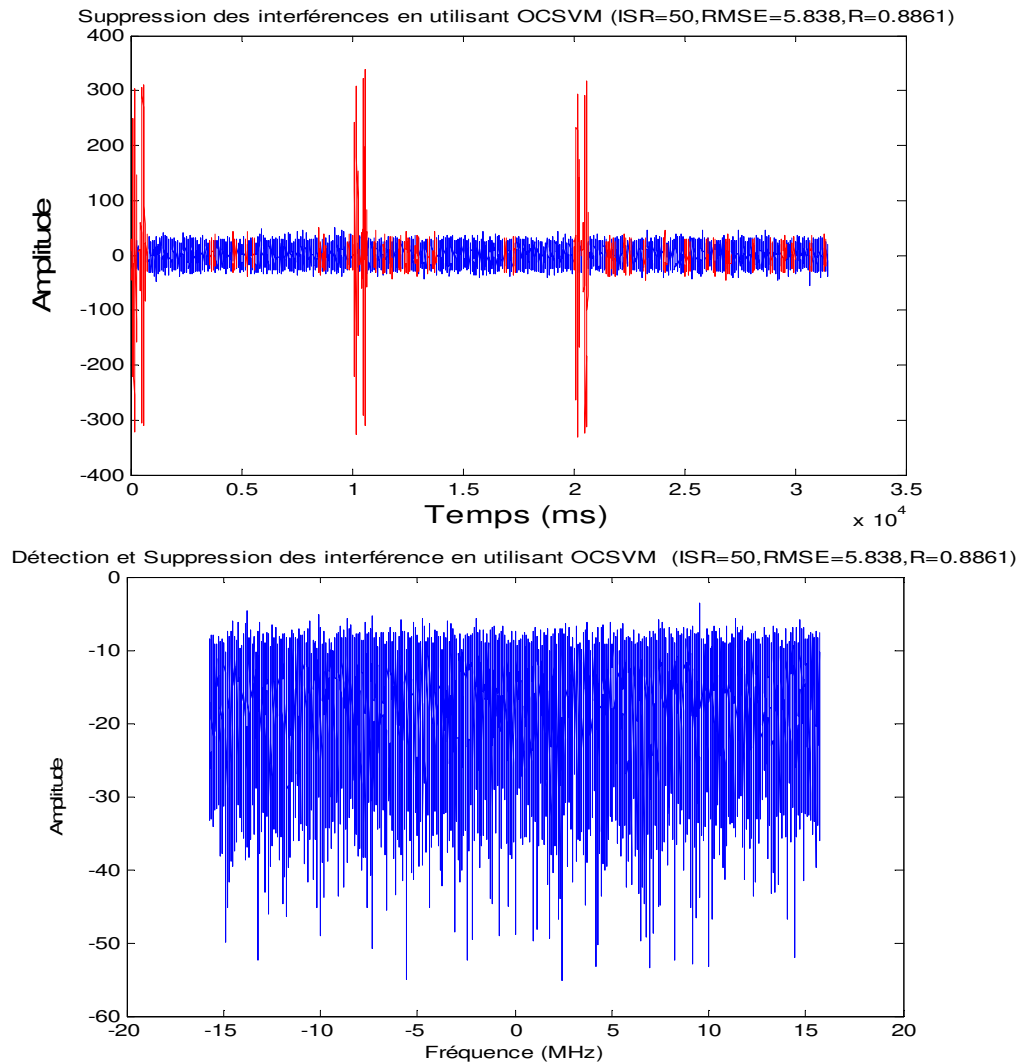


Figure IV.8. Détection et suppression des interférences dans le cas $ISR=50$ dB en utilisant OCSVM.

IV.4.2. Optimisation simple objectif GA

Dans le deuxième scénario, nous avons fait la même étape précédente : classification par l'utilisation d'OCSVM, mais dans ce cas les paramètres de OCSVM sont optimisés par l'algorithme génétique GA. Cependant, comme indiqué dans les sections précédentes, nous avons quatre fonctions fitness à minimiser ($RMSE$, r , ERR , D).

Pour la première fonction RMSE, nous avons calculé RMSE, et R sur les signaux GNSS, avec les différentes ISR comme mentionnés dans le tableau IV.5. En observation générale, il apparaît clairement que la méthode GA-RMSE est la plus performante en termes de RMSE et R dans le cas $ISR=60$ dB. La Figure IV.9 montre le comportement de convergence vers la solution optimale (l'évolution de la fonction fitness) au fur et à mesure des générations, avec une erreur RMSE égale à 5.11, et une $ISR=60$ dB. La Figure IV.10 représente l'identification et la suppression des interférences dans le domaine fréquentiel.

Tableau IV.5. Les résultats obtenus de GA-RMSE pour les différents cas d'ISR 10 à 60 dB.

ISR (dB)	RMSE	R
10	8.06	0.77
20	8.02	0.77
30	8.15	0.76
40	8.12	0.76
50	8.24	0.76
60	5.11	0.91

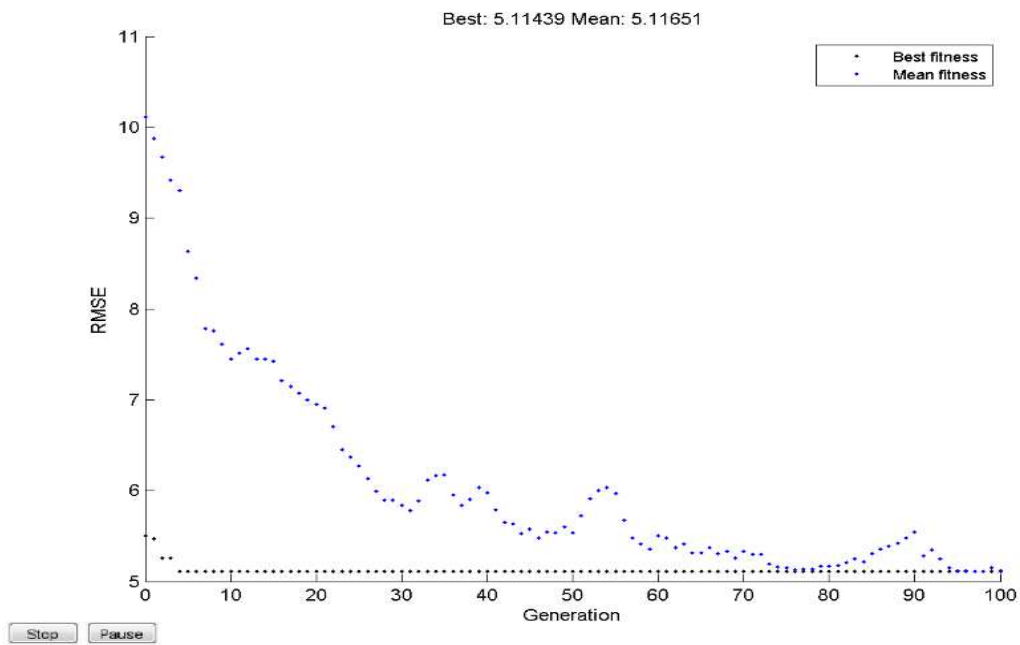
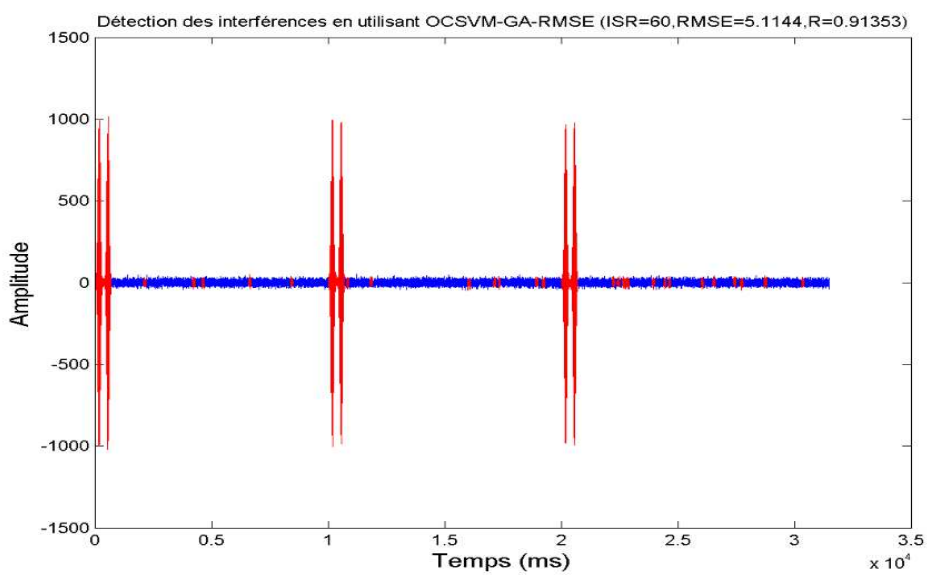


Figure IV.9. Evolution de fitness et la détection et suppression (ISR=60 dB) par GA-RMSE.



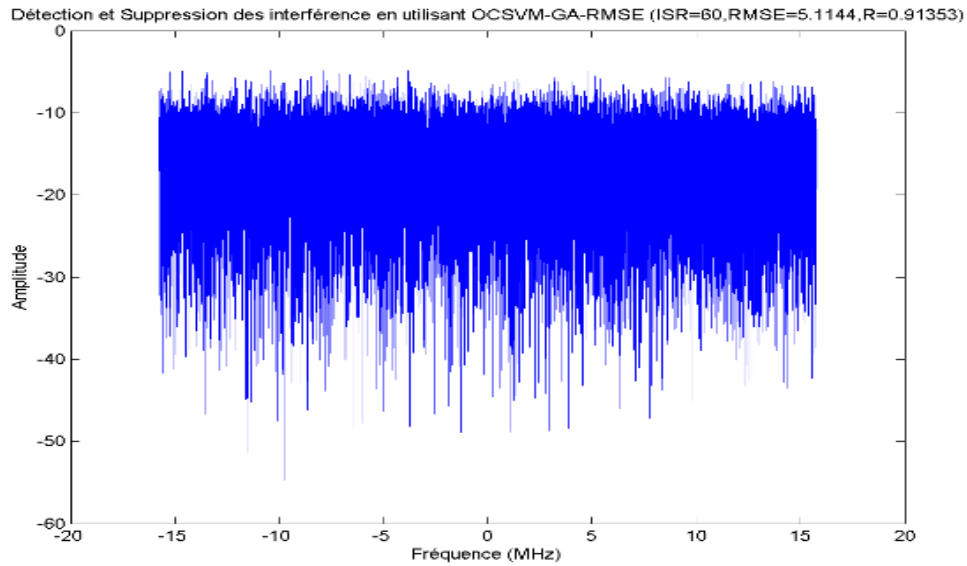


Figure IV.10. Détection et suppression des interférences en utilisant l'algorithme GA-RMSE dans le cas $ISR=60$ dB.

Les Figures IV.11, IV.12 et IV.13, montrent le comportement de convergence vers la solution optimale (l'évolution des fonctions fitness : r , ERR , D) au fur et à mesure des générations, nous avons illustré les meilleurs résultats trouvés par les différentes fonctions d'évaluations.

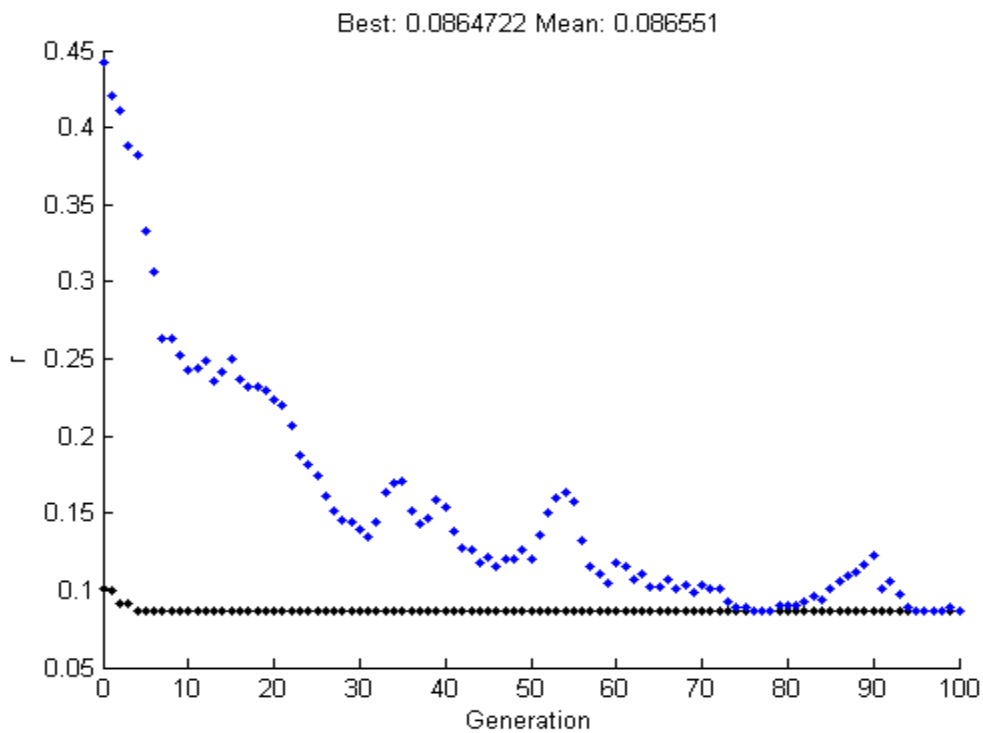


Figure IV.11. Evolution de la fonction fitness r en fonction de nombre de générations, $ISR=60$ dB.

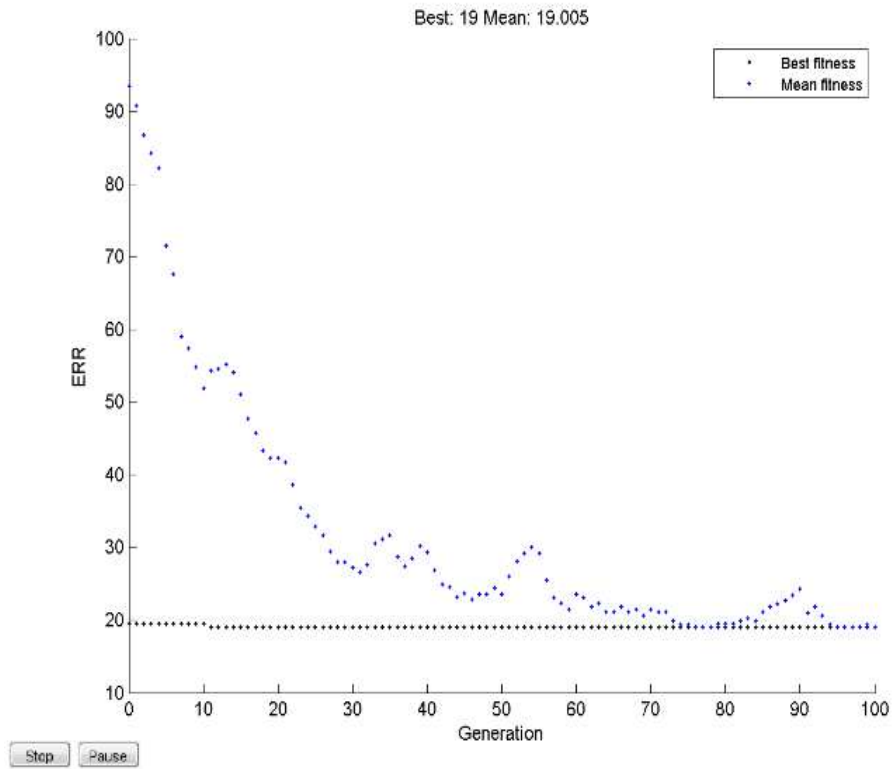


Figure IV.12. Evolution de la fonction fitness ERR en fonction de nombre de générations, $ISR=60dB$.

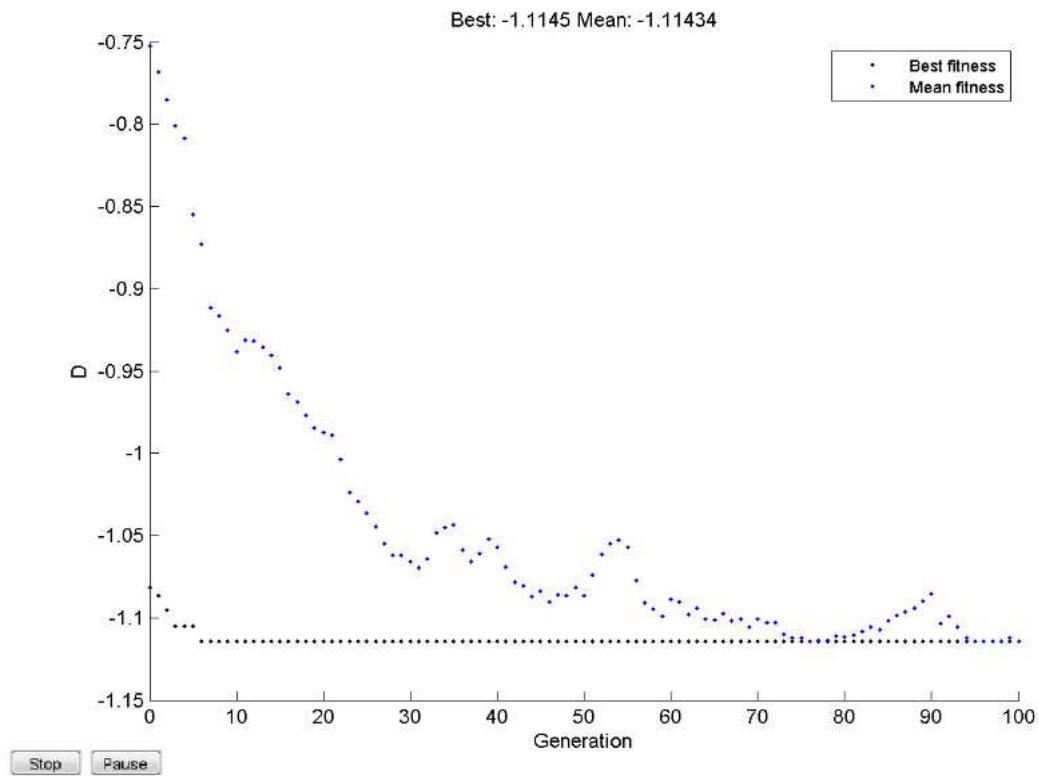


Figure IV.13. Evolution de la fonction fitness D en fonction de nombre de générations, $ISR=60dB$.

Le tableau IV.6 illustre bien les résultats obtenus avec les différentes méthodes d'optimisations pour les critères d'erreurs : r , ERR , et D .

L'analyse des résultats dans les Tableaux IV.5 et IV.6, montrent que les résultats ne sont pas forcément améliorés si on utilise une uni-objectif par rapport le classifieur standard OCSVM (voir le Tableau IV.4).

Tableau IV.6. Les résultats obtenus par les différentes méthodes d'optimisation GA- r , GA- ERR et GA- D .

ISR (dB)	RMSE	R	RMSE	R	RMSE	R
Méthode	GA- r		GA- ERR		GA- D	
10	8.06	0.77	11.26	0.45	11.26	0.45
20	8.02	0.77	11.14	0.45	11.13	0.45
30	8.15	0.76	11.30	0.43	12.51	0.06
40	8.12	0.76	11.23	0.44	8.13	0.76
50	8.24	0.76	9.52	0.65	8.24	0.76
60	5.11	0.91	5.41	0.90	5.11	0.91

IV.4.3. Optimisation Multi-objectif NSGAI

Pour comprendre le comportement de l'algorithme NSGAI, nous commençons par choisir une combinaison entre deux critères d'évaluations par exemple r et D . La Figure IV.14 représente l'évolution de cette combinaison D - r en fonction de nombre de générations.

Nous essayons de minimiser simultanément de nombreuses fonctions de fitness concurrentes, il est difficile de trouver la solution optimale qui satisfasse parfaitement à toutes les fonctions de fitness. L'objectif principal de l'algorithme d'optimisations multi-objectif est de trouver le front de Pareto, qui est l'ensemble des solutions non dominées. La Figure IV.14 montre le comportement de convergence vers la solution optimale (l'évolution de la fonction fitness), veut dire que les paramètres de classifieur OCSVM (μ et σ) pour chaque solution non dominée (le meilleur front de Pareto à la fin du processus d'optimisation). Les meilleures solutions sont celles qui sont les plus proches de l'origine dans le front de Pareto et sont mises en évidence par des points rouges.

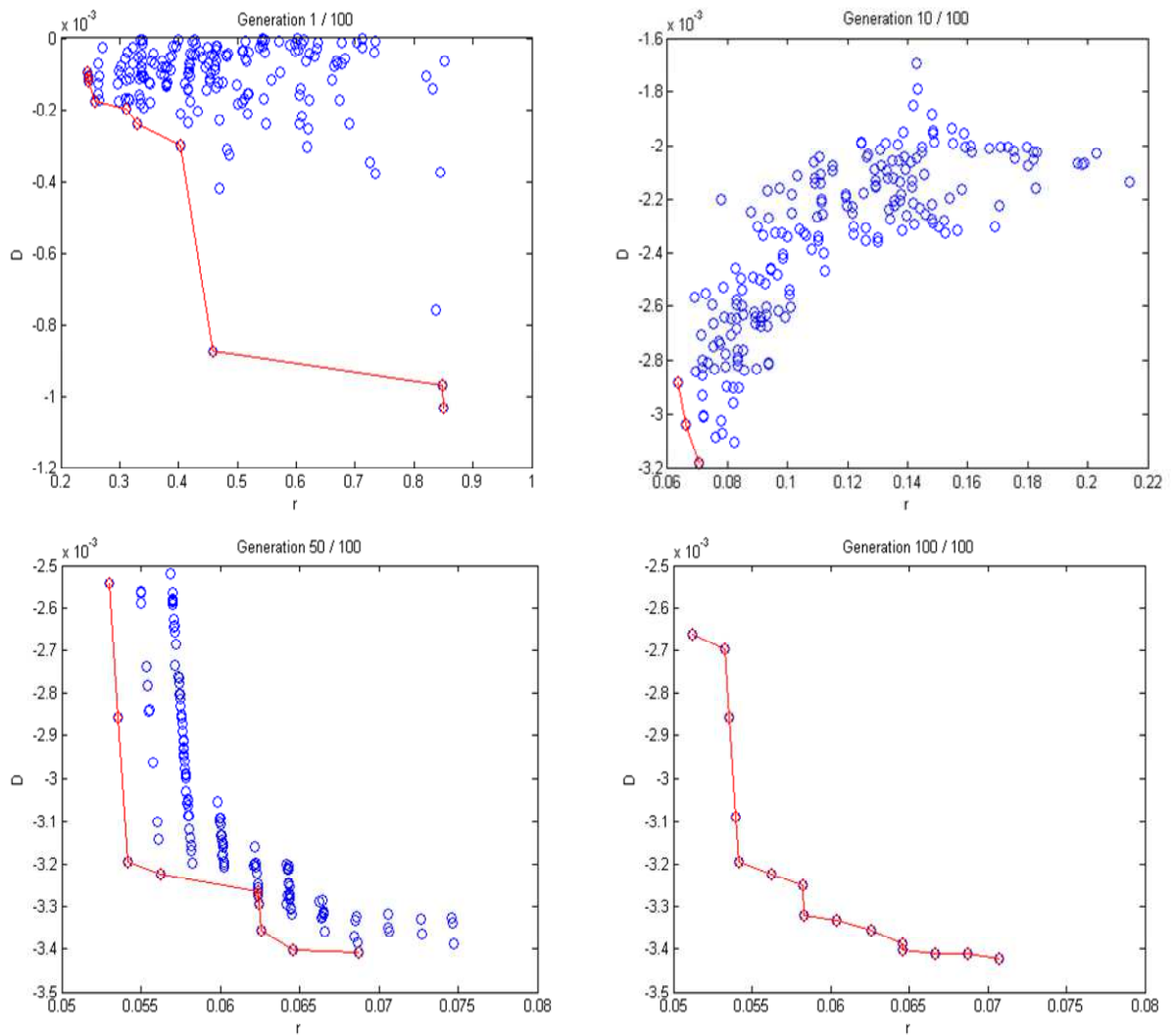


Figure IV.14. Evolution de l’algorithme multi-objectif NSGAI-D-r en fonction de nombre de générations.

Nous avons fait les mêmes simulations pour les combinaisons des fonctions fitness : D-RMSE, D-ERR, RMSE-ERR et r-ERR. Les Figures IV.15, IV.16, IV.17 et IV.18, montrent les meilleurs résultats en termes de RMSE et R, les représentations de ces dernier est en fonction de front de Pareto pour les quatre combinaisons précédentes et les différents ISR.

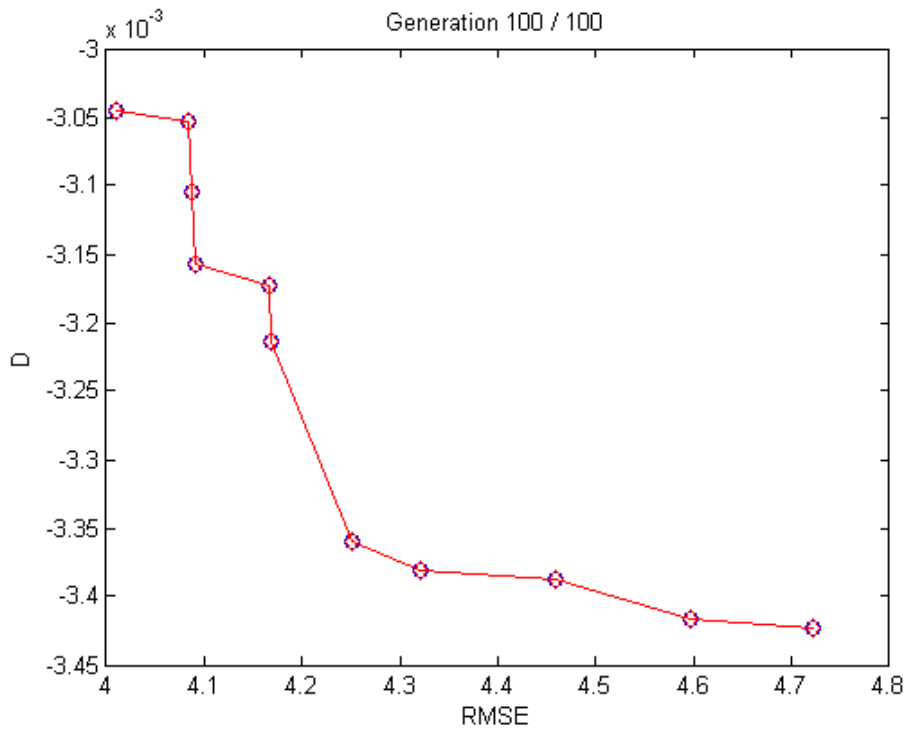


Figure IV.15. Front de Pareto pour l’algorithme génétique multi-objectif NSGAIID-RMSE.

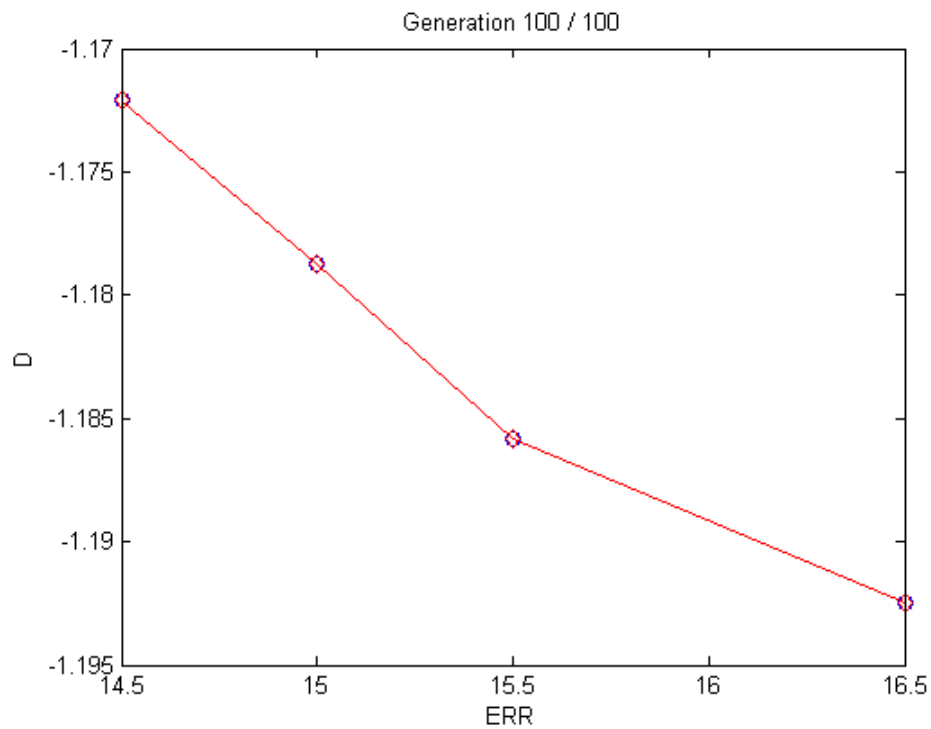


Figure IV.16. Front de Pareto pour l’algorithme génétique multi-objectif NSGAIID-ERR.

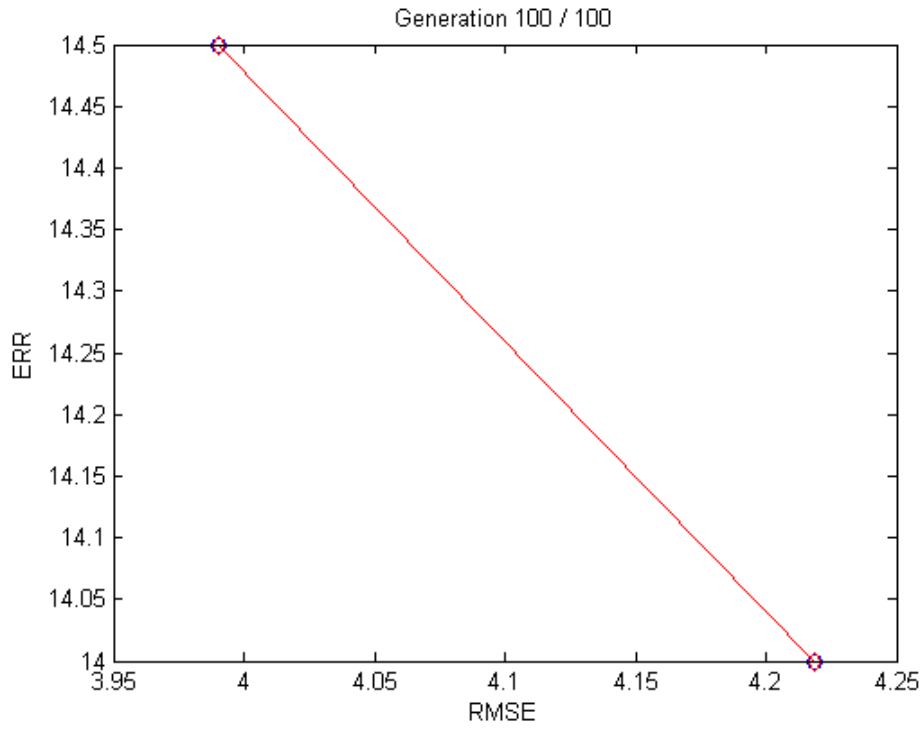


Figure IV.17. Front de Pareto pour l’algorithme génétique multi-objectif NSGAI-ERR-RMSE.

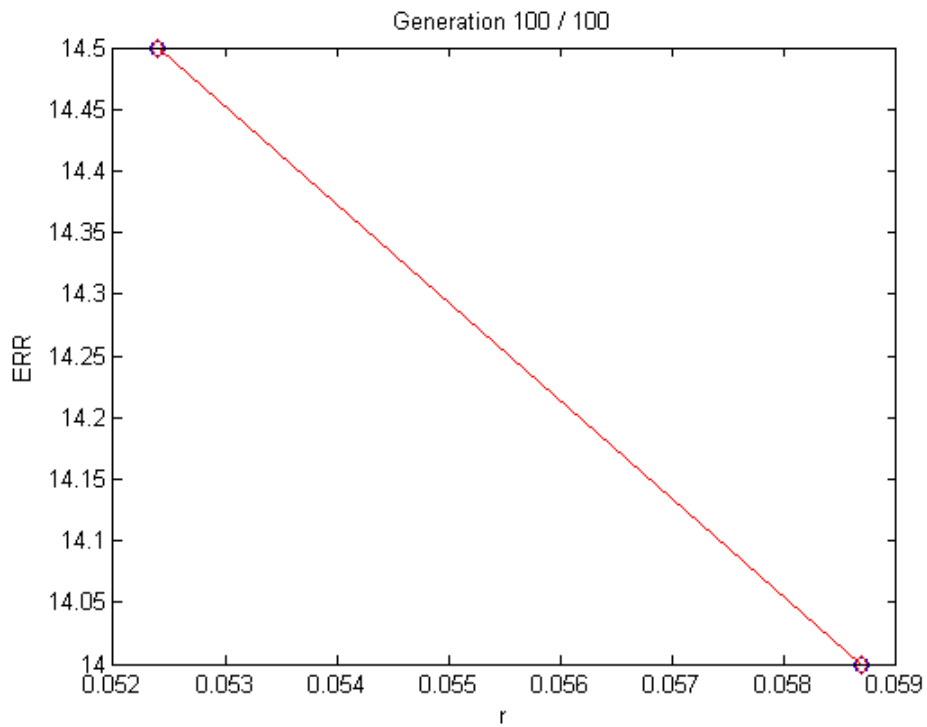


Figure IV.18. Front de Pareto pour l’algorithme génétique multi-objectif NSGAI-ERR-r.

D'après les figures précédentes, pour la combinaison D-RMSE, nous avons trouvé 11 solutions, la meilleure solution que nous avons choisie pour le classifieur OCSVM est le plus proche de centre. Dans le cas de NSGAI-D-ERR, nous avons trouvé quatre solutions, pour les deux combinaisons qui restent on a trouvé seulement deux solutions. Dans le cas NSGAI-RMSE-ERR et NSGA-r-ERR, on ne trouve pas le Front de Pareto.

L'optimisation NSGAI donne meilleur Pareto front, lorsqu'on combine le critère de similarité r avec la distance Bhattacharyya.

Le tableau IV.7 illustre les erreurs RMSE et les R obtenus pour les différents méthodes d'optimisations NSGAI et les différents rapports ISR.

Tableau IV. 7. Les résultats obtenus par les différentes méthodes d'optimisation multi-objectif NSGAI-D-r, NSGAI-D-RMSE, NSGAI-D-ERR, NSGAI-RMSE-ERR et NSGAI-r-ERR.

ISR(dB)	NSGAI-D-r		NSGAI-D-RMSE		NSGAI-D-ERR		NSGAI-RMSE-ERR		NSGAI-r-ERR	
	RMSE	R	RMSE	R	RMSE	R	RMSE	R	RMSE	R
10	3.98	0.95	4.01	0.95	4.78	0.93	4.05	0.95	4.06	0.95
20	3.99	0.95	4.01	0.95	4.86	0.92	3.99	0.95	3.99	0.95
30	4.67	0.93	4.72	0.93	4.74	0.93	4.72	0.93	4.68	0.93
40	4.72	0.93	4.71	0.93	5.14	0.91	4.71	0.93	4.71	0.93
50	4.92	0.92	4.92	0.92	5.16	0.91	5.16	0.91	5.03	0.92
60	4.30	0.94	4.30	0.94	4.50	0.93	4.37	0.94	4.43	0.94
moyenne	4.43	0.93	4.44	0.93	4.86	0.92	4.50	0.93	4.48	0.93

Les Figures IV.19 et IV.20, illustrent la comparaison entre OCSVM standard (sans optimisation) et les différents résultats obtenus par les cinq méthodes d'optimisation multi-objectif, on conclut que notre algorithme dépasse d'une manière significative le standard OCSVM en termes de RMSE et R.

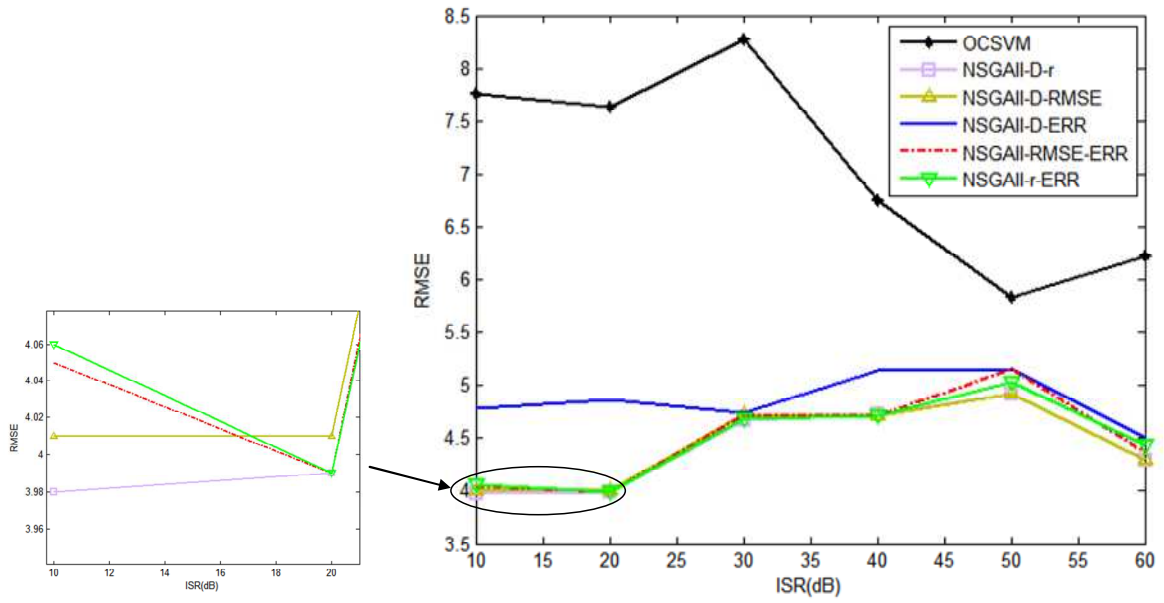


Figure IV.19. Evolution de RMSE en fonction des différentes valeurs ISR.

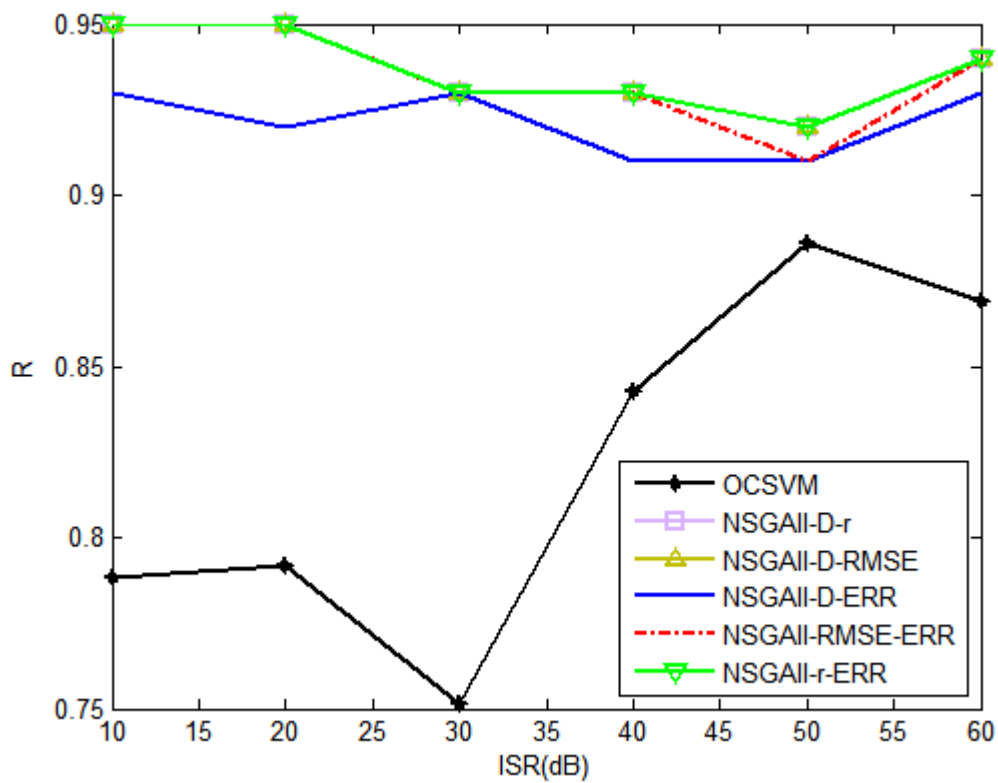
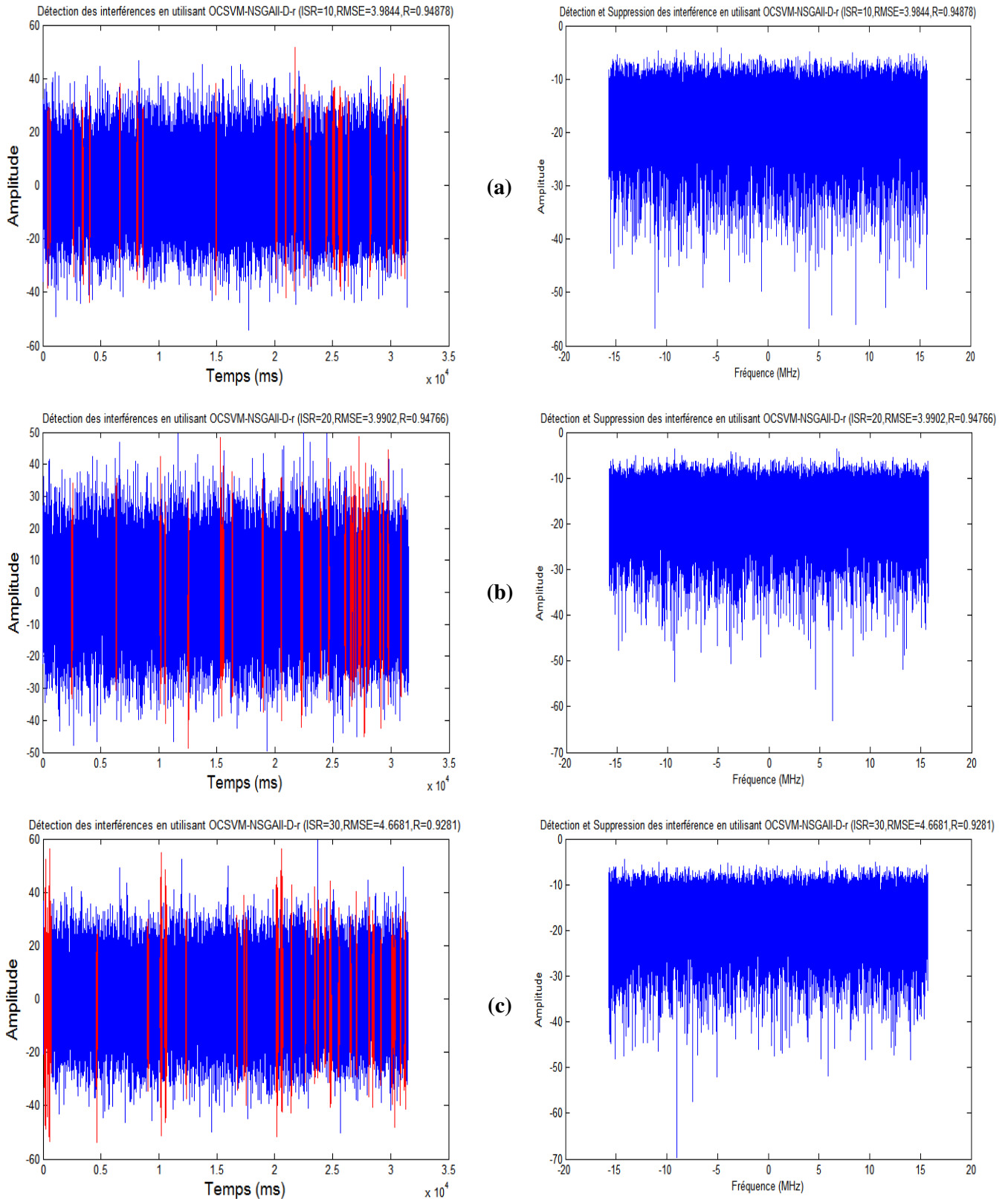


Figure IV.20. Evolution de R en fonction des différentes valeurs ISR.

Dans les simulations précédentes nous avons représentés les résultats obtenus par le classifieur standard OCSVM, les algorithmes génétiques uni-objectif et les algorithmes

génétique multi-objectif, ce qui nous permet de remarquer que la technique d'optimisation NSGAI-D-r est légèrement supérieur en termes RMSE-R en moyenne.

L'application de notre algorithme NSGAI-D-r sur les différentes valeurs d'ISR est montrée dans la figure suivante.



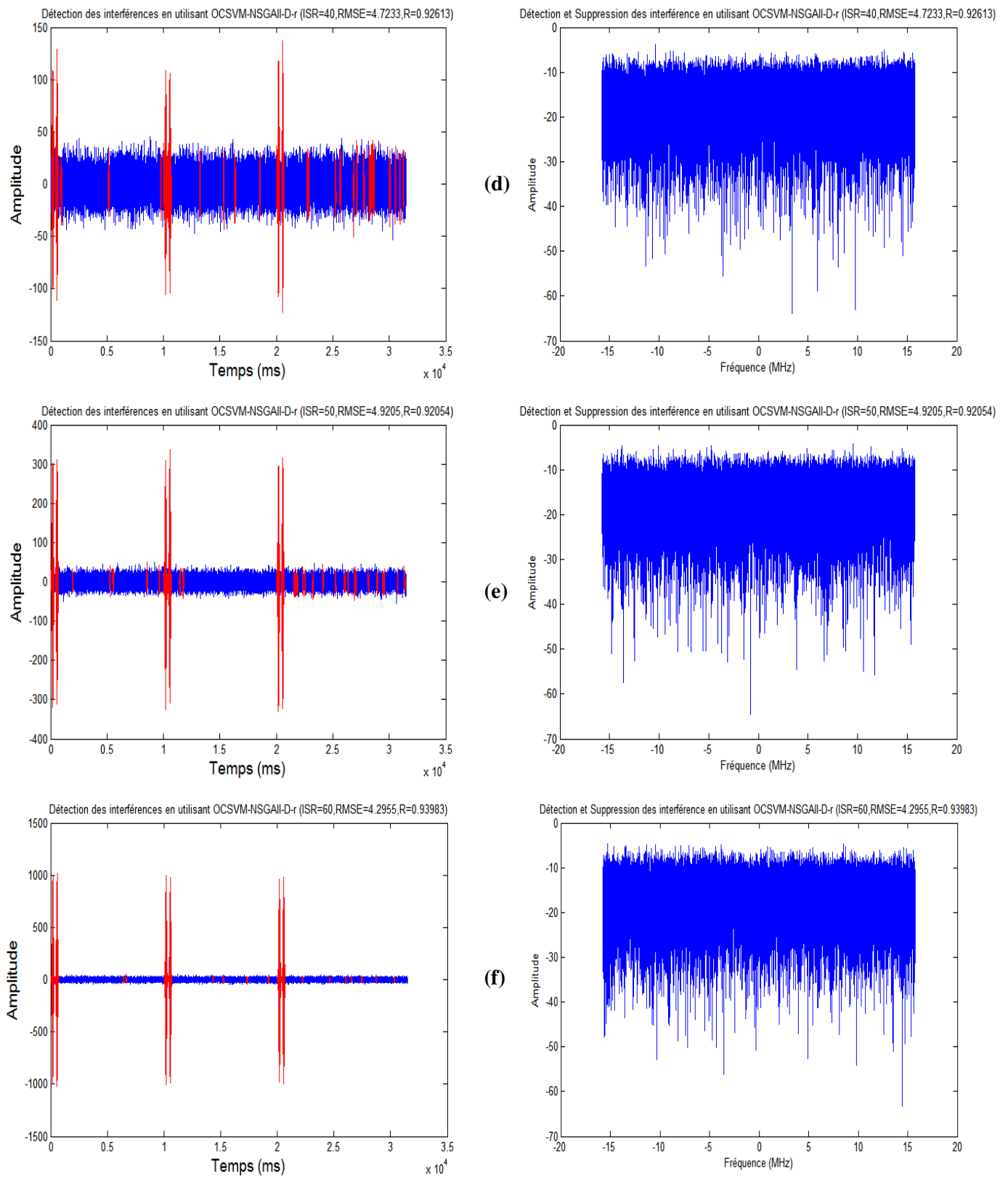


Figure IV.21. Application de l’algorithme NSGAI-D-r pour la détection et la suppression des interférences, (a) $ISR=10dB$, (b) $ISR=20dB$, (c) $ISR=30dB$, (d) $ISR=40dB$, (e) $ISR=50dB$, (f) $ISR=60dB$.

IV. 5. Conclusion

L'objectif principal de ce dernier chapitre, développer un modèle de classification robuste basé sur OCSVM avec l'intermédiaire de la puissance des algorithmes génétiques simple GA ou multi-objectif NSGAI.

A partir de quelques expériences et d'après les critères, qui nous avons utilisés, on a démontré l'efficacité et les performances de notre méthode proposée à notre problème d'identification et de suppression des différents types des outliers (interférences) dans les récepteur GNSS.

Conclusion générale

Conclusion générale

Ces dernières années, plusieurs groupes de recherche et sociétés se sont concentrés sur les contre-mesures d'interférence GNSS, en raison de l'extrême faiblesse des signaux des systèmes GNSS atteignant le récepteur GNSS. Les expériences pratiques, montré que ce récepteur peut facilement être aveuglé et perturbé à cause les interférences intentionnelles et non intentionnelles. Donc notre travail consacré à résoudre le problème de la détection et la suppression des différents types d'interférences aux récepteurs GNSS. Nous avons proposée une méthode basée sur un modèle de classification OCSVM avec les algorithmes d'optimisations uni-objectif (GA) et multi-objectif (NSGAI).

Nous avons vu précédemment dans ce mémoire, les fondamentaux des signaux GNSS et la classification des différents types d'interférences et leurs sources qui peuvent attaquer les récepteurs GNSS. Ensuite les fondamentaux des algorithmes génétiques simples et multi-objectif. Enfin, on a vu les différents méthodes de classifications, on a concentré sur le classifieur OCSVM qui est l'élément le plus important dans l'étape de la détection et la suppression des valeurs aberrantes (interférences).

L'algorithme proposé, contrairement à la plupart des algorithmes existants, optimise simultanément de nombreuses fonctions objectives complémentaires pour une adaptation automatique et donc une meilleure détection des interférences. Il est bien connu que l'élimination des interférences améliore les performances des processus de traitement dans les récepteurs GNSS. Le but de ce travail est donc de développer un nouvel algorithme de classification robuste.

Pour atteindre à notre objectif dans ce travail, nous avons fait plusieurs expériences. Le problème est formulé premièrement avec le classifieur standard OCSVM, ensuite avec l'optimisation par les algorithmes génétiques uni-objectif, et enfin avec l'hybridation d'algorithmes génétiques multi-objectif (NSGAI). Les algorithmes génétiques fonctionnent sur une population de solutions potentielles avec quatre différentes fonctions d'évaluations, pour déterminer automatiquement les paramètres optimaux du classifieur OCSVM.

Cependant, afin d'évaluer notre méthode, nous avons utilisé plusieurs critères qui sont : la distance Bhattacharyya, l'erreur de classification, la racine de l'erreur quadratique moyenne RMSE et le coefficient de corrélation R.

Plusieurs heuristiques on été proposées selon les critères précédents : (i) dans l'espace des caractéristiques, on maximise le critère de la distance Bhattacharyya, pour séparer entre le signal utile (sans interférences) et le signal interférence (outliers), (ii) dans l'espace des

étiquettes (target), on minimise l'erreur de classification, la racine de l'erreur quadratique moyenne et le dernier cas on maximise le critère de similarité R.

Notre approche est très efficace, lorsque on combine le critère de similarité avec un autre concept de distance Bhattacharyya. En combinant les deux avec les méthodes d'optimisation, pour la détection et la suppression des interférences.

Les résultats que nous avons obtenus, et que nous avons présentés dans ce travail, sont assez satisfaisants du point de vue amélioration et sa grande capacité de la localisation des interférences de notre méthode après l'optimisation des coefficients de classifieur OCSVM, par rapport à la méthode de classification standard OCSVM.

Par ailleurs, de nombreuses pistes sont possibles pour améliorer et développer de nouvelles solutions dans le domaine de traitement des signaux GNSS.

- ✚ Utilisation des méthodes de prétraitement des données, comme ICA (Independent Component Analysis), PCA (Principal Component Analysis).
- ✚ Application d'une autre méthode de classification par exemple l'apprentissage profond (Deep Learning).
- ✚ Etude approfondie sur les différents types des outliers dans les signaux GNSS.
- ✚ Utilisation d'autres méthodes robustes aux outliers, comme DONOHO
- ✚ Utilisation d'autre techniques évolutionnaire (Particle Swarm Optimization-PSO, Differential Evolution-DE, Firefly Algorithm-FA, Bat Algorithm).

Bibliographie

Bibliographie

- [1] V. Truong, "DIGUE: Détection d'Interférences Gnss pour Uav autonomE," *Thèse de doctorat, Institut polytechnique de Paris*, 2020.
- [2] Z. A. Khezzar, "Traitement des signaux GNSS," *Thèse de doctorat, Option : Micro-Ondes, Université de Batna 2*, 2021.
- [3] Z. A. Khezzar, R. Benzid, L. Saidi, "New Thresholding Technique in DCT Domain for Interference Mitigation in GNSS Receivers," *Traitement du Signal*, vol. 37, pp. 169-180, 2020.
- [4] D. Ismail and D. Youcef, "Réalisation d'un système de géolocalisation Galileo," *Mémoire de Master, Spécialité : Système de Télécommunications, Université Aboubakr Belkaïd, Tlemcen*, 2019.
- [5] S. Savasta, L. L. Presti, M. Rao, "Interference mitigation in GNSS receivers by a time-frequency approach," *IEEE Transactions on Aerospace and Electronic Systems* vol. 49, pp. 415-438, 2013.
- [6] E. Anyaegbu, G. Brodin, J. Cooper, E. Aguado, S. Boussakta, "An integrated pulsed interference mitigation for GNSS receivers," *The Journal of Navigation*, vol. 61, pp. 239-255, 2008.
- [7] S. Barbarossa and A. Scaglione, "Adaptive time-varying cancellation of wideband interferences in spread-spectrum communications based on time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 47, pp. 957 – 965, 1999.
- [8] G. X. Gao, L. Heng, A. Hornbostel, H. Denks, M. Meurer, T. Walter, P. Enge, "DME/TACAN interference mitigation for GNSS: algorithms and flight test results," *GPS solutions*, vol. 17, pp. 561-573, 2013.
- [9] A. Zinflou, "Système interactif d'aide à la décision basé sur des algorithmes génétiques pour l'optimisation multi-objectifs," *Thèse de doctorat, Université du Québec à Chicoutimi*, 2004.
- [10] C. Chatelain, Y. Oufella, S. Adam, Y. Lecourtier, L. Heutte, "Optimisation multi-objectif pour la sélection de modèles SVM," *Reconnaissance des Formes et Intelligence Artificielle (RFIA'08)*, pp. 67-72, 2008.
- [11] K. Deb, A. Pratap, S. Agarwal, T. A. M. T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE transactions on evolutionary computation*, vol. 6, pp. 182-197, 2002.
- [12] C. Cheng, "GNSS Multipath and Interference Mitigation Using Bayesian Methods," *Thèse de doctorat, Université de Toulouse, ISAE*, 2015.
- [13] A. M. Tobie, "GNSS/5G Hybridization for urban navigation," *Thèse de doctorat, université de Toulouse*, 2021.
- [14] N. Samama, "Applications radars," *Institut Mines-Télécom, édition 4*, pp. 77-78, 2019.
- [15] T. Donal, "Système de positionnement par satellite," *Institut National de l'Information Géographique et Forestière*, vol. 109, pp. 59-84, 2019.
- [16] J.-M. Piéplu, "GPS et Galileo: Systèmes de navigation par satellites," *Editions Eyrolles*, 2011.
- [17] L. Fillatre, "Méthodes de poursuite de phase pour signaux GNSS multifréquence en environnement dégradé," *Thèse de doctorat, Thales Alenia Space*, 2013.
- [18] S. Tay, "Analyse et modélisation de l'utilisation de signaux GNSS en environnement marin," *Thèse de doctorat, Télécom Bretagne, Université de Rennes 1*, 2012.
- [19] M. J. Lesouple, "Estimation parcimonieuse de biais multi trajets pour les systèmes GNSS," *Thèse de doctorat, Université de Toulouse*, 2019.

- [20] S. Roche, "Méthodes de poursuite de phase pour signaux GNSS multifréquence en environnement dégradé," *Thèse de doctorat, Université de Toulouse*, 2019.
- [21] Z. A. Khezzar, "Acquisition Rapide des Signaux GNSS Par un Récepteur Logiciel," *Mémoire de Magister en Electronique, Option : Micro-onde pour Télécommunication, Université de Batna 2*, 2011.
- [22] F. Dovis, "GNSS Interference Threats and Countermeasures," *Artech House*, 2015.
- [23] B. Wang, Y. Sun, Y. Liu, Y. Zhang, S. Li, "Experimental Research on Narrowband Interference Suppression of GNSS Signals," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 3410741, pp. 1-12, 2021.
- [24] K. Yokoyama and Y. Sato, "Using Dominated Solutions to the Uniformity of Non-dominated Solution Distributions in NSGA-II," *Proceedings of the 12th International Joint Conference on Computational Intelligence*, pp. 212-219, 2020.
- [25] F. Douak, N. Ghoggali, R. Hedjam, M. L. Mekhalfi, N. Benoudjit, F. Melgani, "Genetic robust kernel sample selection for chemometric data analysis," *Journal of Chemometrics*, vol. 35, pp. 1-22, 2021.
- [26] A. Layeb, "Introduction aux métaheuristiques," *Cours, Chapitre 5*, pp. 1-15, 2009.
- [27] J. H. Holland, "Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence," *U Michigan Press*, 1975.
- [28] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning," *Addison-Wesley*, 1989.
- [29] I. Chouchani, "Utilisation d'un algorithme génétique pour la composition de services Web," *Mémoire: exigence pratique de la maîtrise en informatique, Université du Québec à Montréal*, 2010.
- [30] N. Zerari, "Les algorithmes génétiques en maintenance," *Thèse de doctorat, Université de Batna 2*, 2006.
- [31] D. Mokeddem, "Contrôle flou des processus biotechnologiques à base d'algorithmes génétiques," *Thèse de doctorat, Université de Jijel*, 2018.
- [32] A. Bouvet, F. Collin, R. Velasco, "Algorithmes génétiques," 2002.
- [33] K. Lamamra, "Optimisation multi-objectifs par les algorithmes génétiques et application à la commande des systèmes," *Thèse de doctorat, Université Mentouri, Constantine*, 2012.
- [34] J.-M. Alliot and N. Durand, "Algorithmes génétiques," *Centre d'Etudes de la Navigation Aérienne*, 2005.
- [35] A. Terki, "Analyse des performances des algorithmes génétiques utilisant différentes techniques d'évolution de la population," *Mémoire de Magister En Electronique, Option : contrôle, Université Mentouri, Constantine*, 2009.
- [36] N. Durand, "Algorithmes Génétiques et autres méthodes d'optimisation appliqués à la gestion de trafic aérien," *Thèse de doctorat, INPT*, 2004.
- [37] J. H. Holland, "Genetic algorithms," *Scientific american*, vol. 267, pp. 66-73, 1992.
- [38] N. Hamdan, "Approche Multi agents pour la reconnaissance du cancer du sein," *Rapport de recherche INRIA*, pp. 1-20, 2011.
- [39] K. S. Tang, K. F. Man, S. Kwong, Q. He, "Genetic algorithms and their applications," *IEEE signal processing magazine*, vol. 13, pp. 22-37, 1996.
- [40] S. Levasseur, "Analyse Inverse en Géotechnique: développement d'une méthode à base d'algorithmes génétiques," *Thèse de doctorat, Université Joseph-Fourier-Grenoble I*, 2007.
- [41] T. Vallée and M. Yildizoğlu, "Présentation des algorithmes génétiques et de leurs applications en économie," *Revue d'économie politique*, vol. 5, pp. 711-745, 2004.

- [42] A. Sioud, "Approches hybrides pour la résolution d'un problème d'ordonnancement industriel," *Thèse de doctorat, Université du Québec à Chicoutimi*, 2011.
- [43] N. Benahmed, "Optimisation de réseaux de neurones pour la reconnaissance de chiffres manuscrits isolés: Sélection et pondération des primitives par algorithmes génétiques," *École de technologie supérieure*, 2002.
- [44] N. B. Mahammed, "Contribution à l'amélioration des processus métiers d'entreprise," *Thèse de doctorat, Université de Djilali Liabes, Sidi Bel Abbes*, 2018.
- [45] J. Dipama, "Optimisation multi-objectif des systèmes énergétiques," *Thèse de doctorat, École Polytechnique de Montréal*, 2010.
- [46] M. Samarra, L. Roseiro, V. Maranha, M. A. Neto, J. Alves, C. Alcobia, "Optimize the position of mechanical components of an external fixator using neural networks and genetic algorithms," *SYMCOMP*, pp. 1-11, 2013.
- [47] I. Boussaid, "Perfectionnement de métaheuristiques pour l'optimisation continue," *Thèse de doctorat, Université de Paris Est*, 2013.
- [48] H. Locteau, R. Raveaux, S. Adam, Y. Lecourtier, P. Héroux, É. Trupin, "Approximation de courbes par Algorithme Génétique Multi-Objectif," *CIFED*, pp. 151-156, 2006
- [49] A. Bensmaïne, "Algorithmes évolutionnaires et méthodes approchées multicritères pour la génération des processus de fabrication dans un environnement reconfigurable," *Thèse de doctorat, Université de Lorraine*, 2013.
- [50] B. Merdjaoui, "Optimisation multi-objectif par algorithmes génétiques et approche pareto des paramètres d'usinage sous contraintes des limitations de production," *Thèse de doctorat, Université de M'hamed Bouguerra, Boumerdes*, 2006.
- [51] C. M. Fonseca and P. J. Fleming, "An overview of evolutionary algorithms in multiobjective optimization," *Evolutionary computation*, vol. 3, pp. 1-16, 1995.
- [52] V. Pareto, "Manuale di Economia Politica," *Societ`a Editrice Libreria, Milan*, 1906.
- [53] V. Pareto, "Manual of Political Economy," *Augustus M. Kelley Publishers, New York*, 1971.
- [54] T. Gräbener and A. Berro, "Optimisation multiobjectif discrète par propagation de contraintes," *JFPC 2008-Quatrièmes Journées Francophones de Programmation par Contraintes*, pp. 403-407, 2008.
- [55] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary computation*, vol. 2, pp. 221-248, 1994.
- [56] O. Debabeche, "La composition de services Cloud par les algorithmes évolutionnaires," *Mémoire de master académique en Informatique, Parcours : Intelligence Artificielle, Université Mohamed Khider, Biskra*, 2020.
- [57] B. Abassi, "Approche de sélection des données biomédicales pour l'identification de pathologies en utilisant les supports vecteurs machines," *Thèse de doctorat. Université Mohamed Boudiaf, M'sila*, 2019.
- [58] J.-M. Commenge, "Big Data et Intelligence Artificielle pour le Génie des Procédés," *Ecole Nationale Supérieure des Industries Chimiques, France. HAL-03107557*, 2020.
- [59] R. Mifdal, "Application des techniques d'apprentissage automatique pour la prédiction de la tendance des titres financiers," *Mémoire de maîtrise électronique, Montréal, École de technologie supérieure*, 2019.
- [60] M. Y. Ammar, "Mise en œuvre de réseaux de neurones pour la modélisation de cinétiques réactionnelles en vue de la transposition batch/continu," *Thèse de doctorat. Institut National Polytechnique, Toulouse, France*, 2007.
- [61] B. S. Belaidi, "Application de la classe des méthodes d'apprentissage statistique SVM (support vector machine) pour la reconnaissance des formes dans les images,"

- Mémoire de Magister En Informatique, Option Intelligence Artificielle et Aide à la Décision (IAAD), Université Abou Bakr Belkaid, Tlemcen, 2011.*
- [62] M. Minsky and S. Papert, "Perceptrons," *M.I.T. Press*, 1969.
- [63] M. L. Minsky and S. A. Papert, "Perceptrons: expanded edition," *Cambridge, MA, United States*, 1988.
- [64] H. Mebarki and O. Ghedir, "Système de communication entre les véhicules (VNET) en utilisant la communication sans fil et les méthodes de l'intelligence artificielle," *Mémoire de Master, Option: Systèmes des Télécommunication, Université Abbes Laghrour, Khenchela*, 2020.
- [65] D. Takouachet and R. Belhabes, "Classification de l'environnement intérieur à base des mesures du signal WiFi," *Mémoire de Master, Option: Systèmes des Télécommunications, Université Abbes Laghrour, Khenchela*, 2021.
- [66] F. Douak, N. Benoudjit, F. Melgani, "A two-stage regression approach for spectroscopic quantitative analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 109, pp. 34-41, 2011.
- [67] B. Wierenga and J. Kluytmans, "Neural nets versus marketing models in time series analysis: A simulation study," *Marketing: Its Dynamics and Challenges*, vol. 1139, pp. 1139-1153, 1994.
- [68] N. Benoudjit, C. Archambeau, A. Lendasse, J. A. Lee, M. Verleysen, "Width optimization of the Gaussian kernels in Radial Basis Function Networks," *ESANN*, vol. 2, pp. 425-432, 2002.
- [69] K. Ferroudji, "Classification des particules par ultrason et réseaux de neurones artificiels (Application aux emboles)," *Mémoire de Magister, Université de Batna 2*, 2011.
- [70] V. Vapnik, "The nature of statistical learning theory," *Springer science & business media*, 1999.
- [71] S. Benyahia, "Application de la classe des méthodes d'apprentissage statistique SVM (support vector machine) pour la reconnaissance des formes dans les images," *Thèse de doctorat, Université de Tlemcen*, 2011.
- [72] O. Bousquet, "Introduction au Support Vector Machines (SVM) " *Centre de mathématiques appliquées, Ecole polytechnique de Palaiseau*, 2001.
- [73] S. Ndiaye, F. Camara, M. L. Samb, M. A. Essegir, Y. Slimani, "Approche de sélection d'attributs pour la classification basée sur l'algorithme RFE-SVM," *Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées*, vol. 17, pp. 197-219, 2014.
- [74] L. M. Manevitz and M. Yousef, "One-class SVMs for document classification," *Journal of machine Learning research*, vol. 2, pp. 139-154, 2001.
- [75] K. L. Li, H. K. Huang, S. F. Tian, W. Xu, "Improving one-class SVM for anomaly detection," *IEEE, Proceedings of the 2003 international conference on machine learning and cybernetics*, vol. 5, pp. 3077-3081, 2003.
- [76] X. He, G. Mourrot, D. Maquin, J. Ragot, P. Beausery, A. Smolarz, "One-class SVM in multi-task learning," *Annual Conference of the European Safety and Reliability Association, ESREL Troyes, France. pp.CDROM. HAL-00593397*, 2012.
- [77] S. Lecomte, C. Richard, R. Lengellé, F. Capman, B. Ravera, "Machines à vecteur de support 1-classe (OC-SVM) pour la détection non supervisée d'événements sonores anormaux," *GRETSI 2011*.
- [78] D. Alonso de Diego, N. G. Ferrara, J. Nurmi, E. S. Lohan, G. Hein, "Interference mitigation in the E5a Galileo band using an open-source simulator," *Inside GNSS*, pp. 55-63, 2016.