

République Algérienne Démocratique et Populaire  
Ministère de l'enseignement supérieur et de la recherche scientifique  
Université Mouloud Mammeri de Tizi-Ouzou  
Faculté de Génie Electrique et d'Informatique  
Département d'Informatique



## THÈSE DE DOCTORAT LMD

Discipline : Informatique

Option : Réseaux, Mobilité et Systèmes Embarqués

Présentée par

**Hassina AIT ISSAD**

### **Déploiement intelligent de drones pour une agriculture du futur**

#### **Jury**

M. Mustapha LALAM,	Professeur, UMMTO	Président
Mme Rachida AOUDJIT-LAGAB,	Professeur, UMMTO	Directrice de thèse
M. Hachem SLIMANI,	Professeur, Univ. Béjaïa	Examineur
M. Mehammed DAOUI,	Professeur, UMMTO	Examineur
M. Saïd TALBI,	MCA, UMMTO	Examineur

Soutenue le 11/10/2020

# Remerciements

**A**vant tout, j'adresse mes remerciements à Allah qui m'a donné la patience et le courage pour compléter ce travail.

Ce travail a été réalisé au sein du laboratoire de recherche en informatique (LARI) de l'Université Mouloud Mammeri de Tizi-Ouzou. Je tiens à remercier son directeur Monsieur Mustapha Lalam pour ses conseils ainsi que les moyens mis à ma disposition pour accomplir ce travail. Je voudrais lui exprimer ma gratitude pour l'honneur qu'il me fait en acceptant de présider le jury de ma thèse.

Je tiens à remercier vivement ma directrice de thèse Mme AOUDJIT Rachida, pour sa disponibilité, son encouragement, ses idées, ses conseils, son aide et sa gentillesse qui m'ont permis de mener à bien cette thèse.

Je remercie également Mr. SLIMANI Hachem, professeur à la Faculté de Bejaïa, d'avoir accepté de juger mon travail de thèse.

Mes remerciements vont également à Mr DAOUI Mohammed, professeur à la Faculté de Génie Electrique et d'Informatique de Tizi-Ouzou et Mr TALBI Saïd, maitre de conférences à la Faculté de Génie Electrique et d'Informatique de Tizi-Ouzou, de l'honneur qu'ils me font en acceptant d'être examinateurs de ma thèse.

J'exprime aussi ma plus profonde gratitude à ma famille qui m'a toujours soutenue et encouragée tout au long de mes études.

J'associe à tous ces remerciements tous les membres du laboratoire LARI ainsi que les enseignants du département d'informatique.

Enfin, Je tiens à remercier tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

# Résumé

Depuis que les êtres humains sont passés à un mode de vie agraire, les progrès technologiques ont permis des évolutions dans l'agriculture, résultant en de plus grandes variétés et rendements des cultures. Après les révolutions de la sélection végétale et de la génétique, une troisième révolution verte prend le contrôle du monde agricole basé sur l'application des technologies de l'information et de la communication (TIC) telles que les capteurs et actionneurs, l'Internet des objets (IoT : Internet of Things), l'intelligence artificielle, les véhicules aériens sans pilote (UAVs, drones). Ces nouvelles technologies ont permis une agriculture basée sur les données : l'agriculture intelligente. L'un des derniers développements est l'utilisation accrue de petits véhicules aériens sans pilote, communément appelés drones, pour l'agriculture. Ceux-ci ont un énorme potentiel dans l'agriculture pour fournir des données précieuses qui peuvent ensuite être utilisées et exploitées par des techniques de Data Mining pour influencer les politiques et les décisions agricoles.

Les travaux de recherche présentés dans cette thèse portent sur l'adoption des nouvelles technologies dans l'agriculture en examinant les progrès de ces technologies clés dans le contexte de l'agriculture intelligente. Nos objectifs à travers le travail rapporté dans ce manuscrit peuvent être résumés en deux aspects. Le premier objectif est de montrer que les techniques de Data Mining et l'accès à des informations en temps opportun pour mener à bien des opérations de planification et de décisions cruciales sont la clé du succès d'une agriculture intelligente. Pour y parvenir, une discussion détaillée des changements apportés par les technologies de traitement de données et de prise de décisions dans le secteur agricole est présentée. Cette partie est une opportunité d'explorer certaines solutions et applications de Data Mining et de comprendre comment ce concept peut être introduit dans des systèmes agricoles durables. Le deuxième objectif est la conception d'une solution de déploiement optimal d'une flotte de drones pour assurer la surveillance d'une zone agricole. En effet, la qualité des services fournis par une flotte de drones dépend de leurs emplacements et leur distribution sur le terrain. Pour mesurer les performances de toute solution, une fonction de déploiement peut être définie en fonction de plusieurs paramètres souvent contradictoires. Une modélisation du problème de déploiement en un problème d'optimisation multi-objectif avec contraintes est présentée, suivie de sa résolution en se reposant sur des méta-heuristiques puissantes pour trouver la stratégie de déploiement optimale avec quatre objectifs. L'approche proposée permet de maximiser la couverture totale, de minimiser la consommation énergétique, de minimiser les chevauchements de couverture et de garantir la connectivité entre les drones.

**Mots clés :** Agriculture intelligente, Agriculture de précision, Déploiement de drones, UAVs, Data Mining, Optimisation multi-objectif.

# Abstract

Since humans moved to an agrarian way of life, technological advances have allowed evolutions in agriculture, resulting in greater varieties and yields of crops. After the revolutions in plant breeding and genetics, a third green revolution takes control of the agricultural sector based on the application of information and communication technologies (ICT) such as sensors and actuators, Internet of Things (IoT), artificial intelligence, unmanned aerial vehicles (UAVs, drones). These new technologies have enabled data-driven agriculture : smart agriculture.

One of the latest developments is the increased use of small unmanned aerial vehicles, commonly known as drones, for agriculture. These have enormous potential in agriculture to provide valuable data which can then be used by Data Mining techniques to influence agricultural policies and decisions.

The research presented in this thesis focuses on the adoption of new technologies in agriculture by examining the progress of these key technologies in the context of smart agriculture. Our objectives through the work reported in this manuscript can be summarized in two aspects. The first objective is to show that Data Mining techniques and access to timely information to carry out crucial planning and decision-making are the key to the success of smart agriculture. To achieve this, a detailed discussion of the changes brought about by data processing and decision-making technologies in the agricultural sector is presented. This part is an opportunity to explore some Data Mining solutions and applications and understand how this concept can be introduced into sustainable agricultural systems. The second objective is the design of an optimal deployment solution for a fleet of drones to monitor an agricultural area. Indeed, the quality of services provided by a fleet of drones depends on their locations and their distribution in the field. To measure the performance of any solution, a deployment function can be defined according to several parameters that are often contradictory. A modeling of the deployment problem into a constrained many-objective optimization problem is presented, followed by its resolution based on powerful metaheuristics to find the optimal deployment strategy with four objectives. The proposed approach maximizes total coverage, minimizes energy consumption, minimizes overlap of coverage and ensures connectivity between drones.

**Keywords :** Smart agriculture, Precision agriculture, Drones deployment, UAVs, Data Mining, Many-objective optimization.

# Table des matières

<b>Table des matières</b>	<b>vi</b>
<b>Table des figures</b>	<b>viii</b>
<b>Liste des tableaux</b>	<b>ix</b>
<b>Introduction générale</b>	<b>1</b>
Contexte et objectifs . . . . .	1
Contributions . . . . .	4
Organisation de la thèse . . . . .	5
<b>I État de l’art</b>	<b>7</b>
<b>1 Les nouvelles technologies au service de l’agriculture intelligente</b>	<b>8</b>
1.1 Introduction . . . . .	8
1.2 L’agriculture intelligente . . . . .	8
1.2.1 Définition de l’agriculture intelligente . . . . .	9
1.2.2 Les technologies de l’agriculture intelligente . . . . .	11
1.2.2.1 Les technologies d’automatisation et d’acquisition . . . . .	12
1.2.2.2 Les technologies de transmission de données . . . . .	17
1.2.2.3 Les technologies de traitement de données . . . . .	18
1.2.2.4 Les réseaux de capteurs sans fils . . . . .	18
1.2.2.5 L’internet des objets . . . . .	19
1.2.2.6 La combinaison de trois technologies : IoT, UAV, RCSF . . . . .	22
1.3 Les drones dans l’agriculture intelligente . . . . .	24
1.3.1 Classification des drones . . . . .	24
1.3.2 Différentes applications de drones dans l’agriculture intelligente . . . . .	25
1.3.2.1 Les drones pour la gestion des maladies et parasites des cultures	25
1.3.2.2 Les drones pour la surveillance et la création de cartes de rendement	26
1.3.2.3 Les drones pour la fertilisation et l’irrigation . . . . .	26
1.3.2.4 Les drones pour la reconnaissance et la classification . . . . .	27
1.4 Conclusion . . . . .	27

<b>2</b>	<b>Data Mining dans l’agriculture intelligente</b>	<b>28</b>
2.1	Introduction . . . . .	28
2.2	Les différentes tâches du Data Mining . . . . .	28
2.2.1	La classification . . . . .	29
2.2.2	Le clustering . . . . .	31
2.2.3	La recherche d’associations . . . . .	31
2.2.4	La prédiction . . . . .	32
2.3	Data Mining et l’agriculture intelligente . . . . .	32
2.3.1	Contrôle de l’irrigation . . . . .	32
2.3.2	Surveillance des maladies des plantes (Détection et prédiction des maladies)	35
2.3.3	Surveillance des organismes nuisibles et gestion optimisée des intrants (en-	
	grais et pesticides) . . . . .	41
2.3.4	Prédiction du rendement des cultures et des effets du changement clima-	
	tique sur la productivité . . . . .	45
2.4	Discussion . . . . .	49
2.5	Défis de Data Mining dans l’agriculture intelligente . . . . .	51
2.5.1	Confidentialité . . . . .	51
2.5.2	Qualité et précision des données . . . . .	52
2.5.3	Importance des données spatiales . . . . .	52
2.5.4	Intégration des connaissances du domaine agricole dans l’exploration de	
	données . . . . .	53
2.5.5	Évolutivité des algorithmes de Data Mining . . . . .	53
2.6	Conclusion . . . . .	53
<b>3</b>	<b>Techniques de déploiement dans les réseaux (RCSF, UAVs)</b>	<b>55</b>
3.1	Introduction . . . . .	55
3.2	Préliminaire / Concepts de Base . . . . .	55
3.2.1	Réseaux de capteurs et réseaux de drones . . . . .	55
3.2.2	Les modèles de couverture . . . . .	56
3.2.3	Les modèles de communication . . . . .	57
3.3	Objectifs et contraintes d’une stratégie de déploiement . . . . .	57
3.3.1	Couverture . . . . .	58
3.3.2	Connectivité . . . . .	59
3.3.3	Efficacité énergétique . . . . .	59
3.3.4	La durée de vie du réseau . . . . .	60
3.3.5	Le coût de déploiement et le nombre de nœuds . . . . .	60
3.3.6	Fiabilité des données . . . . .	60
3.3.7	Tolérance aux pannes et équilibrage énergétique . . . . .	60
3.4	Les stratégies de déploiement des RCSF . . . . .	61
3.4.1	Déploiement basé sur des techniques de déploiement classiques . . . . .	62
3.4.1.1	Techniques basées sur la force . . . . .	62
3.4.1.2	Techniques basées sur la grille . . . . .	65
3.4.1.3	Techniques basées sur la géométrie de calcul . . . . .	66
3.4.2	Techniques basées sur des méta-heuristiques . . . . .	68
3.4.3	Techniques basées sur l’auto-programmation . . . . .	71

3.5	Le déploiement des réseaux de drones . . . . .	71
3.5.1	Déploiement de plusieurs drones et planification de la trajectoire d'un seul drone . . . . .	72
3.5.2	Les stratégies de déploiement dans les drones . . . . .	73
3.6	Conclusion . . . . .	77
<b>II</b>	<b>Contribution</b>	<b>79</b>
<b>4</b>	<b>Une approche de déploiement d'une flotte de drones basée sur des méta-heuristiques</b>	<b>80</b>
4.1	Introduction . . . . .	80
4.2	Concepts de base . . . . .	81
4.2.1	Problème d'optimisation multi-objectif . . . . .	81
4.2.2	Concept de Pareto . . . . .	82
4.3	Définition du problème de déploiement . . . . .	83
4.4	Description et formulation mathématique . . . . .	84
4.4.1	Couverture . . . . .	84
4.4.2	Consommation d'énergie . . . . .	85
4.4.3	Connectivité . . . . .	86
4.4.4	Chevauchement . . . . .	87
4.5	Les méta-heuristiques utilisées . . . . .	88
4.5.1	MOABC . . . . .	89
4.5.2	MOPSO . . . . .	91
4.5.3	SPEA II . . . . .	93
4.5.4	NSGA II . . . . .	95
4.5.5	NSGA III . . . . .	97
4.6	Adaptation des méta-heuristiques . . . . .	99
4.6.1	Représentation d'une solution . . . . .	99
4.6.2	Initialisation de la population . . . . .	100
4.6.3	Les opérateurs de reproduction . . . . .	100
4.6.3.1	Le croisement . . . . .	100
4.6.3.2	La mutation . . . . .	100
4.7	Conclusion . . . . .	101
<b>5</b>	<b>Évaluation des performances de l'approche proposée</b>	<b>102</b>
5.1	Introduction . . . . .	102
5.2	Description des scénarios . . . . .	102
5.3	Paramètres de simulation . . . . .	102
5.4	Les métriques d'évaluation . . . . .	103
5.5	Résultats des simulations . . . . .	104
5.5.1	Partie 1 . . . . .	104
5.5.2	Partie 2 . . . . .	106
5.6	Conclusion . . . . .	112
	<b>Conclusion</b>	<b>113</b>

<b>Publications</b>	<b>116</b>
<b>Bibliographie</b>	<b>117</b>

# Table des figures

1.1	Diagramme conceptuel de l’agriculture intelligente basée sur les réseaux [239] . . .	10
1.2	La position de l’agriculture intelligente [187] . . . . .	10
1.3	Les services fournis et les différentes technologies intégrées dans l’agriculture intelligente . . . . .	11
1.4	Exemples de capteurs utilisés . . . . .	14
1.5	Quantité estimée de données générées par une ferme moyenne par jour[146] . . . .	21
1.6	Les tendances de l’agriculture intelligente [75] . . . . .	22
1.7	Concept d’un système intégré de réseau de capteurs sans fil - véhicule aérien sans pilote (RCSF – UAV)[176] . . . . .	23
1.8	Convergence UAV, RCSF et IoT dans l’agriculture intelligente [176] . . . . .	23
1.9	Types de drones agricoles . . . . .	25
2.1	Les différentes tâches du Data Mining . . . . .	29
3.1	Les modèles de couverture . . . . .	57
3.2	Les différents types de couverture [178] . . . . .	58
3.3	Exemple d’un réseau représenté par un graphe connecté et non connecté . . . . .	59
3.4	La classification des différentes techniques de déploiement [76] . . . . .	61
3.5	Un exemple de trois forces différentes sur un nœud [154] . . . . .	62
3.6	Les modèles les plus utilisés [154] . . . . .	65
3.7	Exemple d’un diagramme de Voronoi et d’une triangulation de Delaunay [76] . . . .	66
4.1	Le concept de Pareto pour un problème d’optimisation à deux objectifs [127] . . . .	82
4.2	Un exemple de déploiement de drones dans une zone agricole . . . . .	84
4.3	Un schéma décrivant la procédure NSGA II [61] . . . . .	96
4.4	Étape de croisement . . . . .	100
4.5	L’opérateur de mutation à un point . . . . .	101
5.1	Valeurs des objectifs lors de l’optimisation de tous les objectifs ensemble . . . . .	105
5.2	Evolution de la moyenne des valeurs IGD en fonction du nombre de générations pour 8 drones . . . . .	106
5.3	Evolution de la moyenne des valeurs HV en fonction du nombre de générations pour 8 drones . . . . .	107
5.4	Evolution de la moyenne des valeurs IGD en fonction du nombre de générations pour 12 drones . . . . .	107

5.5	Evolution de la moyenne des valeurs HV en fonction du nombre de générations pour 12 drones . . . . .	107
5.6	Evolution de la moyenne des valeurs IGD en fonction du nombre de générations pour 15 drones . . . . .	108
5.7	Evolution de la moyenne des valeurs HV en fonction du nombre de générations pour 15 drones . . . . .	108
5.8	Coordonnées parallèles des solutions obtenues par l'algorithme NSGA II (IGD= $2.03e^{-3}$ ) . . . . .	109
5.9	Coordonnées parallèles des solutions obtenues par l'algorithme SPEA II (IGD= $1.8e^{-3}$ ) . . . . .	109
5.10	Coordonnées parallèles des solutions obtenues par l'algorithme NSGA III (IGD= $1.9e^{-3}$ ) . . . . .	110
5.11	Coordonnées parallèles des solutions obtenues par l'algorithme MOPSO (IGD= $4.3e^{-3}$ ) . . . . .	110
5.12	Coordonnées parallèles des solutions obtenues par l'algorithme MOABC (IGD= $3.9e^{-3}$ ) . . . . .	111

# Liste des tableaux

2.1	Un résumé des articles traitant de la tâche d'irrigation . . . . .	34
2.2	Résumé de certains articles traitant de la détection et de la classification des maladies à l'aide des méthodes de Data Mining . . . . .	36
2.3	Un résumé d'articles traitant de la détection et de la classification des maladies à l'aide de l'apprentissage en profondeur . . . . .	40
2.4	Un résumé de certains articles traitant de la surveillance des parasites et de la gestion des intrants . . . . .	44
2.5	Résumé de certains articles traitant de la prédiction du rendement des cultures et des effets du changement climatique sur la productivité . . . . .	48
3.1	Résumé des objectifs atteints et les limites des articles cités précédemment avec les différentes stratégies utilisées basées sur la force . . . . .	64
3.2	Résumé des objectifs atteints et les limites des articles cités précédemment avec les différentes stratégies utilisées basées sur la grille et sur la géométrie de calcul . . . . .	67
3.3	Résumé des objectifs atteints et les limites des articles cités précédemment avec les différentes stratégies utilisées . . . . .	70
3.4	Comparaison entre le RCSF et le réseau de drones coopératifs . . . . .	72
3.5	Un résumé des stratégies de déploiement des drones . . . . .	76
4.1	Le tableau de la solution . . . . .	99
5.1	Configuration des paramètres des algorithmes . . . . .	103

# Introduction générale

## Contexte et objectifs

Depuis toujours, l'humanité a été menacée par la combinaison de la croissance démographique et de la diminution des ressources. Cette menace continue se traduit par une diminution de la sécurité alimentaire au fil du temps. Comme le montre l'histoire, cette menace a été toujours repoussée par des solutions innovantes développées par l'être humain.

Grâce à ces solutions innovantes, la production agricole augmente depuis des siècles. Une première «révolution verte» a commencé en 1870 en Europe, au cours de laquelle la production et la productivité ont augmenté [223]. L'utilisation d'engrais, de nouvelles semences et de nouvelles machines a augmenté les rendements tout en nécessitant moins de travail. Dans les années 1960, l'introduction de nouvelles variétés de cultures, un apport accru d'engrais, d'eau, de pesticides et d'autres technologies ont entraîné une deuxième révolution verte qui a considérablement augmenté la production [74].

En raison d'une population mondiale croissante qui devrait atteindre 9,6 milliards en 2050, l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO : Food and Agriculture Organization of the United Nations) a déclaré que pour nourrir les habitants du monde d'ici 2050, la production alimentaire devra augmenter d'environ 70% [135]. De plus, selon l'évaluation du climat de 2018 par le programme américain de recherche sur le changement climatique [221], le changement climatique présente de nombreux défis pour maintenir et améliorer la productivité des cultures, la santé du bétail et la vitalité économique mondiale.

Bien que certaines régions puissent voir des conditions propices à une productivité accrue ou alternative des cultures au cours des prochaines décennies, dans l'ensemble, les rendements des principales cultures mondiales devraient diminuer en raison de l'augmentation des températures et peut être des changements dans la disponibilité de l'eau, l'érosion des sols et les maladies et épidémies de ravageurs [87]. Une troisième révolution verte est alors nécessaire pour augmenter la production agricole : **l'agriculture intelligente**.

Cette nouvelle révolution verte impliquera l'utilisation des technologies de l'information et de la communication (TIC) qui transforment les capacités de production agricole. Cette nouvelle ère permettrait une productivité considérablement plus élevée tout en faisant face à une pénurie importante de ressources pour atteindre une production agricole durable.

L'agriculture intelligente est perçue comme un des moyens d'atteindre ces objectifs. Les systèmes agricoles intelligents et de précision devraient jouer un rôle important dans l'amélioration des activités agricoles. L'optimisation des ressources pour une production durable à coûts maîtrisés est l'un de ses principes de base. Elle est basée sur l'utilisation de différentes technologies d'automatisation, de capture de données, de transmission et de traitement de données et de prise de décision. Ces technologies incluent des capteurs de différentes sortes qui peuvent être déployés dans l'air et au sol, des systèmes de navigation géographiques pour la localisation, des systèmes d'information géographique, des drones, des smartphones, des technologies sans fil, l'informatique en nuage, des technologies de l'IoT (Internet of Things) et d'autres technologies [49, 58].

Les systèmes d'agriculture intelligente et de précision devraient jouer un rôle important dans l'amélioration des activités agricoles. C'est la convergence de l'IoT et des technologies de l'information. Son objectif est de collecter des données provenant de sources hétérogènes pour comprendre, prévoir et mieux organiser l'activité agricole. Grâce à l'IoT, des capteurs sont déployés partout pour collecter différents types de données (température du sol, humidité du sol, humidité et verdure des feuilles, rayonnement solaire, direction du vent, niveau des précipitations) en temps réel.

Certes, l'agriculture intelligente envisage de tirer profit des technologies de l'information et de la communication pour permettre aux entreprises agricoles d'être plus productives, plus efficaces et plus rentables, cependant de telles technologies ne suffisent pas à elles seules ; elles doivent être combinées judicieusement avec des outils puissants d'exploration de données pour fournir des informations pertinentes, utiles et fiables.

En effet, les données ramenées de l'environnement de surveillance et collectées par divers équipements (drones, capteurs, etc.) jouent un rôle important. Les informations météorologiques et les données sur l'état des sols peuvent mettre en garde les agriculteurs contre les maladies infectieuses. Les agriculteurs peuvent également utiliser ces données collectées pour la protection des cultures et la fertilisation personnalisée, ce qui se traduit par un rendement plus élevé avec un impact environnemental réduit.

Les données collectées (météorologiques, agronomiques, risque de bio-agresseurs, économiques, etc.) permettent une agriculture de précision, qui vise à optimiser les rendements en apportant un traitement adapté à chaque zone. Ils permettent également de prévenir les risques sanitaires et climatiques, d'avoir une idée précise des besoins de chaque parcelle agricole, de gérer et d'optimiser les ressources (eau, nutriments et pesticides) et d'ajuster les interventions de l'agriculteur.

Cependant, le traitement manuel de ces énormes données est difficile voire impossible pour les humains, il est donc essentiel d'automatiser cette tâche en utilisant des méthodes et des outils d'analyse pour transformer ces données en connaissances qui serviront au processus décisionnel. En effet, pour la prise d'innombrables décisions quotidiennes, des techniques d'analyse solides capables de traiter et d'analyser de grandes quantités de données pour obtenir des informations et des connaissances plus fiables et des prédictions beaucoup plus précises sont essentielles.

Le travail effectué dans cette thèse a commencé avec un objectif principal qui était d'étudier

l'intégration des nouvelles technologies dans des environnements agricoles. Pour y parvenir, nous avons besoin d'étudier une combinaison d'outils, de modèles et de moyens qui ont permis l'émergence et la révolution de l'agriculture intelligente.

D'une manière générale, nous avons étudié différentes tâches et activités agricoles telles que la prédiction des maladies des cultures, la détection des ravageurs, la gestion des intrants ( planification de l'irrigation et des pesticides), la prédiction des rendements, etc. Nous avons analysé plus particulièrement comment ces tâches ont tiré profit des nouvelles technologies dans le cadre d'une agriculture intelligente. Nous avons constaté que les données collectées de l'environnement sont considérées comme une nouvelle source d'informations utiles pouvant affecter directement l'efficacité des activités agricoles, offrant des résultats de meilleure qualité. En conséquence, l'exploration de données (Data Mining) a suscité l'intérêt de beaucoup de chercheurs en raison de sa forte capacité à extraire des informations significatives et robustes. Dans ce contexte, nous avons réalisé une étude complète et détaillée sur les techniques de Data Mining (DM) dans l'agriculture intelligente [12].

D'autre part, L'utilisation de véhicules aériens sans pilote (UAVs : Unmanned Aerial Vehiculs) dans le secteur agricole a suscité beaucoup d'intérêt ces dernières années. Ils sont devenus une technologie importante jouant un rôle clé dans différentes applications dans différents domaines, notamment la surveillance de l'environnement, le secteur médical, en particulier dans certaines situations d'urgence, la gestion des catastrophes naturelles, la surveillance des incendies de forêt, la surveillance du trafic routier en milieu urbain, etc. Ils semblent être précieux pour l'avenir de l'agriculture.

Les drones placés d'une manière optimale peuvent jouer un rôle d'un réseau pour collecter des informations à partir des capteurs déployés sur terre, ils les envoient au centre de contrôle (Cloud par exemple) pour l'analyse et reçoivent des commandes à communiquer à une infrastructure terrestre (un réseau de capteurs). Ils peuvent même servir à exécuter des tâches comme la pulvérisation de la surface agricole avec des quantités précises des produits phytosanitaires pour éviter la dégradation de la santé des cultures et les protéger contre tout ravageur.

Malgré que les drones offrent beaucoup d'avantages, la qualité des services qu'ils fournissent dépend de leurs emplacements. Ainsi, leurs performances sont grandement influencées par les stratégies de placement car elles peuvent affecter directement les mesures de la qualité de la couverture, la consommation d'énergie, la durée de vie, etc. En conséquence, le choix d'une stratégie de déploiement n'est pas une tâche aisée et c'est l'un des problèmes les plus fondamentaux des réseaux de drones.

Le déploiement optimal des drones garantissant que les paramètres importants tels que la connectivité, la consommation d'énergie et le chevauchement sont pris en considération tout en garantissant qu'une zone est entièrement couverte par tous les drones fait partie des problèmes d'optimisation difficiles (NP\_difficile) [207].

De nombreuses études [258, 179, 257] relatives au déploiement de drones dédiées particulièrement au suivi des cibles et portent sur l'optimisation d'un à deux critères.

Dans cette thèse, nous nous intéressons à planifier de manière optimale une mission agricole, couvrant une zone donnée avec un ensemble de drones, en utilisant une combinaison de paramètres de performance, tels que l'énergie, la couverture, la connectivité réseau et le chevauchement pour la prise de décision de déploiement, contrairement à la littérature où seuls l'énergie et le nombre de drones sont généralement étudiés mais très peu de travaux ont considéré le chevauchement et la connectivité.

L'énergie et le nombre de drones sont des métriques très importantes, en particulier l'énergie qui est une ressource cruciale pour tous les appareils mobiles, en particulier les drones. Puisque le système est autonome, l'énergie devient une préoccupation majeure et il devient alors vital d'assurer le succès de la mission. Il est également nécessaire de minimiser le nombre de drones afin de réduire le coût d'investissement.

Le maintien de la connectivité d'un réseau est une exigence essentielle pour certaines applications. La connectivité a un impact important sur le succès de la mission et la coordination des drones [242]. Cependant, peu d'articles l'ont considérée [43, 189]. La couverture est également une mesure importante qui reflète le degré de surveillance de la zone d'intérêt.

Le chevauchement est une autre contrainte à prendre en compte. Les UAVs doivent être placés de manière à maximiser la couverture totale tout en évitant le chevauchement dans leurs zones de couverture. Avec le déploiement de drones, les chevauchements de couverture, où la même surface de la zone de surveillance peut être couverte par plusieurs drones en même temps, peuvent être très gênants. En effet, la qualité de service dans un système agricole est critique et le chevauchement des zones entre drones peut augmenter la consommation d'énergie, provoquer une toxicité et une pollution du milieu agricole, et générer un coût supplémentaire relatif à l'application des surdoses de produits phytosanitaires lors d'une mission de pulvérisation par exemple.

## **Contributions**

Les principales contributions apportées par cette thèse sont :

- Un aperçu des techniques de Data Mining qui sont une approche essentielle pour la réalisation de solutions efficaces et pratiques aux problèmes agricoles. Il souligne les caractéristiques particulières des données agricoles, ce qui rend très important leur prise en compte dans l'analyse des données en agriculture intelligente. Il a mis en évidence des techniques prometteuses utilisées pour découvrir et comprendre les relations entre les différents facteurs environnementaux et biologiques et leur influence sur la production végétale saine et l'estimation du rendement pour prendre d'innombrables décisions chaque jour. Il mentionne également quelques orientations pour des travaux futurs dans ce domaine.
- Une solution pour le problème de déploiement optimal d'un ensemble de drones maximisant la couverture totale tout en minimisant les chevauchements et la consommation d'énergie et

en maintenant la connectivité dans un environnement agricole.

- Contrairement aux stratégies présentées dans la littérature qui se concentrent généralement sur une ou deux métriques de performance (nombre de drones et l'énergie) et prêtent peu d'attention au problème de chevauchement et de connectivité, cette thèse propose d'utiliser une combinaison de quatre métriques de performance pour résoudre le problème. Les UAVs doivent être positionnés de manière à maximiser la couverture totale et à éviter les chevauchements dans leurs zones de couverture. En conséquence, les effets indésirables d'une surdose de pesticides seront évités.
- Notre approche tente de couvrir l'ensemble d'une zone d'intérêt contrairement aux techniques qui se concentrent uniquement sur la couverture des cibles représentées par des points.
- Enfin, la thèse donne une comparaison des algorithmes d'optimisation méta-heuristiques, à savoir Non-Dominated Sorting Genetic Algorithm II(NSGA II), Non-Dominated Sorting Genetic Algorithm III (NSGA III), Strength Pareto Evolutionary Algorithm II (SPEA II), Multi-Objective Particle Swarm Optimization (MOPSO) et Multi-Objective Artificial Bee Colony (MOABC) pour résoudre le problème de couverture à plusieurs objectifs afin de déterminer le meilleur emplacement 3D d'une flotte de drones, contrairement aux approches proposées dans la littérature qui explorent peu de méta-heuristiques et peu de fonctions de coût sont considérées.

## Organisation de la thèse

Cette thèse est organisée en deux parties. La première partie, composée de trois chapitres, est dédiée à la présentation de l'agriculture intelligente (chapitre 1), du Data Mining dans le contexte de l'agriculture intelligente (chapitre 2) et enfin des techniques de déploiement dans les réseaux de capteurs et de drones (chapitre 3). La deuxième partie, présente notre contribution. Elle est divisée en deux chapitres 4 et 5, dédiés respectivement à la présentation de notre solution et à l'exposition des différents résultats d'évaluation de l'approche proposée.

Le détail de cette organisation est donné comme suit :

Dans le chapitre 1, nous présenterons les différentes innovations susceptibles d'apporter des réponses aux problèmes agricoles ainsi que les aspects liés à l'agriculture intelligente. Nous nous attacherons en particulier à présenter les différentes technologies adoptées dans l'agriculture intelligente en mettant en évidence le Data Mining et les drones. Dans ce chapitre, nous consacrerons une section pour la technologie des drones dans laquelle nous présenterons d'abord la classification des drones, puis nous discuterons les différentes applications des drones dans l'agriculture intelligente.

Le chapitre 2 sera dédié à la présentation et à la discussion de l'application des techniques de Data Mining dans l'agriculture intelligente. Ces techniques sont utilisées afin de résoudre plusieurs problèmes agricoles comme la gestion de l'irrigation, la gestion des maladies et des ravageurs, la planification des intrants (eau, pesticides) et la prédiction du rendement. À la fin de ce chapitre, nous aborderons les défis de l'application de Data Mining dans l'agriculture intelligente.

Dans le chapitre 3, quelques concepts et connaissances de base pour le problème de déploiement seront fournis. Puis les différentes stratégies dédiées au déploiement des réseaux de capteurs sans fil (RCSF ou WSN pour Wireless Sensor Networks) seront classées en adoptant la classification de farsi et al. [76] en trois grandes catégories à savoir : les techniques classiques (techniques basées sur la force, techniques basées sur la grille, et techniques basées sur la géométrie de calcul), les techniques basées sur des méta-heuristiques et les techniques basées sur l'auto-programmation. De plus, les stratégies liées au déploiement des drones seront également présentées et discutées.

Le chapitre 4 sera consacré à la présentation du modèle de déploiement basé sur des méta-heuristiques que nous proposons dans le cadre de cette thèse. Dans ce chapitre, nous rappellerons brièvement certaines notions et concepts relatifs à l'optimisation multi-objectif. Ensuite, nous présenterons la description et la formulation mathématique du problème de déploiement. Nous exposerons également dans ce chapitre cinq méta-heuristiques à savoir MOABC, MOPSO, SPEA II, NSGA II, NSGA III, ainsi que leur adaptation pour résoudre le problème de placement de drones visant à optimiser la couverture, la consommation d'énergie et les chevauchements et à maintenir la connectivité.

Le chapitre 5 exposera les différents résultats des expérimentations réalisées pour évaluer notre approche qui a pour objectif de trouver le placement optimal des drones tout en respectant certaines exigences décrites dans le chapitre 4. Plus particulièrement, le premier volet de ce chapitre sera dédié à la description des scénarios d'application suivi de réglages des paramètres des différentes méta-heuristiques utilisées. Ensuite, à l'aide de deux indicateurs de performance, à savoir la métrique Hypervolume (HV) et la métrique distance générationnelle inversée (IGD : Inverted Generational Distance), les approximations du front de Pareto obtenues par les différents algorithmes d'optimisation multi-objectif pour chaque scénario seront comparées en termes de la diversité et de la convergence. Pour finir, un outil de visualisation sera utilisé pour voir les différents compromis fournis par les algorithmes.

Enfin, en conclusion générale, nous présenterons les perspectives de nos présentes propositions.

# **Première partie**

## **État de l'art**

# Chapitre 1

## Les nouvelles technologies au service de l'agriculture intelligente

### 1.1 Introduction

L'agriculture reste un secteur vital pour la plupart des pays. Elle présente la principale source de nourriture pour la population mondiale. Cependant, elle fait face à un grand défi : produire plus et mieux tout en augmentant la durabilité avec une utilisation raisonnable des ressources naturelles, réduire la dégradation de l'environnement ainsi que s'adapter au changement climatique.

Il est donc extrêmement important de passer des méthodes agricoles traditionnelles à une agriculture moderne. L'agriculture intelligente est l'une des solutions pour répondre à la demande croissante de nourriture tout en répondant aux exigences de durabilité. Elle est incarnée par l'utilisation de différentes technologies comme l'Internet des objets, l'intelligence artificielle et de Data Mining, les véhicules aériens sans pilote (drones), la robotique, etc.

Ce chapitre est dédié aux différentes technologies de l'agriculture intelligente. Ces technologies couvrent les technologies d'automatisation et d'acquisition de données, les technologies de transmission de données, les technologies de traitement de données, les RCSF, l'IoT, et les drones. Dans la section réservée à la technologie des drones, nous discutons la classification des drones, ainsi que les différentes applications de ces derniers dans le secteur agricole.

### 1.2 L'agriculture intelligente

L'agriculture intelligente est un concept qui fait rapidement son chemin dans le secteur agricole. Offrant un contrôle des cultures de haute précision, une collecte de données utile et des techniques agricoles automatisées, il y a clairement de nombreux avantages qu'une agriculture intelligente a à offrir. Un récent rapport de Beecham intitulé "Towards Smart Farming : Agriculture Embracing the IoT Vision" [135] prévoit que la production alimentaire doit augmenter de 70% en 2050 afin de répondre à la population mondiale estimée à 9,6 milliards de personnes. Il décrit également les préoccupations croissantes concernant l'agriculture à l'avenir : changement climatique, terres arables limitées, coûts et disponibilité de l'énergie, en particulier des combustibles fossiles. Alors,

quelle est la solution ? Agriculture intelligente.

### 1.2.1 Définition de l'agriculture intelligente

Le concept de l'agriculture intelligente n'est pas formellement défini pour tous ses aspects. De plus, cette notion peut être définie de différentes façons, toutefois nous avons retenu les définitions suivantes qui nous semblent être les plus appropriées :

- L'agriculture intelligente est un concept moderne dans lequel les technologies de l'information et de la communication sont utilisées pour gérer tous les processus et toutes les activités liées au domaine agricole [26].
- L'agriculture intelligente est une agriculture basée sur les réseaux. Elle est définie dans [239] comme « un service qui utilise des réseaux pour actualiser un service de convergence dans le domaine agricole pour faire face à divers problèmes, tels que le changement climatique, les conditions de croissance des produits agricoles et les maladies à l'aide des technologies de traitement de l'information et de contrôle autonome ». Elle doit être considérée sur la base des interactions entre des entités étroitement liées au domaine agricole c.à.d. les producteurs agricoles, les prestataires de services, les agents logistiques, les distributeurs de marché, les clients et les réseaux de télécommunications qui les interconnectent (voir FIGURE 1.1).
- L'agriculture intelligente adopte une démarche pluridisciplinaire, axée sur les nouvelles technologies. Elle intègre trois aspects technologiques différents qui sont la biotechnologie, la nanotechnologie et les technologies de l'information et de la communication, comme l'illustre la FIGURE 1.2 [187].
- L'agriculture intelligente va au-delà du concept d'agriculture de précision qui est décrit dans la littérature par plusieurs autres termes, à savoir l'application d'intrants spécifiques au site, la technologie agricole spécifique au site et le traitement à taux variable [170]. Ce dernier fait référence à un concept de gestion axé sur l'observation, la mesure et les réponses à la variabilité des cultures, des champs et des animaux [25]. Autrement dit, les animaux et les cultures reçoivent précisément le traitement dont ils ont besoin. L'agriculture intelligente ne se focalise pas seulement sur la détermination des différences dans le champ ou entre les animaux individuels ou la mesure précise mais également sur les données déclenchées par des événements en temps réel [208]. Cette agriculture recourt à des réseaux et à des outils de gestion des données. Elle implique toutes les opérations agricoles. En utilisant des appareils mobiles comme les tablettes et les smartphones, les agriculteurs peuvent accéder à des données en temps réel sur le climat, l'état des plantes et du sol, l'utilisation des ressources, etc.

L'utilisation des technologies numériques dans l'agriculture devrait apporter un certain nombre d'avantages, tels qu'une rentabilité et une productivité accrue avec une empreinte environnementale réduite. L'utilisation de ces technologies numériques est recommandée non seulement afin de réduire le risque de sécurité alimentaire auquel sont confrontées certaines régions du monde mais

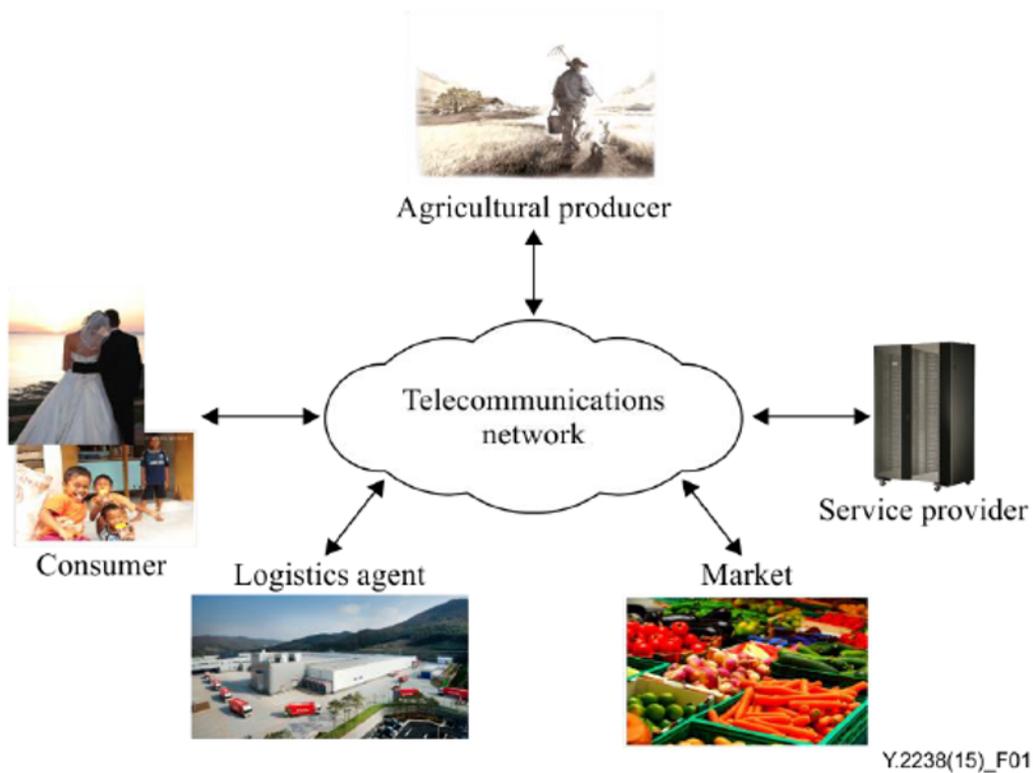


FIGURE 1.1 Diagramme conceptuel de l'agriculture intelligente basée sur les réseaux [239]

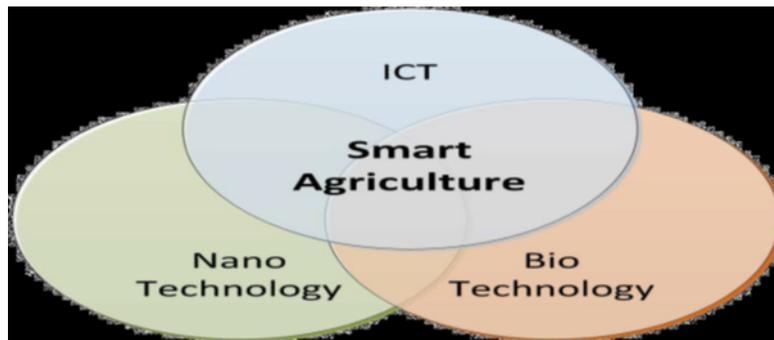


FIGURE 1.2 La position de l'agriculture intelligente [187]

aussi pour réduire l'empreinte environnementale et augmenter la rentabilité et la durabilité des exploitations et entreprises. Les technologies numériques ont un grand potentiel d'aider les agriculteurs à relever des défis plus spécifiques surtout avec l'intégration de l'IoT, des UAVs, des UGVs (Unmanned Ground Vehiculs ou véhicules terrestres sans pilote), les RFIDs (Radio Frequency Identification), les systèmes de positionnement (GPS : Global Positioning System), l'intelligence artificielle, le Data Mining, les smartphones, les tablettes, les capteurs, etc.

Dans cette forme d'agriculture qui portait aussi le nom agriculture numérique ou digitale, une

gamme de services pourrait être fournie, par exemple la gestion de l'irrigation, la gestion des maladies, la gestion des intrants (pesticides, fongicides, herbicides), la prédiction des rendements, la surveillance à distance, le contrôle des machines agricoles, la traçabilité des produits, le guidage des véhicules, le suivi des animaux, etc (voir FIGURE 1.3).

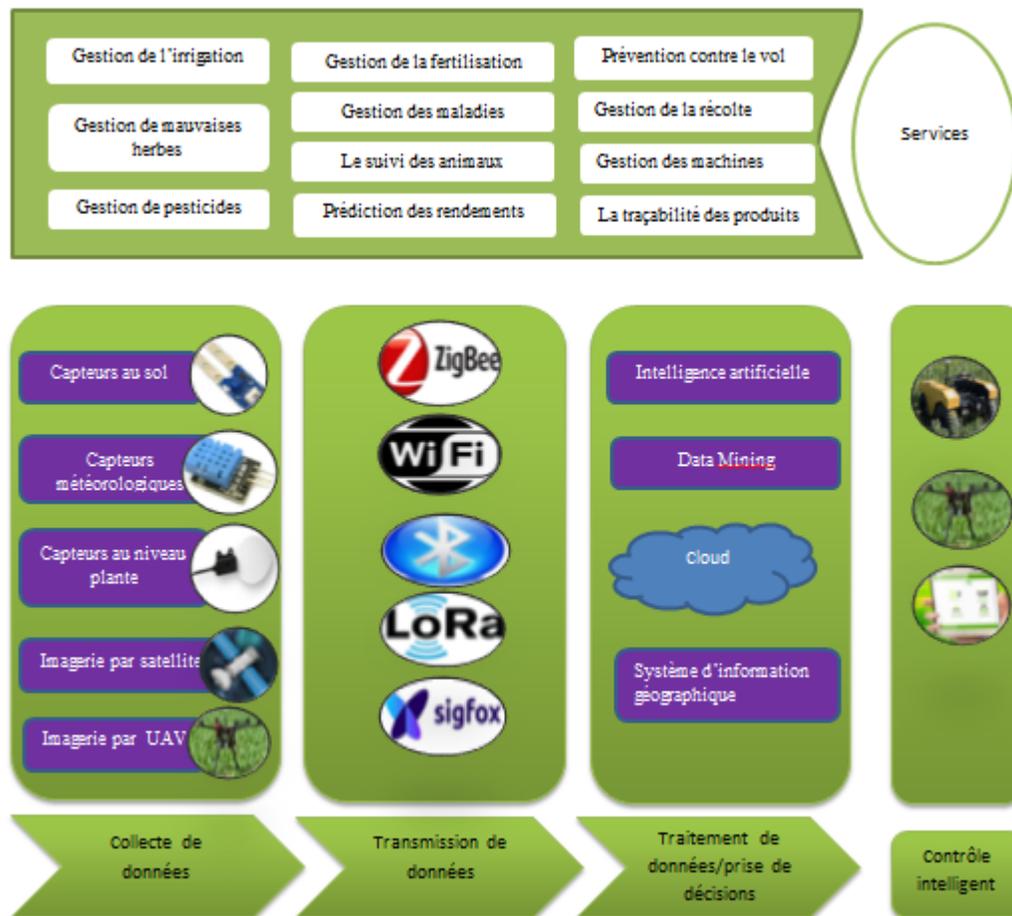


FIGURE 1.3 Les services fournis et les différentes technologies intégrées dans l'agriculture intelligente

Du point de vue de l'agriculteur, l'agriculture intelligente devrait lui apporter une valeur ajoutée sous la forme d'une meilleure prise de décision ou d'une exploitation et d'une gestion plus efficaces. Les applications d'agriculture intelligente ne ciblent pas seulement les grandes exploitations agricoles, mais pourraient également cibler l'agriculture familiale et l'agriculture biologique [9].

### 1.2.2 Les technologies de l'agriculture intelligente

Dans le passé, l'agriculture était basée sur les animaux et la force humaine. Cependant, au fil du temps, l'évolution technologique a progressivement modifié le mode de production agricole. Cette

évolution a ramené la mécanisation comme les tracteurs et les moissonneuses, et de nombreuses autres technologies qui peuvent être utilisées ensemble afin de rendre l'agriculture intelligente. Ces technologies sont réparties en trois catégories principales à savoir les technologies d'automatisation et d'acquisition, les technologies de transmission de données et les technologies de traitement de données. Ces technologies sont listées ci-dessous :

### 1.2.2.1 Les technologies d'automatisation et d'acquisition

L'introduction des technologies dans l'agriculture a permis l'acquisition de différents types de données. Ces technologies incluent les technologies de détection, de cartographie, de navigation et de machines.

#### a) Les technologies de capture

Les capteurs représentent un outil précieux pour la surveillance dans le domaine agricole. Ils permettent d'obtenir des informations objectives que ce soit sur les animaux, les cultures, le sol et le climat. Ces capteurs peuvent être déployés au sol ou dans l'air ou sur la plante. Dans l'agriculture intelligente, plusieurs types de capteurs sont utilisés :

- **Capteurs de lumière, capteurs sensibles à la pression (fluide), capteur de CO<sub>2</sub>, capteurs de vent** : Ces capteurs sont généralement utilisés afin de collecter des informations sur le climat dans les serres. A l'aide de ces capteurs, le contrôle des équipements comme celui de la ventilation, l'ouverture et fermeture des fenêtres, la commande de rideau, etc, est effectué.
- **Capteurs de température et d'humidité de l'air** : La température et l'humidité de l'air sont des paramètres importants pour la croissance de la plante. L'humidité relative fournit des informations sur la vapeur d'eau dans l'air, et qui est fortement interdépendante de la température.
- **Capteur d'humidité du sol** : Le capteur d'humidité peut mesurer la quantité d'humidité du sol qui l'entoure. Cette mesure est importante en agriculture pour aider les agriculteurs à gérer efficacement leurs systèmes d'irrigation.
- **Capteurs de vitesse de déplacement** : Ils permettent de déterminer la distance parcourue par exemple par le tracteur ou la moissonneuse-batteuse. Parfois, la vitesse de déplacement est mesurée avec un récepteur GNSS (Global Navigation Satellite System) ou un capteur à ultrasons ou un radar [170].
- **Capteurs de fluorescence** : La fluorescence des plantes causée par le rayonnement UV a été utilisée pour estimer l'état des plantes. Plus précisément, elle est utilisée pour l'identification des espèces et la croissance des plantes, le manque de nutriments, le manque d'eau, les effets de la température sur les plantes et la détection des attaques

- d'agents pathogènes des plantes [170]. Ainsi, beaucoup de capteurs pour l'étude de la fluorescence ont été utilisés.
- **Capteurs d'humidité des feuilles** : Ce type de capteurs permet de mesurer l'humidité au niveau de la feuille de la plante.
  - **Capteurs de lumière visible (RGB : Red Green Blue)** : Les cameras RGB fournissent des images de haute résolution qui peuvent être acquises dans différentes conditions. Elles sont peu coûteuses par rapport à d'autres types de capteurs. Elles sont légères et faciles à utiliser. De plus, les informations acquises nécessitent un traitement simple [218]. L'inconvénient principal de ces capteurs est le fait qu'ils sont insuffisants pour analyser de nombreux paramètres de végétation qui nécessitent des informations spectrales dans le spectre non visible [218]. Ce sont les capteurs les plus utilisés dans les drones.
  - **Capteurs thermiques** : Ce type de capteurs peut lire la température rayonnée d'un objet. Les caméras thermiques utilisent des capteurs infrarouges et une lentille optique pour recevoir l'énergie infrarouge [218]. Des images colorées peuvent également être générées par de nombreux capteurs thermiques. Les capteurs thermiques sont utilisés pour des tâches spécifiques comme la prédiction du stress hydrique des plantes [111].
  - **Capteurs multi-spectraux** : Ils captent la lumière réfléchie : rouge, vert, et deux bandes infrarouges. Ils sont les plus utilisés par rapport aux capteurs hyper-spectraux en raison de leur coût [218]. Ils permettent de produire l'indice de végétation par différence normalisée (NDVI : Normalized Difference Vegetation Index) qui est utilisé pour calculer plusieurs paramètres liés à la plante comme la biomasse, la maturation des plantes, l'état des nutriments, la teneur en eau, etc. Les rapports NDVI sont utiles surtout dans la détection de la présence des parasites, des mauvaises herbes, la création d'une prescription d'engrais et beaucoup d'autres problèmes liés aux cultures.
  - **Capteurs hyper-spectraux** : Les capteurs hyper-spectraux enregistrent de nombreuses longueurs d'onde de lumière visible et invisible [145]. Les caméras hyper-spectrales peuvent produire un grand nombre d'indices de végétation par rapport aux caméras multi-spectrales [170]. Les images prises par ces deux types de capteurs nécessitent des méthodes complexes de prétraitement pour extraire des informations utiles [218]. Ces capteurs permettent d'identifier le type spécifique de plante en mesurant la couleur de la lumière qu'elle reflète, ce qui faciliterait le repérage de choses comme les mauvaises herbes résistantes aux herbicides.
  - **Capteurs LiDAR (Light Detection And Ranging)** : **LiDAR** signifie la télédétection par laser ou détection et estimation de la distance par la lumière. Concernant leur application dans le domaine agricole avec les UAVs, par rapport aux autres types de capteurs utilisés pour l'acquisition de données, ce type de capteur n'a pas été amplement utilisé

en raison de son coût élevé [218]. A titre d'exemple il est utilisé avec d'autres capteurs afin de compter le nombre de fruits [57].

- **Capteurs de position, d'orientation et de distance :** Les machines agricoles telles que les drones et les UGVs peuvent être équipées d'un GPS, d'un accéléromètre et de capteurs de mesure inertielle pour localiser la machine ainsi que les données acquises, et détecter tout changement dans l'orientation. Des capteurs de distance et de vitesse sont aussi utilisés afin de mesurer la distance entre la machine et un objet ou de détecter sa vitesse.
- Etc.

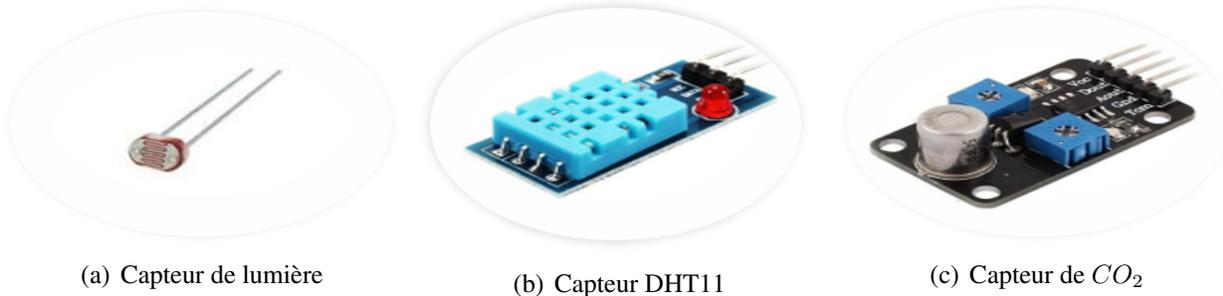


FIGURE 1.4 Exemples de capteurs utilisés

## b) Les technologies de la cartographie

Les systèmes d'informations géographiques (SIG), les GNSS et les capteurs permettent d'avoir des cartographies qui ont déjà prouvées leur valeur dans le domaine agricole pour la gestion des ressources et la prédiction des rendements. Avec ce type de technologie, la création de cartes d'élévation, de sols, de rendements et de cultures est possible. L'élévation est très utile pour comprendre la réponse au rendement. Elle influence la formation du sol, le mouvement de l'eau et les aspects cultureux [235]. Elle peut déterminer le risque d'érosion et les zones gorgées d'eau [170]. Les techniques photogrammétriques sont particulièrement utilisées pour construire des modèles numériques d'élévation d'une ferme ou d'un champ afin d'exploiter les informations 3D concernant la végétation. Plus précisément, elles peuvent être utilisées pour classer les caractéristiques du terrain telles que l'aspect, la pente, la courbure, les directions d'écoulement de l'eau, l'interception du rayonnement solaire et les indices d'humidité topographique. Les cartes d'élévation peuvent aider à identifier comment la topographie peut affecter les résultats agronomiques dans un champ [235].

La création de cartes des sols représente une étape très importante surtout aux applications à taux variables. L'échantillonnage du sol fournit des informations sur la disponibilité des nutriments pour la croissance des cultures, la texture du sol, et d'autres propriétés chimiques du sol [170]. La cartographie peut être réalisée en utilisant la technique en grille, aléatoire ou

adaptative. Avec le développement technologique, la cartographie des propriétés du sol d'un champ peut être réalisée en utilisant des capteurs (capteurs électriques et électromagnétiques, capteurs optiques et radiométriques, capteurs électrochimiques, etc.) qui peuvent aussi être combinés avec un récepteur GNSS.

La cartographie du rendement est une technologie de collecte de données géo-référencées sur les caractéristiques et le rendement des cultures [78]. Les cartes peuvent être créées à l'aide des données collectées sur différents types de machines comme la moissonneuse-batteuse et les arracheuses de betteraves sucrières et de pommes de terre, etc. Pour cartographier les rendements en grain, des composants principaux devraient être utilisés notamment le capteur de débit de grain mesurant le volume de grain récolté, le capteur d'humidité du grain quantifiant la variation d'humidité, une antenne GPS, capteur de vitesse de déplacement qui fournit la distance que la moissonneuse a parcourue pendant un certain intervalle d'enregistrement, le capteur de vitesse d'élévateur à grain, l'affichage du moniteur de rendement avec un récepteur GNSS, le capteur de position d'en-tête qui lance la mesure du grain lorsque la tête est abaissée<sup>1</sup>.

La cartographie des cultures est basée sur des données provenant de satellites, de drones, d'avions ou de tracteurs. Ainsi, grâce à la télédétection, le type de cultures et leur stade de développement peuvent être identifiés.

#### c) **Les technologies de navigation**

Ces technologies permettent d'enregistrer la position réelle qui peut être utilisée à plusieurs fins comme la cartographie, le guidage, etc. Les systèmes mondiaux de navigation par satellite sont utilisés pour localiser l'emplacement géographique du récepteur d'un utilisateur partout dans le monde. Actuellement, il existe deux systèmes GNSS opérationnels (GPS et GLONASS) et deux systèmes en développement (Galileo et BeiDou) [170].

#### d) **Les technologies de machines**

Partant de smartphones et de tablettes, l'agriculture intelligente nous présente également un monde de robots et d'avions dotés de plusieurs capteurs pour accomplir différentes missions. Les technologies qui envahissent ces dernières années le domaine agricole sont :

- ◇ **Tracteurs sans conducteur** : Les tracteurs aujourd'hui constituent le cœur battant de l'agriculture. Au cours des dernières années, le travail du sol avec des tracteurs automobiles a dominé partout dans le monde développé, dans certaines parties du monde en développement et dans les grandes exploitations agricoles du monde entier [120]. Il s'agit essentiellement de tracteurs semi-autonomes guidés par le conducteur, équipés de moteurs électriques ou à essence. Ils s'appuient sur le travail humain pour assurer un bon ameublissement du sol, semer les graines aussi uniformément que possible et

---

1. <https://cropwatch.unl.edu/ssm/mapping>

effectuer des perturbations entre les rangs (mise à la terre). Avec le développement technologique, des tracteurs sans conducteurs sont introduits dans l'agriculture. Cette technologie peut être conduite à l'aide de GPS et d'un ordinateur à bord<sup>2</sup>.

- ◇ **Les drones ou UAVs :** Ils sont des aéronefs sans pilote humain à bord. Les drones sont des appareils qui peuvent voler à l'aide du pilote automatique et des coordonnées GPS sur un parcours prédéfini ou peuvent être utilisés manuellement avec des signaux radio à l'aide de la télécommande ou de l'application pour smartphone [185]. Avec la disponibilité de tant de capteurs, cette technologie peut nous approvisionner avec des informations en temps réel, plus précises avec plus de détails et moins d'erreurs. Ils sont une bonne alternative à l'utilisation d'avions qui ont été utilisés depuis les années 1920 comme un moyen efficace d'appliquer des pesticides et des herbicides pour la lutte contre les ravageurs et les maladies sur de grandes surfaces [35]. Bien que les avions agricoles soient utiles pour des applications telles que l'application rapide des produits chimiques nécessaires aux cultures sans perturber ou compacter le sol autour des plantes, ils ne conviennent pas aux applications agricoles qui nécessitent une plus grande précision [35]. De plus, d'un point de vue économique, ces avions ne sont pas un bon choix pour les petites exploitations. L'utilisation de drones fait partie des tendances les plus récentes et semble se propager rapidement. Les UAVs sont une technologie précieuse, en particulier pour surveiller les cultures et les animaux, analyser le sol, planter les cultures, identifier les mauvaises herbes, pulvériser les produits chimiques et effectuer l'irrigation. Un aperçu des différentes applications des drones dans l'agriculture est donné dans la section 1.3 de ce chapitre. Contrairement aux avions, les drones sont généralement alimentés par des batteries qui limitent leur temps de vol.
  
- ◇ **Les UGVs :** Ces véhicules sans opérateur humain sont devenus populaires pour l'utilisation agricole. Ils sont développés afin de surmonter le problème de la pénurie de la main-d'œuvre humaine dans l'ensemble de l'industrie agricole et d'améliorer la sécurité sanitaire des aliments tout au long du cycle de production des produits agricoles. La capacité unique des UGVs à voyager à travers les champs, tout en prenant en charge des charges utiles importantes, les rend idéals pour un environnement agricole [34]. Tout comme les UAVs, les UGVs sont aussi dotés de différentes sortes de capteurs qui leur permettent d'accueillir des données pertinentes à leur mission. Le contrôle du mouvement d'un véhicule terrestre sans pilote est fait à l'aide d'une commande manuelle ou autonome [34]. Pour les UGVs mobiles, le contrôle manuel peut être exécuté en utilisant un appareil mobile ou une tablette. Les actions de contrôle peuvent être envoyées au véhicule à proximité ou à distance, selon la technologie sans fil. Pour le second type de contrôle, la position ou l'emplacement du véhicule doit être déterminé et interprété comme une entrée pour réaliser une navigation autonome. Ils doivent être conçus pour interagir avec l'environnement particulier où ils sont utilisés [35]. Ces véhicules servent à exécuter plusieurs tâches dans le domaine agricole comme l'échantillonnage du sol, la gestion de l'irrigation, la pulvérisation de précision, le désherbage mécanique et la récolte des cultures. Les caractéristiques principales de ces robots typiques sont leur

---

2. <http://www.merlofarminggroup.com/agri-view-driverless-tractors>

petite taille, leur légèreté, et leur autonomie énergétique [170].

Les technologies introduites dans l'agriculture ont aussi permis l'automatisation des équipements agricoles pour contrôler les paramètres liés aux cultures des végétaux. Cette automatisation est à base d'un domaine multidisciplinaire de la science comme l'électronique, la mécanique, l'informatique, etc.

Le contrôle des installations dans les champs ou dans les serres comme moteur hors-bord, ventilateur, pompe à eau, ouverture et fermeture de portes, fenêtres et toit, réglages de l'éclairage, déclenchement/arrêt du chauffage par des systèmes de commande est réalisé suivant les conditions du milieu concerné.

### 1.2.2.2 Les technologies de transmission de données

Les déploiements de capteurs se font dans des zones éloignées. Pour recueillir leurs informations ou données, les technologies de communication sans fil constituent le meilleur support. Selon [187], avant de sélectionner une technique de communication des capteurs ou des nœuds, trois contraintes devraient être prises en compte autre que le prix à savoir la portée de communication, la batterie et la sécurité des données. Les technologies de communication sans fil utilisées dans le domaine agricole sont les suivantes :

- **ZigBee** : c'est la plus utilisée et la plus adaptée pour le domaine agricole [187, 97]. Compte tenu des exigences des applications agricoles, Zigbee peut jouer un rôle vital, en particulier en ciblant la surveillance des conditions environnementales, l'installation d'irrigation automatisée, etc [102]. Les réseaux Zigbee prennent en charge trois types de topologie à savoir la topologie en étoile, la topologie maillée et la topologie en arbre. Les périphériques basés sur ZigBee peuvent être de trois types, particulièrement coordinateur, routeur et utilisateur final.
- **Bluetooth** : est une technologie de communication sans fil qui connecte des appareils sur des distances plus courtes, coopérant généralement à proximité. En raison de ses avantages de faible coût, facile à utiliser et faible consommation d'énergie, cette technologie est utilisée dans de nombreuses applications d'agriculture intelligente comme l'irrigation [97].
- **Wi-Fi** : est une autre technologie de communication sans fil avec une portée de plusieurs dizaines de mètres qui est utilisée chaque fois que des communications LAN (Local Area Network) sont nécessaires dans une agriculture intelligente.
- **GPRS** (General Packet Radio Service)/ **GSM** (Global System for Mobile Communications) : sont des systèmes de téléphonie cellulaire longue distance qui trouvent aussi une large application dans l'agriculture intelligente. A titre d'exemple, le GPRS a été utilisé pour concevoir et implémenter un système de surveillance et de contrôle d'un ensemble de serres [130].
- **LoRa** (Long Range Radio) et **sigfox** sont aussi utilisées dans le domaine agricole [23]. Lorsque les informations agricoles doivent être transmises sur de longues distances, il sera de préfé-

rence d'utiliser LoRa en raison de sa faible consommation d'énergie et de sa grande portée [97].

### 1.2.2.3 Les technologies de traitement de données

Avec la prolifération des technologies numériques, des réseaux et des capteurs, de la technologie d'identification par radiofréquence, du système de positionnement mondial, etc, l'agriculture est devenue un monde axé sur les données. Cependant, le volume et la complexité croissants des ensembles de données ont rendu la tâche d'extraction d'informations utiles pour soutenir la prise de décision effectuée par des analystes très ardue. Dans une telle situation, l'intelligence artificielle pourrait aider à surmonter ce défi avec des techniques (comme par exemple les algorithmes génétiques (Genetic Algorithms (GA)), la logique floue, l'apprentissage en profondeur (Deep Learning (DL)) et les systèmes experts) permettant d'explorer les données et d'extraire des informations utiles.

Le Data Mining permet d'extraire des informations précieuses à partir de données aussi volumineuses et de découvrir des modèles inconnus et des relations cachées dans les données qui peuvent être importantes pour différents problèmes [252].

Il s'agit d'un processus qui implique des méthodes et des outils issus de différents domaines de l'informatique, des statistiques ou de l'intelligence artificielle. Le Data Mining est une approche interdisciplinaire qui a été adoptée dans différents domaines : optimisation des systèmes industriels, analyse de marché pour découvrir des modèles financièrement intéressants, analyse de données médicales, web (Text mining). Son application en agriculture est le sujet principal de certaines études récentes [63, 227, 148, 80, 147, 105, 188]. L'application de Data Mining dans l'agriculture intelligente et ses défis seront présentés et discutés dans le chapitre 2.

La clé du succès d'une agriculture intelligente est l'accès à des informations en temps opportun et aux décisions élaborées. Les systèmes d'information de gestion des exploitations agricoles (SIGF) constituent un grand support de diffusion des données pour mener à bien les opérations agricoles. Ils sont définis comme un système planifié de collecte, de traitement, de stockage et de diffusion des données sous une forme nécessaire [170].

### 1.2.2.4 Les réseaux de capteurs sans fils

Les réseaux de capteurs sans fil est l'une des technologies intégrées, ils se basent sur des capteurs déployés dans l'environnement de surveillance ou de contrôle, ils acquièrent des données pertinentes qui seront utilisées pour automatiser les tâches agricoles, ou pour effectuer le suivi, ou pour aider le fermier à prendre des décisions pertinentes.

Cette technologie est utilisée aussi bien dans l'élevage des bétails et de volailles [153, 160] que dans les serres et les champs agricoles. Dans le cadre de ce travail nous nous limiterons à la production végétale. Elle est utilisée dans les champs, par exemple pour donner un support à l'irrigation de précision [236]. [251] l'a développée et déployée pour détecter le stress des plantes à faible coût et à faible énergie en temps réel dans les fermes à petite échelle de l'Afrique subsaharienne qui ne

sont pas connectées au réseau électrique et ont des options limitées de connectivité internet.

Xiao et Guo [237] ont décrit la conception d'un nouveau réseau de capteurs sans fil pour la surveillance de l'environnement agricole. L'évaluation de la robustesse et de la longévité du réseau sur le terrain a montré que le système améliore l'efficacité de la production agricole par rapport à l'agriculture traditionnelle.

Dans les serres, en se basant sur le RCSF, [130] a conçu et implémenté un système de surveillance intelligent en temps réel. Ce système est capable de surveiller et contrôler un grand nombre de serres divisées en régions. Chaque région est équipée d'un ensemble de capteurs dans le but de collecter la température, l'humidité et la lumière, qui sont envoyées au puits de chacune. Ces derniers à leur tour les envoient via GPRS au centre de surveillance. Les puits peuvent aussi recevoir les commandes de contrôle envoyées par le centre de surveillance ou par les terminaux distants des utilisateurs et contrôler les dispositifs correspondants.

Dans [142], le Rgroupe a utilisé des appareils appelés « sensicast » pour surveiller la température de l'air, l'humidité, et la température du sol dans une serre de tomates au sud de l'Italie grâce à un réseau de capteurs sans fil. Les données collectées par les nœuds capteurs sont transmises via LAN (Local Area Network) à une station de base. L'un des objectifs de Rgroupe est de développer une architecture pour le système d'aide à la décision dans l'agriculture basée sur les réseaux de capteurs sans fil et de développer un modèle de maladies plus précis aidant le fermier à améliorer la qualité des cultures.

Une autre forme du RCSF a émergée dans l'agriculture WUSN (Wireless Underground Sensor Network). [247] a donné un aperçu sur les réseaux de capteurs souterrains sans fil. Les nœuds WUSN sont enterrés et ne nécessitent aucune connexion câblée. L'architecture a été présentée et une classification WUSN a été proposée. Les résultats expérimentaux montrent que la profondeur d'enfouissement et la fréquence, la teneur volumétrique en eau du sol sont des paramètres importants qui affectent la communication WUSN.

### 1.2.2.5 L'internet des objets

L'internet des objets est un concept qui permet à des milliards d'objets d'être connectés à Internet. L'IoT implique l'utilisation de différents équipements tels que des capteurs et des actionneurs, des microcontrôleurs, etc [156].

L'agriculture intelligente est déjà de plus en plus courante chez les agriculteurs, et l'agriculture de haute technologie devient rapidement la norme grâce à l'IoT. Ci-dessous, nous avons décrit les applications de l'IoT dans l'agriculture et comment « l'agriculture de l'Internet des objets » aidera les agriculteurs à répondre aux demandes alimentaires mondiales dans les années prochaines.

Les agriculteurs ont déjà commencé à utiliser certaines techniques et technologies agricoles de haute technologie afin d'améliorer l'efficacité de leur travail quotidien. Par exemple, des capteurs placés dans les champs et connectés à internet permettent aux agriculteurs d'obtenir des cartes détaillées des ressources et de la topographie de la zone, ainsi que des variables telles que la tem-

pérature et l'acidité du sol. Ils peuvent également accéder aux prévisions climatiques pour prévoir les conditions météorologiques dans les prochains jours et semaines.

Les agriculteurs peuvent utiliser leurs smartphones pour surveiller à distance leurs équipements, leurs cultures, ainsi que pour obtenir des statistiques sur l'alimentation et les produits phytosanitaires. Ils peuvent même utiliser cette technologie pour effectuer des prévisions statistiques sur leurs cultures.

Dans le but de fournir une solution à l'attaque des insectes et des ravageurs, à l'attaque des animaux sauvages et des oiseaux lors de la croissance des cultures et à la possibilité de vol de cultures lors de la récolte, [85] a proposé un système qui inclut :

- Un robot contrôlé à distance basé sur GPS, il est doté de divers capteurs et dispositifs comme la caméra, le capteur d'obstacle, la sirène, le coupeur et le pulvérisateur pour réaliser des tâches comme : garder la vigilance, effrayer les oiseaux et les animaux, désherber et pulvériser.
- Une irrigation intelligente avec un contrôle et une prise de décision basée sur des données précises en temps réel.
- Une gestion intelligente d'entrepôt qui comprend des capteurs comme celui de l'humidité, de la température, de la lumière et le détecteur de mouvement, ainsi que le chauffage et le ventilateur pour la maintenance de la température et de l'humidité.

Le contrôle de toutes ces opérations se fait via une machine distante ou d'un ordinateur connecté à Internet. Les opérations sont effectuées à l'aide de capteurs, caméras et actionneurs, modules Wi-Fi ou ZigBee, un microcontrôleur et Raspberry pi. Toutes les observations et essais expérimentaux montrent que le projet est une solution complète aux activités sur le terrain, aux problèmes d'irrigation, et de stockage. La mise en œuvre d'un tel système sur le terrain peut absolument participer à améliorer le rendement des cultures et la production globale.

Shruti et Patil [203] ont proposé une nouvelle solution basée sur l'IoT pour surveiller les paramètres environnementaux dans le champ agricole et contrôler l'irrigation. Le système proposé peut être exécuté sur les smartphones, les tablettes ou PC au niveau du client. La solution fournie peut être surveillée en temps réel par tous les périphériques distants y compris les portables et les tablettes. Ce qui fournit la flexibilité pour la compréhension et la visualisation des données et l'analyse prédictive sur l'état futur de l'environnement.

L'agriculture intelligente prend son envol, mais elle pourrait simplement être le précurseur d'une utilisation encore plus grande de la technologie dans le monde agricole. BI Intelligence [146], le service de recherche premium de Business Insider, prévoit que les installations d'appareils IoT dans le monde agricole passeront de 30 millions en 2015 à 75 millions en 2020, pour un taux de croissance annuel de 20%. OnFarm, une entreprise de technologies agricoles, s'attend à ce qu'une ferme moyenne pourrait générer en moyenne 4,1 millions de points de données par jour en 2050, contre 190 000 en 2014 [146].

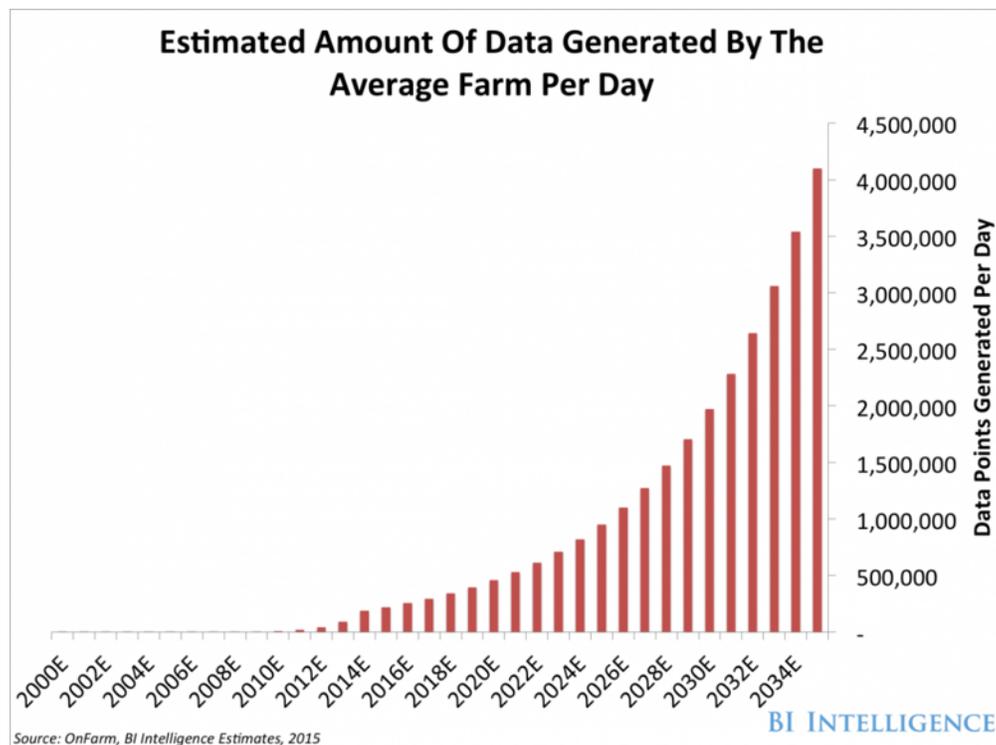


FIGURE 1.5 Quantité estimée de données générées par une ferme moyenne par jour[146]

Parmi les nombreux avantages que l’IoT apporte à la figure, sa capacité à innover dans le paysage des méthodes agricoles actuelles est absolument révolutionnaire. Les capteurs IoT capables de fournir aux agriculteurs des informations sur les rendements des cultures, les infestations de ravageurs, les précipitations et la nutrition des sols qui sont inestimables pour la production et offrent des données précises qui peuvent être utilisées pour améliorer les techniques agricoles au fil du temps.

L’IoT garantit une planification efficace, ce qui garantit que la meilleure utilisation des objets connectés améliore la productivité. La FIGURE 1.6 montre un aperçu schématique indiquant les tendances agraires qui fournissent des interactions pratiques simples et rentables grâce à une connectivité à travers des observations individuelles de serre, de bétail, d’agriculteur et de terrain. Par conséquent, le système agricole basé sur l’IoT utilise les composants sans fil pour permettre la surveillance des fermes en temps réel.

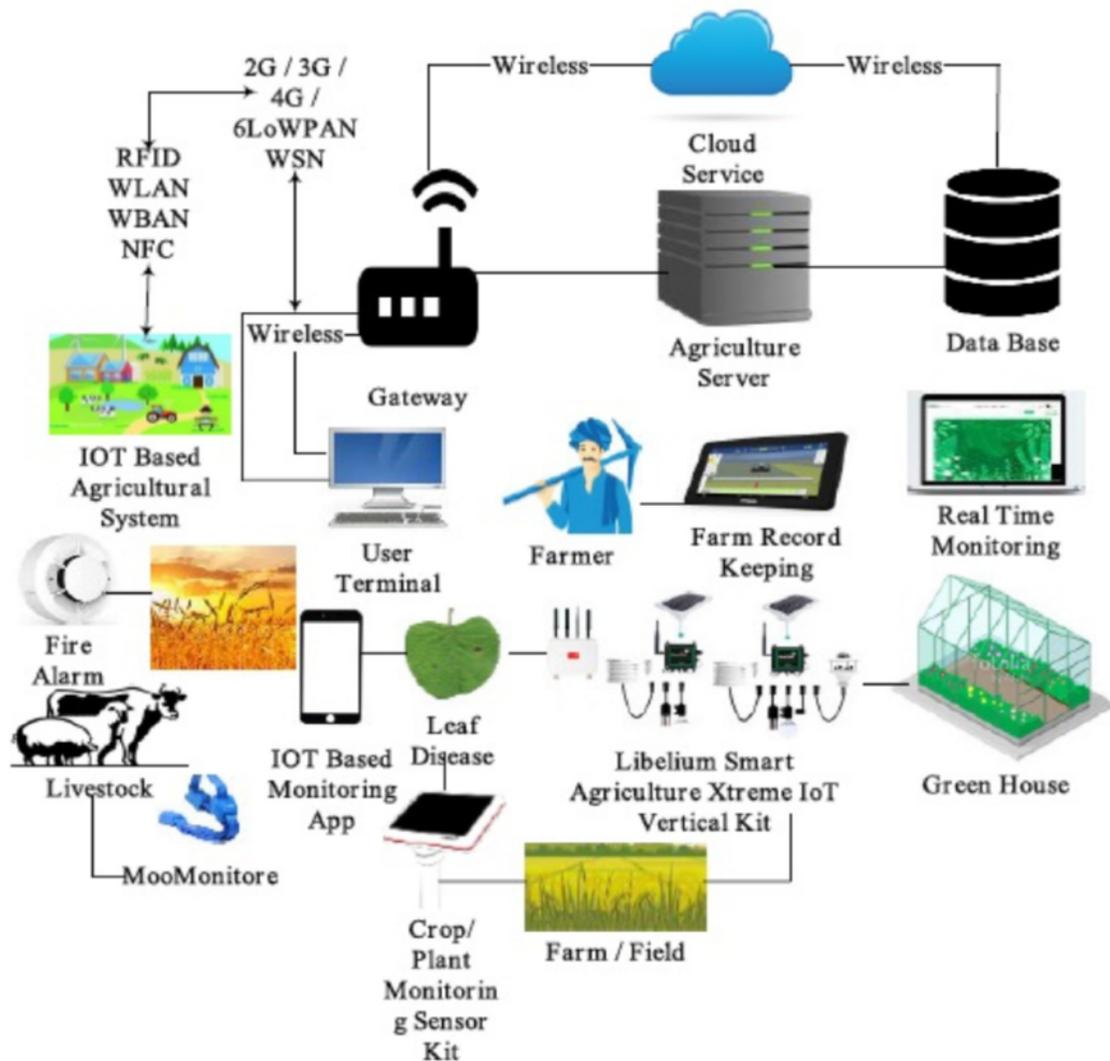


FIGURE 1.6 Les tendances de l'agriculture intelligente [75]

### 1.2.2.6 La combinaison de trois technologies : IoT, UAV, RCSF

Une nouvelle tendance actuelle qui pourrait être importante pour la gestion des fermes et des exploitations agricoles consiste à concevoir un système multiforme en se basant sur la combinaison de technologies comme l'IoT, UAV et RCSF. Le déploiement des réseaux de capteurs sans fil dans le domaine agricole permet la surveillance et le contrôle des cultures. La collecte de différentes données de divers capteurs, le contrôle des actionneurs et l'accès aux données peuvent se faire via internet, mettant en évidence l'internet des objets.

Selon [176], un schéma directeur d'un système de surveillance collaboratif aussi complexe est composé de quatre éléments principaux qui sont : UAV, RCSF au sol, stations de contrôle au sol (GCS : Ground Control Stations) et Internet / système mondial de réseaux de communications mobiles (GSM), comme illustré par la FIGURE 1.7.

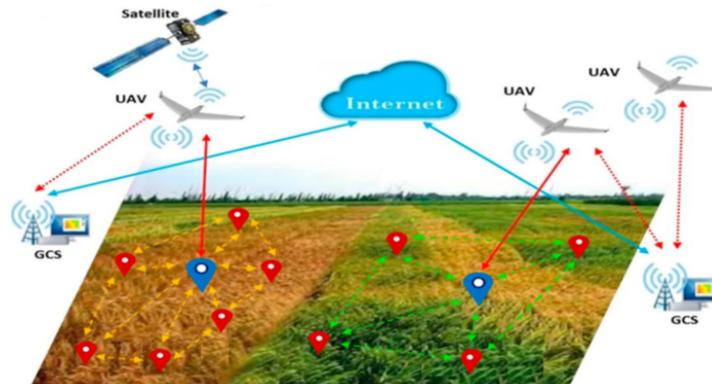


FIGURE 1.7 Concept d'un système intégré de réseau de capteurs sans fil - véhicule aérien sans pilote (RCSF – UAV)[176]

D'après une étude faite par les mêmes auteurs de [176], le développement d'un système d'agriculture intelligente puise ses principes méthodologiques et conceptuels dans différents champs disciplinaires et il se trouve au croisement de plusieurs disciplines convergentes : UAV, RCSF et IoT (voir FIGURE 1.8).

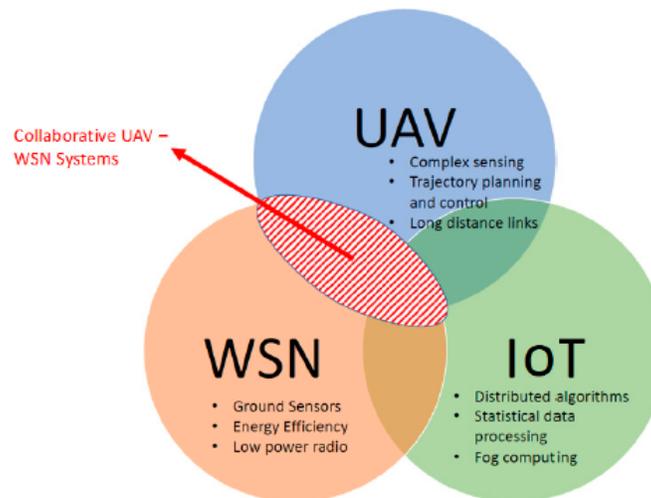


FIGURE 1.8 Convergence UAV, RCSF et IoT dans l'agriculture intelligente [176]

Avec l'avancement technologique, de nouvelles portes s'ouvrent aussi pour l'utilisation d'un nouveau paradigme nommé Internet des Drones (IoD : Internet of Drones) [5]. Ce dernier représentant une branche de l'Internet des objets est utilisé pour la communication entre drones.

## 1.3 Les drones dans l'agriculture intelligente

Les drones apportent chaque jour une réelle valeur ajoutée aux agriculteurs en maximisant leur capacité de production, grâce à des analyses de terrain en temps réel basées sur des vues de terrain détaillées. Ils sont équipés de caméras et de différents capteurs pour contrôler et analyser les cultures et faire un diagnostic de leur état de santé. Ils survolent les champs et capturent des images de haute précision.

Ces images aériennes de drones peuvent être de précieuses sources d'informations. Des images en temps réel de haute qualité offrent une bonne reconnaissance et une bonne classification des cultures pour surveiller leur croissance [132, 158] et prévenir les maladies et les ravageurs qui peuvent les endommager [59]. En surveillant constamment les cultures et en localisant les zones à traiter, les agriculteurs sont en mesure de réduire la consommation d'intrants (pesticides et eau), leur donnant la possibilité de dépenser moins et d'améliorer les rendements en ciblant les zones à traiter [15] et en gérant avec précision les quantités.

En plus de tous ces services, ils contribuent de manière significative à la collecte de données à partir de réseaux de capteurs sans fil [20, 53]. Compte tenu des réseaux de capteurs à grande échelle [3], en particulier dans les grandes exploitations et la particularité des environnements de déploiement, la collecte de données aériennes basée sur un drone contrôlable présente un certain nombre d'avantages, notamment la vitesse de collecte des données. Un drone peut également agir comme relais qui transmet les commandes reçues à l'infrastructure terrestre ou comme système de contrôle (s'il dispose d'un système de décision intégré).

### 1.3.1 Classification des drones

Les drones sont devenus un axe de recherche très important. En effet, les fabricants et les chercheurs de drones proposent divers modèles pour le domaine agricole. Puri et al. [181] ont présenté les différents drones disponibles sur le marché et leurs spécifications techniques pour la surveillance et le contrôle agricole.

Différents types de drones sont disponibles aujourd'hui, mais ne sont pas tous adaptés pour l'agriculture. Les drones qui sont convenables sont classés en deux catégories à savoir les drones à rotors multiples et les drones à voilure fixe. Chacune de ces deux classes a ses avantages et ses propres limites. Les drones à voilure fixe ont une capacité de vol à longue portée avec une vitesse très élevée. Par contre les drones à rotors multiples, peuvent survoler à une vitesse constante, ils peuvent décoller et atterrir dans de petits espaces confinés, en plus ils sont plus rapides à installer sur le terrain.

Les drones peuvent être classés aussi selon leur autonomie en drones semi-autonomes et drones autonomes. Les drones de la première classe reçoivent les instructions des opérateurs. Ils sont télécommandés afin de prendre des décisions critiques comme la trajectoire de vol et la photographie. Par contre dans la deuxième classe, les trajectoires de vol des drones sont prédéterminées. Les trajectoires de vol et d'autres activités (photographie) sont dirigées par des programmes informatiques.

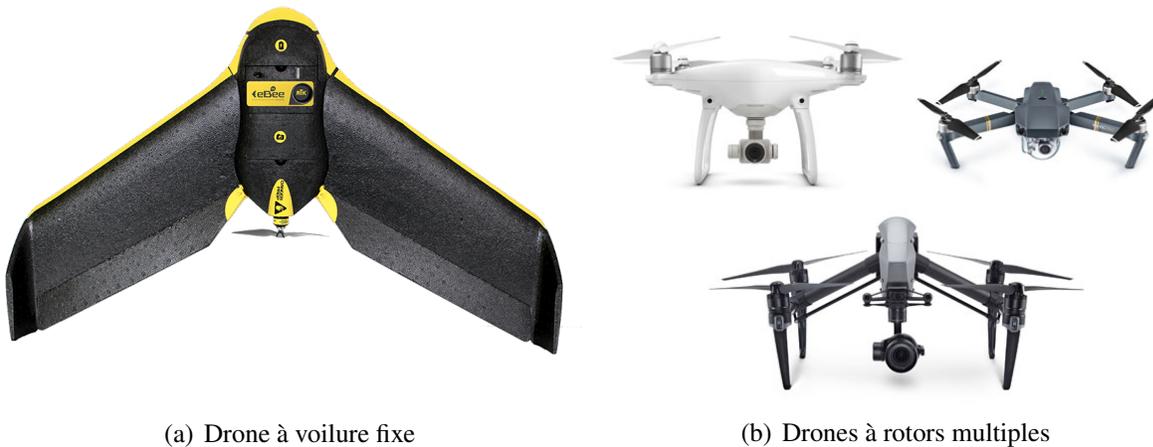


FIGURE 1.9 Types de drones agricoles<sup>3</sup>

Différents composants matériels sont intégrés dans les drones afin de leur permettre d'accomplir plusieurs tâches comme la cartographie, la surveillance, etc, et parmi ces composants matériels on trouve les capteurs. Les capteurs montés sur drone les plus couramment utilisés sont : les capteurs de lumière visible (RGB), les capteurs thermiques, les capteurs multi-spectraux, les capteurs hyper-spectraux, les capteurs LiDAR et les capteurs de position, d'orientation et de distance.

### 1.3.2 Différentes applications de drones dans l'agriculture intelligente

Les drones sont une nouvelle technologie passionnante qui pourrait changer la façon dont les agriculteurs gèrent leurs fermes et pourrait devenir aussi omniprésente que les tracteurs. Le rôle joué par un drone dans le secteur de l'agriculture n'est pas moins important que celui qu'il joue dans d'autres secteurs. Ils peuvent accomplir plusieurs missions et remplir plusieurs tâches en facilitant le travail de l'agronome et de l'agriculteur.

#### 1.3.2.1 Les drones pour la gestion des maladies et parasites des cultures

Une estimation faite par FAO indique que les parasites et les maladies causent une perte de 20 à 40 % du rendement total des cultures chaque année malgré l'application de plusieurs tonnes de pesticides [117]. L'utilisation des équipements intelligents comme les drones permettrait une application précise des produits chimiques et l'élimination des parasites pour obtenir des aliments sains et de qualité.

Les drones couvrent plusieurs hectares dans le but de scanner ou de prendre des images des différentes parcelles agricoles. Ils sont capables de voir les cultures avec différents indices [149]. En se basant sur l'analyse des données collectées et les coordonnées GPS de chaque image prise, une pulvérisation précise des produits chimiques est appliquée.

---

3. <https://blog.dronedeploy.com/choosing-the-right-mapping-drone-for-your-business-part-i-multi-rotor-vs-fixed-wing-aircraft-6ec2d02eff48>

Un système basé sur un drone permettant de gérer les parasites et de réduire les effets néfastes de pulvérisation des produits chimiques sur la santé de pulvérisateur est proposé dans [149]. Il est doté de deux composants principaux à savoir un système de surveillance qui est basé sur une caméra multi-spectrale permettant de prendre des images analysées par l'indicateur géographique NDVI et un système de pulvérisation qui est monté sur la région inférieure de l'UAV. Un autre système autonome est développé dans [177] afin de contrôler les parasites. Les auteurs proposent un drone à quatre hélices et un module de pulvérisation.

Dans [15], les auteurs ont présenté un drone autonome avec un module de décision embarqué basé sur la vision. Ce drone est capable de détecter et réaliser des actions dans une position spécifique de la cible par exemple l'application des herbicides. Les auteurs de [53] ont proposé une architecture basée sur des drones pour l'application de produits chimiques dans une zone agricole. La pulvérisation de produits chimiques est effectuée sur la base des données fournies par un réseau de capteurs sans fil déployé sur le terrain. Ces informations étaient importantes pour ajuster la route des UAVs pour une pulvérisation guidée.

Dans le but de soutenir le domaine agricole dans la surveillance des champs pour vérifier et contrôler la présence de parasites qui peuvent endommager la culture, [59] propose un simulateur pour la gestion d'une flotte de drones. Il a présenté plusieurs techniques de coordination pour la surveillance et le contrôle.

### 1.3.2.2 Les drones pour la surveillance et la création de cartes de rendement

Les drones aujourd'hui remplacent la méthode classique d'observation des cultures en croissance qui consiste à marcher sur le champ de la ferme avec un cahier à la main et qui constitue l'un des obstacles d'augmentation de la productivité et d'amélioration de la qualité. Comme nous l'avons dit précédemment, les drones peuvent couvrir jusqu'à des centaines d'hectares en un seul vol, et acquérir des données qui ne sont pas percevables par l'œil humain en utilisant par exemple NDVI. Ils permettent d'assurer une visualisation et une collection régulière d'informations ce qui donne la possibilité de surveiller la croissance des cultures et de détecter la variabilité dans plusieurs paramètres du champ. Les données collectées sont utiles pour préparer des cartes qui devraient aider les agriculteurs à avoir une idée du rendement final prévu des cultures.

Les drones sont utilisés afin de déterminer la vigueur des cultures de maïs et aussi pour estimer les rendements à ses divers stades de croissance dans les petites exploitations [228]. Ils sont utilisés pour évaluer la qualité des rizières vietnamiennes [214, 132] et pour estimer leur rendement [66]. Leur utilisation tend également à d'autres cultures comme le blé [121] et la grenade [68]. En plus, ils aident à la détermination du moment optimal de la récolte en se basant sur la mesure de la maturation des plantes<sup>4</sup>.

### 1.3.2.3 Les drones pour la fertilisation et l'irrigation

Le manque d'un ou plusieurs nutriments essentiels comme l'eau et l'azote limite la production agricole [125]. Les drones constituent un support majeur dans l'évaluation des exigences de la

---

4. <https://www.pix4d.com/blog/drones-variable-rate-application>

variabilité spatiale des cultures. En effet, une carte de prescription spatiale est générée grâce aux images prises par drones. Cette carte permet de répondre aux besoins variables de la culture en nutriments dans différentes places du champ. L'application à taux variable d'engrais et de l'eau permet d'éviter la sur-fertilisation et la sur-irrigation, de gagner du temps et de réduire les coûts en terme de nutriments et l'impact environnemental (réduit le ruissellement des nutriments par exemple).

Plusieurs études ont été menées pour gérer d'une manière optimale ces intrants en se basant sur les drones. A titre d'exemples, [111] a proposé un système d'imagerie UAV multi-capteurs autonome pour fournir des informations spectrales sur la gestion de l'eau pour un verger de grenade. Les indices de stress hydrique et de la végétation sont obtenus à partir des données spectrales, thermiques et multi-spectrales collectées. [175] a présenté un schéma pour les estimations de la teneur en azote et de la biomasse à partir d'images hyper-spectrales. Les vols ont été menés sur le terrain d'essai finlandais au cours des étés 2011 et 2012.

#### **1.3.2.4 Les drones pour la reconnaissance et la classification**

La reconnaissance et la classification des plantes à l'aide de drones peuvent être une étape importante dans l'agriculture intelligente. La distinction entre la culture et la mauvaise herbe et entre la culture et le sol est essentielle pour accomplir certaines tâches comme la gestion spécifique au site et l'estimation du rendement [70]. En plus, la classification en différents types de cultures peut aider à les traiter différemment (par exemple en optimisant l'utilisation des engrais, eau et herbicides) [158]. Les images prises par drones peuvent aussi être segmentées afin de détecter les maladies des cultures [166], d'estimer l'état physiologique des plantes [195] et de classer et cueillir des fruits [1].

## **1.4 Conclusion**

Le secteur agricole connaît actuellement un afflux de technologies de l'information et de la communication, principalement basé sur le système de positionnement mondial, les avancées technologiques dans les capteurs, la robotique et les outils et les techniques d'analyse de données, etc. A travers ce chapitre nous avons donné un aperçu sur les technologies adoptées dans l'agriculture intelligente tout en mettant en évidence l'exploration de données et les drones.

Compte tenu de l'importance d'analyse de données dans la gestion des opérations agricoles, le chapitre suivant sera consacré à l'exploration de données agricoles, dans lequel nous discuterons les différentes applications des techniques de Data Mining ainsi que les défis rencontrés.

# Chapitre 2

## Data Mining dans l'agriculture intelligente

### 2.1 Introduction

Les changements apportés par les nouvelles technologies, en particulier les technologies d'acquisition de données, de traitement de données et de prise de décisions dans le secteur agricole ont été extrêmement importants pour une agriculture durable. Les technologies de détection et d'acquisition de données ont permis d'acquérir de grandes quantités de données auprès de fermes, de petites et de grandes exploitations agricoles. L'extraction d'informations utiles et précieuses dans un grand volume de données nécessite l'introduction de méthodes d'analyse et de traitement des données afin d'aider l'agriculteur à prendre les bonnes décisions et à faire de bonnes prédictions. Le Data Mining devrait jouer un rôle important dans l'analyse de ces grandes masses de données. Il implique des outils et des méthodes issus de différents domaines de l'informatique, des statistiques et de l'intelligence artificielle. Dans ce chapitre, nous présentons et discutons les applications de Data Mining dans l'agriculture intelligente, ainsi que les défis auxquels le Data Mining est confronté.

### 2.2 Les différentes tâches du Data Mining

Au fil du temps, les agriculteurs doivent produire plus de nourriture en utilisant moins de ressources. Ils doivent consommer moins d'eau et utiliser moins de produits chimiques. Les agriculteurs veulent maximiser leur productivité et réduire leurs coûts, tandis que les consommateurs exigent des aliments sains. En conséquence, l'industrie agricole est à la recherche de nouveaux produits, pratiques et technologies. Ces différentes exigences peuvent être satisfaites par l'agriculture intelligente.

L'agriculture intelligente dépend d'une série de technologies qui fonctionnent ensemble pour permettre la collecte et l'analyse de données. Ensemble, ces nouvelles technologies génèrent d'énormes quantités de données. La Startup Farmobile a répertorié 22 catégories de données de machines agricoles et 21 catégories de données agronomiques saisonnières sur son site Web (Farmobile) [173]. L'agriculture intelligente diffère de l'agriculture conventionnelle en se concentrant sur la collecte et l'utilisation de données pour prendre des décisions.

L'exploitation d'une grande quantité de données sur les cultures, les sols et le climat et l'analyse

de nouvelles données non expérimentales optimisent la production et rendent l'agriculture plus résiliente au changement climatique. En même temps, les données sont rapidement accumulées dans des ensembles trop volumineux et complexes pour être analysées sans logiciels adéquats. Les données seules ne génèrent pas d'informations. Des analyses sont nécessaires pour aider les agriculteurs à utiliser toutes ces données. Les applications basées sur Data Mining deviennent de plus en plus intelligentes. Le Data Mining permet de remplir diverses tâches, notamment la classification, le clustering, la recherche d'associations et la prédiction (voir FIGURE 2.1).

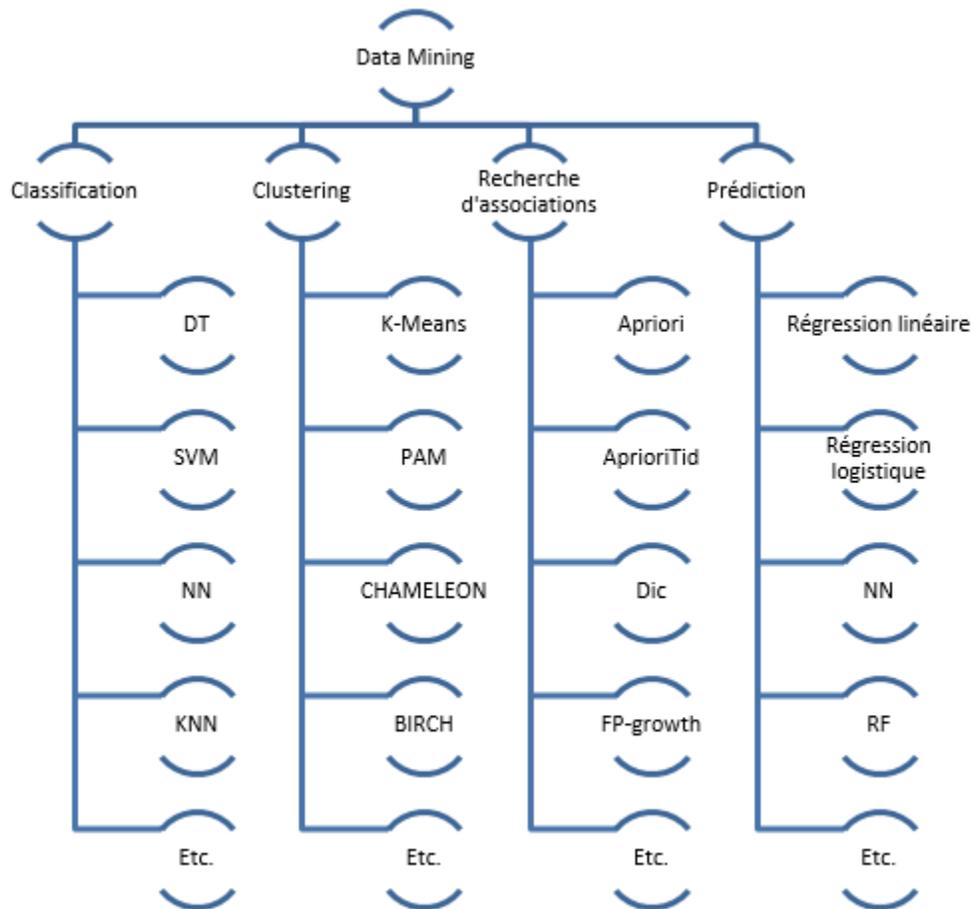


FIGURE 2.1 Les différentes tâches du Data Mining

### 2.2.1 La classification

La classification est un processus d'apprentissage supervisé qui permet de prédire une étiquette de classe à partir d'un ensemble de données de formation [252]. Elle consiste à mapper chaque élément des données sélectionnées dans l'une des classes prédéfinies. Dans le domaine agricole, la classification a de larges applications, notamment dans la classification des maladies des cultures (feuillage, fruits).

Parmi les techniques de classification les plus utilisées, on peut citer : les arbres de décision (Decision Trees (DT)), la machine à vecteurs de support (Support Vector Machine (SVM)), les réseaux de neurones (Neural Networks (NN)), les forêts aléatoires (Random Forests (RF)), la méthode des k plus proches voisins (K-Nearest Neighbors (KNN)), le modèle de Markov (Markov Model (MM)) et récemment l'apprentissage en profondeur.

Les arbres de décision, également appelés arbres de classification et de régression [6], sont classés comme des techniques prédictives et descriptives. Ils fournissent un modèle sous la forme d'une structure arborescente [252]. Les règles de classification sont basées sur des tests organisés dans l'arborescence. Les arbres de décision sont faciles à assimiler et à comprendre grâce à une présentation visuelle. Parmi les algorithmes utilisés pour les construire, nous citons : ID3 (Iterative Dichotomiser 3) [182], C4.5 [183] et CART (Classification & Regression Tree) [40].

La méthode SVM proposée par Vladimir Vapnik est applicable aux problèmes de régression et de classification [161]. SVM produit des fonctions de mappage d'entrée - sortie à partir d'un ensemble de données d'apprentissage étiqueté. La classification multiclasse est également possible et consiste essentiellement en plusieurs SVM de deux classes pour résoudre le problème (en utilisant un contre tous ou un contre un). Les SVM ont montré de très bonnes performances dans de nombreuses applications du monde réel, comme le diagnostic médical, la bioinformatique, la reconnaissance faciale, le traitement d'image et récemment dans l'agriculture intelligente.

Les réseaux de neurones sont l'une des techniques d'apprentissage automatique qui ont été largement utilisées. Les réseaux de neurones artificiels tentent d'imiter la structure de calcul des systèmes nerveux animaux pour résoudre différentes tâches [161]. Ces derniers trouvent une large application dans le secteur agricole, en particulier pour la prédiction et la classification des maladies et ravageurs.

La forêt aléatoire est considérée comme l'une des méthodes utilisées pour la classification et la régression [174]. Elle convient pour classer de grandes quantités de données. Elle possède une collection d'arbres structurés. Les arbres dans la forêt aléatoire votent pour la catégorie la plus populaire.

Le KNN permet de classer une nouvelle donnée d'entrée, en fonction de K données d'entraînement les plus proches, où la valeur de K est définie par l'utilisateur [252]. Ainsi, la classe la plus votée parmi les K voisins les plus proches de l'ensemble de données d'apprentissage est attribuée à la nouvelle donnée d'entrée.

Le modèle de Markov est un processus stochastique dans lequel l'état suivant du système dépend uniquement de l'état actuel. Le modèle de Markov caché (HMM : Hidden Markov Model) est une extension du modèle de Markov [106] et est considéré comme le réseau bayésien dynamique le plus simple [6]. Il a été utilisé dans plusieurs domaines, en particulier dans la reconnaissance vocale.

Le deep learning est un sous-domaine de l'apprentissage automatique (machine learning) basé sur l'apprentissage de plusieurs niveaux de représentation. Il utilise des algorithmes pour modéliser les abstractions de haut niveau dans les données [109]. Il utilise généralement des réseaux

de neurones artificiels avec des couches cachées. Le deep learning a récemment suscité beaucoup d'intérêt, en particulier dans le domaine agricole où il a été appliqué pour effectuer plusieurs tâches, notamment la détection des maladies des cultures et l'identification des mauvaises herbes [10].

### 2.2.2 Le clustering

Contrairement à la classification, le clustering est un processus d'apprentissage non supervisé où les classes ne sont pas connues à l'avance. Il consiste à diviser les objets en groupes (clusters) à partir des informations trouvées dans les données qui décrivent ces objets et leurs relations. Les objets appartenant au même cluster doivent présenter une similitude entre eux tout en étant aussi différents que possible des objets associés à d'autres clusters. Les techniques de clustering sont classées selon plusieurs aspects en méthodes de partitionnement, méthodes hiérarchiques, méthodes basées sur la densité, méthodes basées sur une grille et méthodes basées sur un modèle.

Dans les méthodes de partitionnement, un certain nombre de clusters est formé à partir des données. Certains des algorithmes bien connus de cette classe sont k-Means [91] et PAM (Partitioning Around Medoids) [128]. K-Means est l'une des techniques les plus utilisées qui vise à segmenter en K clusters (groupes) une collection de n objets. Elle est utilisée dans le domaine agricole pour remplir plusieurs tâches telles que la segmentation d'images pour détecter les maladies.

Les méthodes hiérarchiques créent une hiérarchie de clusters appelée dendrogramme. Cette catégorie est divisée en méthodes divisives ou agglomératives [107]. Les algorithmes de cette catégorie englobent CHAMELEON [110] et BIRCH (Balanced Iterative Reducing and Clustering Using Hierarchies) [255].

Les méthodes basées sur la densité regroupent les objets selon des fonctions d'objectif de densité spécifiques [107]. Certains algorithmes de cette catégorie sont DBSCAN (Density-Based Spatial Clustering of Applications with Noise) [73], DENCLUE (DENSITY CLUSTERING) [92] et OPTICS (Ordering Points To Identify the Clustering Structure) [21].

Les méthodes basées sur une grille divisent l'espace en un nombre fini de cellules qui forment une structure de grille sur laquelle toutes les opérations de regroupement sont effectuées. Certains algorithmes de cette catégorie incluent STING (STATISTICAL INFORMATION GRID APPROACH) [233] et CLIQUE (CLUSTERING IN QUEST) [7].

Les méthodes basées sur des modèles font référence à l'utilisation de modèles de mélanges finis pour effectuer des regroupements [32, 155]. COBWEB [79] et CLASSIT [82] sont des exemples d'algorithmes utilisés dans cette classe.

### 2.2.3 La recherche d'associations

Elle est utilisée pour explorer des relations significatives entre des objets appartenant à une grande quantité de données. L'analyse des règles d'association consiste à déterminer les associations ou les modèles entre les éléments de l'ensemble de données en cherchant à identifier les groupes d'éléments les plus courants qui se produisent ensemble [252]. Parmi les algorithmes

d'exploration de règles d'association existants figurent Apriori [6], AprioriTid [8], DIC (Dynamic Itemset Counting) [41], Eclat (Equivalence Class Transformation) [249]) et FP-growth (Frequent Pattern growth) [89].

## 2.2.4 La prédiction

Elle consiste à découvrir des modèles de données pouvant conduire à des prédictions raisonnables sur l'avenir. Elle est également connue sous le nom d'analyse prédictive.

L'analyse des séries chronologiques et l'analyse de régression sont parmi les techniques de prédiction les plus utilisées. Une série chronologique est définie comme une séquence d'observation ordonnée dans le temps [37]. Elle est utilisée pour découvrir des informations et des connaissances cachées. L'analyse des séries chronologiques est la méthode utilisée pour analyser les données statistiques dans une période de temps donnée. La prévision des séries chronologiques est un modèle pour prédire les valeurs futures sur la base des valeurs précédemment observées.

L'analyse de régression est l'une des techniques utilisées pour l'analyse prédictive. Elle est utilisée pour découvrir et analyser la relation d'une variable avec une ou plusieurs variables. La régression linéaire est une méthode statistique utilisée pour prédire la valeur d'une variable dépendante à partir d'une variable indépendante, où la relation entre les variables peut être représentée avec un modèle linéaire. Ainsi, elle corrèle le changement d'une variable dépendante à une variable indépendante.

En plus de ces techniques, il existe d'autres techniques utilisées pour accomplir la tâche de prédiction telles que les réseaux de neurones [161], SVM [161], l'arbre de décision [6]) et les forêts aléatoires [174]. Les différentes applications de la prédiction en agriculture comprennent la prédiction du rendement des cultures, la prédiction des maladies, la prédiction des ravageurs et la prédiction de l'humidité du sol.

## 2.3 Data Mining et l'agriculture intelligente

L'exploration de données dans l'agriculture intelligente est principalement utilisée pour planifier l'utilisation des sols et de l'eau, surveiller la santé des cultures, réduire et optimiser l'utilisation des ressources naturelles, limiter l'utilisation de polluants (pesticides, herbicides, etc.), améliorer la qualité de la production, etc. Cette section présente les méthodes d'exploration de données utilisées pour résoudre divers problèmes agricoles.

### 2.3.1 Contrôle de l'irrigation

L'eau est un facteur essentiel pour la croissance des plantes. L'application d'une quantité d'eau élevée ou insuffisante a une influence négative sur sa croissance. En effet, une irrigation inadéquate ou mal conçue peut être source de nombreux problèmes. La sous-irrigation augmente le risque de salinisation. La sur-irrigation peut être une source de propagation d'agents pathogènes (pseudomonas, kystes d'amibes et larves d'anguilles et d'œufs de parasites, polluants (résidus de médicaments,

biocides, etc.)) dans les cultures. De nombreux systèmes d'irrigation intelligents basés sur Data Mining ont été développés afin de définir les besoins des cultures en fonction du climat et des cycles végétatifs. Data Mining joue un rôle important dans l'amélioration de la gestion de l'irrigation afin d'évaluer la consommation d'eau en utilisant des méthodes faisant intervenir des éléments climatiques, des facteurs de culture et des objectifs économiques.

Afin de surveiller et de contrôler avec précision les besoins en eau d'un champ, un système d'irrigation automatisé est proposé dans [164]. Ce système fonctionne en installant des capteurs dans le champ pour surveiller différents paramètres, à savoir : l'humidité, la température du sol et le type de sol (PH). L'algorithme Naive Bayes est appliqué pour l'estimation de la quantité précise d'eau en fonction des besoins de la culture. Cette technique prend en compte les prévisions météorologiques pour réguler automatiquement la quantité d'eau fournie aux cultures et suggère également les engrais nécessaires. Le système d'irrigation intelligent proposé garantit une utilisation judicieuse de l'eau.

[171] a développé un modèle pour prédire les périodes d'irrigation. Il est basé sur l'arbre de décision qui est combiné avec un algorithme génétique (NSGA II) [61]. Ce dernier est utilisé afin d'optimiser les différents paramètres de l'arbre de décision. Différentes données sont utilisées, notamment des données climatiques, des données sur le type de culture, des données phénologiques sur l'état des plantes et d'autres données. Les résultats ont montré que leurs modèles donnent des résultats prometteurs.

Xie et al. [238] ont proposé un système d'irrigation comprenant les composants suivants : Estimation de la demande d'irrigation, estimant la quantité d'énergie et le temps nécessaires à l'exécution de la prochaine opération, Prévision de l'énergie solaire pour prévoir l'énergie solaire à l'aide de la méthode de régression vectorielle SVR (Support Vector Regression), et Optimisation de la planification à la demande afin de réduire les coûts d'irrigation. Le système utilise la prévision numérique du temps (NWP : Numerical Weather Prediction) et le modèle de prix TOU (Time Of Use) de l'électricité. Le résultat a montré que les ressources en énergie et en eau peuvent être économisées de 7,97% et que le coût peut être réduit de 25,34% par rapport à la méthode d'irrigation basée sur l'humidité du sol.

Kokkonis et al. [119] ont proposé un algorithme de réseau de neurones flou pour gérer l'irrigation appelé FITRA. L'algorithme contrôle le système d'irrigation en fonction des données des capteurs. Plusieurs capteurs d'humidité du sol sont utilisés pour minimiser la consommation et augmenter la production. Le système est capable de s'adapter automatiquement aux conditions environnementales changeantes.

[86] décrit un système d'aide à la décision pour une irrigation exacte afin de maximiser le rendement des cultures et d'économiser de l'eau. Ce système est basé sur un réseau de capteurs / actionneurs sans fil (WSAN : Wireless Sensor and Actuator Network) pour mettre en œuvre une irrigation précise et autonome dans une serre. Le système développé fournit une surveillance et un contrôle en temps réel des plantes pour une irrigation de précision basée sur les techniques de Data Mining. Différents capteurs sont utilisés, notamment des capteurs d'humidité du sol, d'humidité de l'air et de température. Les résultats des tests dans le champ de fraises ont montré une réduction

significative de la consommation d'eau (20%) par rapport à l'irrigation traditionnelle.

Le système proposé dans [84] a pour objectif de prédire les recommandations d'irrigation en appliquant des techniques d'exploration de données. Plusieurs modèles de prédiction ont été appliqués en utilisant des algorithmes de régression et de classification. Différentes données, y compris des données météorologiques et d'humidité collectées par des capteurs, ont été utilisées pour assurer la surveillance, le contrôle et la prédiction en temps réel des besoins des plantes. Les modèles Gradient Boosted Regression Trees (GBRT) et Boosted Trees Classifier (BTC) ont donné de meilleures performances par rapport aux modèles de régression linéaire.

Le Tableau 2.1 résume ces articles.

**Tableau 2.1** – Un résumé des articles traitant de la tâche d'irrigation

Référence	Objectif	Techniques	Données utilisées	Prétraitement des données
[164]	Estimation de la quantité précise d'eau et suggestion des engrais nécessaires	L' algorithme Naive Bayes	Données de capteurs (humidité, température du sol et type de sol) et données des sites de prévisions météorologiques	-
[171]	Prédiction des périodes d'irrigation	Arbres de décision, algorithme génétique	Données sur les cultures, les jours juliens, les jours fériés, les jours de la semaine et les conditions climatiques (température, humidité, précipitations)	-
[238]	Minimisation des coûts d'irrigation	La méthode SVR, un algorithme d'estimation d'irrigation et un modèle d'optimisation	Informations sur l'humidité du sol et l'irradiance sans nuages, informations météorologiques (couverture nuageuse, humidité, précipitations ...) et données sur l'énergie solaire.	-
[119]	Gestion de l'irrigation	L'algorithme FITRA	Capteurs d'humidité du sol	Suppression des mesures erronées
[86]	Gestion de l'irrigation spécifique à la zone	Algorithmes de DM	Données de capteurs (humidité du sol, humidité de l'air, température) et autres données de la plante	Réduction de données
[84]	La prédiction du plan d'irrigation hebdomadaire	GBRT, BTC et des modèles de régression linéaire	Données des stations météorologiques, enregistrements d'irrigation actuelle et historique des données des capteurs	Nettoyage des données et transformation en fonctionnalités requises

### 2.3.2 Surveillance des maladies des plantes (Détection et prédiction des maladies)

La plante est affectée par plusieurs maladies au cours de sa croissance. Leur détection était l'objectif de nombreuses recherches. Elles reposent sur la combinaison de techniques de Data Mining et de traitement d'images pour pallier le manque d'observation humaine et même pour réduire les coûts [165].

Barbedo et Garcia [29] étudient l'exploitation de techniques courantes de traitement d'images numériques combinées à l'exploration de données dans le secteur agricole pour détecter, classer et quantifier les maladies des plantes.

Dans [199], un système de détection robotique des maladies est présenté. Il est basé sur un manipulateur permettant d'atteindre différentes poses de détection. Plusieurs algorithmes basés sur l'analyse en composantes principales (PCA : Principal Components Analysis) et le coefficient de variation sont utilisés afin de détecter le virus de la maladie bronzée de la tomate et l'oïdium qui menacent les poivrons de serre.

Baquero et al. [28] ont proposé une stratégie pour la récupération d'image des feuilles de tomate afin de renforcer le diagnostic de la maladie. Cette stratégie utilise le descripteur de structure de couleur, qui capture les informations spatiales et de couleur, ainsi que l'algorithme du plus proche voisin pour la tâche de récupération. Dans ce travail, six maladies de la tomate ont été examinées : la brûlure alternarienne, la chlorose, les moisissures fuligineuses, l'oïdium, la nécrose et les mouches blanches. Sur la base des résultats expérimentaux, l'approche proposée peut partiellement fournir des informations sur diverses maladies telles que la chlorose, la brûlure alternarienne et les moisissures fuligineuses. D'autre part, les performances sont faibles pour l'oïdium et les mouches blanches.

Dans [151], un système intelligent basé sur des capteurs multiples est développé afin de détecter les maladies ou le stress des plantes. Deux classificateurs sont utilisés : les réseaux de neurones et la discrimination quadratique (QDA : Quadratic Discriminant Analysis). Le système utilise des informations hyper-spectrales et multi-spectrales, ainsi que le système de positionnement géostationnaire. Une stratégie de pulvérisation spécifique au site est également proposée pour prévenir les maladies fongiques.

Dhau et al. [62] ont utilisé le classificateur forêt aléatoire avec des données satellites RapidEye pour cartographier et détecter le virus de la striure du maïs en Afrique du Sud. Pour compléter les données de télédétection, des données de terrain sont également collectées. Les résultats ont montré que la précision de la classification est améliorée grâce à l'intégration des indices de végétation. L'étude a montré que l'indice de végétation ajusté au sol, l'indice rouge, l'indice de végétation amélioré et l'indice de végétation normalisé sont les facteurs les plus importants pour la détection du virus de la striure du maïs.

Liu et al. [136] ont proposé une méthode bayésienne pour déduire la probabilité de stress dû au cadmium (Cd) du riz en utilisant des caractéristiques spatio-temporelles et des données satellites.

La probabilité de stress dû au Cd est vérifiée à l'aide de ses concentrations dans le sol. Les résultats montrent qu'une précision globale de 81,57% est obtenue pour le stress dû au Cd dans le riz.

Les avantages de l'automatisation de la détection des maladies font du traitement d'images un sujet d'étude plus intéressant. La plupart des articles traitant de la détection et de la classification des maladies [196, 205, 165, 234, 112, 31, 98, 169, 190, 167, 197, 243, 99, 22, 67] font la combinaison entre le traitement d'images et les techniques de Data Mining telles que SVM, ANN (Artificial Neural Network), etc.

Le Tableau 2.2 résume les articles traitant de la détection et de la classification des maladies à l'aide des méthodes de Data Mining.

**Tableau 2.2 – Résumé de certains articles traitant de la détection et de la classification des maladies à l'aide des méthodes de Data Mining**

Référence	Maladies traitées	Méthodes	Étapes de traitement	Précision
[196]	La tavelure de la pomme, la pourriture de la pomme, la tache de la pomme	Traitement d'image + RF	<ul style="list-style-type: none"> <li>Extraction des caractéristiques de texture et de couleur.</li> <li>La fusion de deux ou plus de deux de ces caractéristiques.</li> <li>La classification en utilisant le classificateur forêt aléatoire.</li> <li>Pour le fruit infecté la technique K-means est utilisée pour segmenter la partie infectée.</li> </ul>	-
[165]	Maladies de la feuille de vigne : oïdium, mildiou	Traitement d'image + SVM	<ul style="list-style-type: none"> <li>L'image d'entrée est prétraitée (c.à.d. le redimensionnement, le seuillage et le filtre gaussien sont appliqués).</li> <li>Elle est segmentée en trois clusters en utilisant la technique de clustering K-means.</li> <li>Extraction des caractéristiques de texture et de couleur et les donner au classificateur SVM.</li> </ul>	88.99%
[234]	Maladies des feuilles de coton : tache rouge, la cercosporiose, tache alternarienne	Traitement d'image + NN	<ul style="list-style-type: none"> <li>Prétraitement de l'image (Égalisation des histogrammes, redimensionnement.)</li> <li>Segmentation avec k-means.</li> <li>Extraction des caractéristiques.</li> <li>La classification avec NN.</li> </ul>	89.56%
[67]	Pourriture de pomme, tavelure de pomme, tache de pomme	Traitement d'image + SVM	<ul style="list-style-type: none"> <li>La segmentation en utilisant K-means.</li> <li>Extraction des caractéristiques (CCV (Color Coherence vector), GCH (Global Color Histogram), LBP (Local Binary Pattern), CLBP (Complete Local Binary Pattern)) qui seront utilisées par le classificateur SVM.</li> </ul>	93%

## CHAPITRE 2. DATA MINING DANS L'AGRICULTURE INTELLIGENTE

[205]	Les maladies des feuilles sur différentes cultures	Traitement d'image+ GA	-	<ul style="list-style-type: none"> <li>• Prétraitement de l'image.</li> <li>• La segmentation en utilisant un algorithme génétique.</li> </ul>
[98]	3 maladies du raisin et 2 maladies de la pomme	Traitement d'image + ANN	-	<ul style="list-style-type: none"> <li>• Extraction des caractéristiques (forme, couleur, texture).</li> <li>• La classification à l'aide d'ANN.</li> </ul>
[169]	L'oïdium et le mildiou du concombre	Traitement d'image + ANN	80.45%	<ul style="list-style-type: none"> <li>• Les caractéristiques de texture sont extraites en utilisant les moments statistiques de premier ordre et de second ordre (GLCM : Grey Level Co-occurrence Matrix).</li> <li>• Le réseau de neurones est utilisé pour la classification des maladies en utilisant comme entrée les caractéristiques extraites.</li> </ul>
[31]	Telya (brûlure bactérienne) de grenade	Traitement d'image + SVM	82%	<ul style="list-style-type: none"> <li>• Prétraitement de l'image (redimensionnée).</li> <li>• Extraction des caractéristiques (couleur, morphologie, CCV).</li> <li>• La classification avec SVM .</li> </ul>
[190]	Maladies des taches de la feuille de coton : verticilliose, fusariose, brûlure des feuilles.	Traitement d'image + NN	98.1%	<ul style="list-style-type: none"> <li>• Après le prétraitement de l'image, la segmentation est faite en utilisant Canny et Sobel.</li> <li>• La classification est réalisée avec le réseau de neurones.</li> </ul>
[167]	Tache sur les feuilles de coton	Traitement d'image + SVM	89% [234]	<ul style="list-style-type: none"> <li>• Conversion en une image en niveaux de gris, segmentation avec le seuil de niveau de gris et aussi application du filtre médian et d'opérateurs morphologiques pour éliminer les taches inutiles.</li> <li>• Extraction des caractéristiques (forme, couleur).</li> <li>• Analyse statistique (pour le choix de bons paramètres).</li> <li>• La classification avec SVM.</li> </ul>
[112]	La pomme de terre tardive (feuille)	Traitement d'image + ANN	100%	<ul style="list-style-type: none"> <li>• Après la normalisation et l'amélioration du contraste, Ils ont appliqué l'étirement de la décorrélation</li> <li>• Conversion en espace LAB.</li> <li>• Définition du vecteur caractéristique qui sera utilisé pour le clustering en utilisant Fluffy C-mean (FCM).</li> <li>• Conversion des clusters en une image HSV (Hue Saturation Value).</li> <li>• Extraction des caractéristiques de gradient et de texture et les utiliser pour former ANN.</li> </ul>

[197]	Maladies de la feuille de vigne : mildiou, mildiou	Traitement d'image + BPNN (Back Propagation Neural Network)	<ul style="list-style-type: none"> <li>• L'enlèvement de fond (le seuillage / maskage)</li> <li>• Prétraitement de l'image (par 5 itérations de la diffusion anisotrope).</li> <li>• La segmentation avec K-means.</li> <li>• Extraction des lésions.</li> <li>• Extraction des caractéristiques de texture de la partie malade extraite avec GLCM.</li> <li>• La classification à l'aide de BPNN.</li> </ul>	100%
[243]	Brûlure de la gaine du riz, Brûlure bactérienne des feuilles du riz, souffle du riz	Traitement d'image + SVM	<ul style="list-style-type: none"> <li>• Après l'acquisition de l'image dans la rizière avec une caméra elle est prétraitée et segmentée, ensuite ses caractéristiques de forme et de texture sont extraites.</li> <li>• En se basant sur ces caractéristiques, la classification est réalisée avec SVM.</li> </ul>	97.2%
[99]	Maladies du concombre : tache brune, mildiou, tache angulaire de la feuille	Traitement d'image + SVM	<ul style="list-style-type: none"> <li>• Prétraitement de l'image (Avant d'appliquer la méthode SVM, une série de prétraitement d'image est faite (technologie de coupe d'image, suppression des bruits, méthode de seuillage))</li> <li>• Extraction des caractéristiques (forme, couleur, texture).</li> <li>• La classification à l'aide de SVM.</li> </ul>	-
[22]	Maladies des feuilles sur différentes cultures	Traitement d'image + SVM	<ul style="list-style-type: none"> <li>• Prétraitement de l'image (la conversion de RGB à HSI (Hue Saturation Intensity), le maskage des pixels verts, la suppression des pixels verts masqués).</li> <li>• Segmentation.</li> <li>• Extraction des caractéristiques de texture (contraste, énergie, homogénéité locale, la visibilité des grappes) à partir desquelles les maladies sont classées en différents types.</li> <li>• La classification avec SVM.</li> </ul>	94%

D'autres articles récents portent sur l'utilisation de l'apprentissage en profondeur. Y. Lu et al. [140] ont proposé un modèle basé sur les réseaux de neurones convolutionnels profonds (CNN : Convolutional Neural Networks) afin d'identifier dix maladies du riz, notamment le faux-charbon du riz, le souffle du riz, la tache brune du riz, la brûlure de la gaine du riz, la maladie de bakanae, la brûlure bactérienne des feuilles du riz, la pourriture de la gaine du riz, la pourriture bactérienne de la gaine du riz, le flétrissement bactérien du riz et la brûlure des semis de riz. Les images sont prétraitées en utilisant la méthode PCA et les méthodes de blanchiment afin d'obtenir les caractéristiques de formation et de test. Leur modèle a montré une précision de 95%.

Ferentinos [77] a développé des modèles de réseaux de neurones convolutionnels profonds afin d'identifier les maladies des plantes à partir d'images de feuilles. Ces modèles sont AlexNet [123], AlexNetOWTBn [122], GoogleNet [210], Overfeat [201] et VGG [204]. Plusieurs cultures et maladies ont été intégrées dans le modèle, mais leur système reste applicable pour 25 plantes dont on

trouve globalement 58 classes différentes, y compris les plantes en bonne santé et les maladies des plantes (plante-maladie).

Durmuş et al. [69] ont utilisé des modèles CNN pour détecter des maladies sur la base d'images de feuilles de tomates. Les modèles utilisés sont AlexNet [123] et SqueezeNet [93]. Leur système est conçu pour être utilisé dans les appareils mobiles. Les deux modèles ont montré des résultats comparables.

Kawasaki et al.[113] ont mis au point un système basé sur CNN pour la détection des maladies des feuilles du concombre, le virus de la tache jaune du melon et le virus de la mosaïque jaune de la courgette. Ils ont utilisé deux stratégies de prétraitement qui sont "square crop" et "square deformation crop". Ces stratégies permettent de se focaliser sur la région d'intérêt et de créer d'autres exemples d'entraînement. Leur système a montré une précision de 94,9%.

Amara et al. [18] ont utilisé le modèle LeNet [126] pour détecter deux maladies des feuilles du bananier (tache de banane et sigatoka de banane). Les étapes de prétraitement des images sont utilisées, notamment le redimensionnement des images à 60x60 pixels et la conversion en images en niveaux de gris. Ils ont validé le modèle proposé en utilisant des images de feuilles de bananier qui se trouve sur plant Village. Des résultats intéressants ont été montrés dans leurs expériences.

Brahimi et al.[39] ont proposé des modèles CNN (AlexNet, GoogleNet) pour classer neuf maladies de la tomate. Ils ont utilisé un ensemble de données de 14828 images. Une méthode de visualisation est également utilisée pour visualiser et détecter les symptômes. Les expériences ont montré que les résultats des modèles CNN (avec ou sans formation préalable) sont meilleurs que ceux de SVM et de RF.

Jiang et al. [100] ont proposé un modèle basé sur l'apprentissage en profondeur afin de détecter cinq maladies de la feuille du pommier en temps réel. Ils ont utilisé 26377 images collectées sur le terrain et en laboratoire. Les résultats ont montré que la performance du modèle proposé INAR-SSD atteignait 78,80% mAP (mean Average Precision) avec une vitesse de détection de 23,13 FPS (Feet Per Second).

La détection de la plante (ou d'un organe) pourrait être utile en tant qu'étape préliminaire pour la détection de la maladie. Yahata et al. [240] présentent une méthode de détection d'images pour détecter les fleurs et les gousses dans des environnements extérieurs à Hokkaido, au Japon. Une approche hybride pour détecter le phénotypage de végétation est adoptée, utilisant le groupement itératif linéaire simple (SLIC : Simple Linear Iterative Clustering), la méthode de détection d'objet Viola-Jones, les caractéristiques issues de tests accélérés de segments (FAST : Features from Accelerated Segment Test) et le réseau de neurones convolutifs. Le comptage automatique des fleurs et des gousses a permis d'observer avec précision le processus de croissance du soja. Des résultats prometteurs ont été montrés dans lesquels la précision de la détection des fleurs atteignait 0,9 et le taux de rappel de 0,826.

Le Tableau 2.3 résume les articles basés sur l'apprentissage en profondeur.

**Tableau 2.3** – Un résumé d'articles traitant de la détection et de la classification des maladies à l'aide de l'apprentissage en profondeur

Référence	Maladies traitées	Modèles utilisés	Ensemble de données	Précision
[69]	Maladies de la tomate (10 classes, y compris la classe saine)	AlexNet, SqueezeNet	feuilles de tomate (PlantVillage)	AlexNet 95.65% SqueezeNet 94.3%
[77]	Maladies de 25 plantes différentes (58 classes)	AlexNet, AlexNetOwtBn, GoogleNet, Overfeat, VGG	Base de données de 87848 images de feuilles	Avec les images originales : AlexNet 99.06% AlexNetOwtBn 99.49% GoogleNet 97.27% Overfeat 98.96% VGG 99.53%
[140]	10 maladies du riz	SVR et CNN comparés à d'autres techniques (SVM, PSO (Particle Swarm Optimization), BP)	Base de données de 500 images	95,48%
[39]	9 maladies de la tomate	AlexNet et GoogleNet comparés à d'autres techniques (SVM, RF)	Base de données de 14828 images de feuilles de tomate (PlantVillage)	Avec pré-formation : AlexNet 98.66% GoogleNet 99.185% SVM 94.5% RF 95.54%
[113]	2 maladies des feuilles de concombre	CNN	Un ensemble de données de 800 images de feuilles de concombre	94.9%
[18]	Maladies de la feuille de bananier (3 classes incluant la classe saine)	LeNet	3700 images de la feuille de bananier (PlantVillage)	Images en niveaux de gris >86% Images en couleur >92%
[100]	5 maladies de la feuille de pomme	INAR-SSD et d'autres modèles	26377 images de feuilles de pommier	78.80%

Outre que les images, d'autres chercheurs considèrent les paramètres environnementaux pour surveiller l'état de santé des plantes.

Patil et Thorat [168] ont utilisé le modèle de Markov caché pour la détection précoce des maladies du raisin. Dans leur travail, le modèle analyse les données d'entrée (température, humidité relative, humidité de la feuille) et classe les maladies du raisin (anthracnose, oïdium, rouille, mildiou, tache bactérienne et cancer de la feuille bactérien). Les résultats ont montré que l'utilisation de HMM est très précise par rapport à la méthode statistique utilisée pour la détection.

Dans [231], la méthode SVM a été utilisée pour prédire la rouille jaune du blé en Chine. La prédiction est effectuée à l'aide de trois ensembles de données (données sur la maladie et données

météorologiques). Les auteurs présentent également une comparaison entre la méthode SVM et la méthode d'analyse par régression. Les résultats ont montré que la méthode SVM peut atteindre une plus grande précision de prédiction, et que l'application de cette méthode à la prédiction de la maladie est réalisable et efficace.

Dans [217], différentes techniques d'exploration de données (Gaussian Naive Bayes, Rapid Association Rule Mining et Multivariate Regression Mining (MVR)) ont été appliquées pour étudier les interactions culture-climat-maladie. Un réseau de capteurs sans fil est utilisé pour collecter des données de surveillance des cultures d'arachides pour la maladie des taches foliaires. Le modèle présenté a été validé et comparé au modèle ARI (Agriculture Research Institute).

### **2.3.3 Surveillance des organismes nuisibles et gestion optimisée des intrants (engrais et pesticides)**

Au cours du cycle de production d'une plante, des conditions anormales telles que la température et l'humidité de l'environnement permettent la propagation de plusieurs maladies causées par des insectes, champignons, mauvaises herbes, nématodes et rongeurs, etc., qui affectent la bonne croissance de la plante. Ces parasites font l'objet de certaines recherches.

Boniecki et al. [36] ont proposé un réseau de neurones pour identifier six parasites dans le verger de pommiers. Les six parasites sont le charançon de la fleur du pommier, le meunier de feuille de pomme, la teigne du pommier, le carpocapse de la pomme, l'aile claire du pommier et le puceron de la pomme. Le classificateur est basé sur 23 paramètres, dont 16 caractéristiques de couleur et 7 facteurs de forme. Ils ont obtenu de bons résultats en utilisant la topologie du réseau de neurones MLP (Multi-Layer Perceptron).

Dans [216], des techniques de Data Mining sont utilisées pour comprendre la relation entre la maladie BNV (Bud Necrosis Virus) et les données météorologiques et entre le parasite (thrips) et les données météorologiques dans les cultures d'arachides de l'Inde. Les méthodes Gaussian Naive Bayes et Rapid Association Rule Mining sont utilisées respectivement pour la classification, l'association et l'analyse de corrélation. Un modèle de régression multivariée est développé et utilisé afin de développer un modèle de prédiction empirique. Quelques cas d'utilisation de techniques de Data Mining pour la prédiction des parasites dans les cultures ont été décrits dans [116].

Dans [193], un système intelligent a été développé et appliqué à une étude de cas de la dynamique des populations d'organismes nuisibles. Il décrit un nouveau système d'aide à la décision pour l'agriculture, appelé AgroDSS. L'aspect décisionnel du système proposé repose sur des approches d'exploration de données, qui peuvent extraire des informations utiles à partir de données volumineuses. Les outils mis en œuvre comprennent l'apprentissage supervisé (modélisation prédictive), l'apprentissage non supervisé (clustering) et l'analyse de séries chronologiques. Les données sur les organismes nuisibles, collectées par TrapView, permettent de surveiller les organismes nuisibles dans les vignes et les vergers au moyen de pièges à insectes.

Dans [54], un paramètre important, les «degrés-jours accumulés (thermal integral over air temperature)», est calculé à l'aide de données sur la surface des terres obtenues par un satellite météo-

rologique afin de cartographier les risques liés aux parasites et aux maladies. La régression linéaire est utilisée pour traiter les données manquantes et la régression logistique pour analyser les résultats des données mensuelles (en degrés-jours) de la température de la surface de la terre. Une relation significative est montrée entre les valeurs accumulées calculées à partir des stations météorologiques et les valeurs accumulées obtenues par satellite.

Afin d'identifier les cultures et les mauvaises herbes, Lottes et al. [138] proposent un système qui détecte les plantes de betterave à sucre, en extrayant les caractéristiques pour obtenir une estimation précise de la culture et des mauvaises herbes en combinant la classification forêt aléatoire et en exploitant les informations voisines via un champ de Markov aléatoire (RMF : Random Markov Field). L'approche proposée a été mise en œuvre et évaluée sur un véritable robot de ferme sur trois champs de betteraves sucrières différents. En séparant automatiquement les deux catégories de plantes, un robot peut effectuer la pulvérisation nécessaire ou éliminer les mauvaises herbes.

Dans [244], différents modèles de réseau de neurones à convolution profonde sont utilisés pour détecter les mauvaises herbes. Ils sont basés sur des images de mauvaises herbes prises avec un appareil photo numérique. Ces modèles sont VGGNet, GoogleNet et DetectNet. Les résultats ont montré que le réseau de neurones à convolution profonde convient parfaitement à la détection des mauvaises herbes.

Bah et al. [24] ont proposé un réseau de neurones convolutifs associé à un algorithme appelé "crops lines algorithm" pour identifier les mauvaises herbes dans les champs de betteraves, d'épinards et d'haricots. Des images végétales à haute résolution prises à environ 20m par des drones sont utilisées. Les résultats ont montré que la meilleure précision est atteinte dans le champ de betteraves. Les auteurs mentionnent certaines difficultés à détecter les mauvaises herbes, en particulier au début de la croissance des plantes ou lorsque les mauvaises herbes poussent près de la culture.

H. Lu et al. [139] ont proposé des réseaux de neurones à convolution profonde et un transfert d'apprentissage pour extraire des informations sur les terres cultivées. La méthode proposée est comparée à une autre méthode appelée ECLE (eCognition for cultivated land information extraction) et leurs expériences ont montré des résultats comparables, à l'exception de l'intégrité et de la continuité des terres cultivées, la méthode proposée surpasse la seconde. Des images UAV capturées à haute altitude ont été utilisées.

Les ravageurs et les maladies des cultures étant fortement influencés par les paramètres météorologiques, un système d'aide à la décision en temps réel a été présenté dans [215] pour étudier et prédire le statut des ravageurs et des maladies, alerter tôt et suggérer les mesures de contrôle nécessaires. Diverses techniques de DM ont été utilisées en se basant sur des paramètres météorologiques et d'autres paramètres de surveillance des cultures d'arachides. Des modules de classification à plusieurs niveaux ont été développés pour classer les parasites et les maladies en fonction de leur gravité. La méthode Gaussian Naive Bayes a été utilisée dans l'expérience. Des techniques de classification ont été utilisées en association avec Rapid Association Rule Mining pour découvrir la corrélation de multiples paramètres météorologiques et autres paramètres connexes. Des modèles prédictifs ont été utilisés pour prendre les bonnes décisions au bon moment et protéger la culture, en particulier dans des conditions météorologiques très variables.

Dans l'agriculture intelligente, le développement de robots autonomes repose sur une technologie de base, à savoir des systèmes de vision robustes. Les systèmes développés doivent pouvoir surveiller les cultures et ne contrôler que des plantes spécifiques nécessitant un traitement. [38] a proposé une nouvelle méthode de détection et de classification des cultures et des mauvaises herbes basée sur la morphologie des attributs. Pour la discrimination culture / mauvaise herbe, la technique SVM a été utilisée. SVM classe les régions détectées en trois classes : cultures, mauvaises herbes et mélange. Les résultats montrent des taux de classification efficaces et compétitifs pour la discrimination des cultures et des mauvaises herbes appliquée à deux types de cultures, la betterave à sucre et l'oignon.

Pour lutter contre ces parasites et protéger les cultures, la protection chimique par l'utilisation de produits phytosanitaires (insecticides, fongicides, herbicides, nématicides, rodenticides, etc.) est préférée. Cependant, l'utilisation aléatoire et irrationnelle de ces derniers entraîne une toxicité en cas de surdosage et, dans le cas contraire, un déficit. Afin de réguler les quantités de produits phytosanitaires, des techniques de Data Mining sont envisagées. Cela implique de réaliser une pulvérisation de précision en développant des dispositifs de contrôle travaillant à partir d'informations obtenues en temps réel (capteurs embarqués) ou à partir d'informations a priori (images de télédétection). Ces dispositifs optimiseront les réglages pour améliorer le contrôle de la dose nécessaire en tenant compte des besoins de la culture et en assurant une meilleure complémentarité avec la fertilisation organique.

Dans [164], la méthode de classification Naive Bayes est utilisée pour suggérer la quantité d'engrais en fonction du type de sol et de la culture. L'article estime également la quantité précise d'eau en utilisant des capteurs tels que les capteurs de température, d'humidité et de pH. La prévision météorologique pour prédire la quantité de pluie est adoptée pour optimiser l'utilisation de l'eau.

Pour fournir la quantité et le type de nutriments nécessaires au bon moment, un système hydroponique intelligent a été développé dans [14] en combinant la culture hydroponique avec l'IoT. Il est utilisé dans l'automatisation du processus de croissance des cultures à l'aide d'un modèle d'analyse prédictive basé sur l'algorithme de prédiction réseau bayésien (Bayesian Network (BN)). Le système comprend trois composants principaux : un module d'analyse de données, des capteurs et une interface Web. Les capteurs contrôlent divers paramètres nécessaires au développement de la plante, à savoir la conductivité électrique, le niveau de pH, l'intensité lumineuse, l'humidité et la température de l'eau. Deux sites Web sont utilisés à des fins de visualisation et de contrôle. L'article a conclu que le rendement obtenu avec la commande automatique est meilleur que celui obtenu avec la commande manuelle en raison des différences de gain calculées entre 20% et 60% pour tous les paramètres utilisés pour évaluer des cultures de bonne qualité.

Pour encourager l'adoption de pratiques durables en agriculture et surmonter les limites de la productivité de l'agriculture biologique, D'Arpa et al. [55] sont axés sur l'application de Data Mining pour la gestion intégrée des informations environnementales sur les exploitations agricoles et l'utilisation intelligente des données spatiales. Ils se concentrent sur l'utilisation de K-Means et des réseaux de neurones pour regrouper les superficies cultivées, tout en illustrant le potentiel des algorithmes de DM pour définir la prise de décision typique des interventions de gestion et de planification en tenant compte de la nature physico-chimique du sol, du type de culture utilisée, de

la disponibilité et de la qualité de l'eau, ainsi que des aspects économiques et socio-économiques. Le Tableau 2.4 résume les articles traitant de la surveillance des parasites et de la gestion des intrants.

**Tableau 2.4** – Un résumé de certains articles traitant de la surveillance des parasites et de la gestion des intrants

Référence	Objectif	Techniques	Données utilisées	Prétraitement des données
[36]	Prédiction de 6 types de parasites dans un verger de pommiers	NN	Images numériques des parasites	-
[216]	Prédire les parasites / maladies	Gaussian Naive Bayes, Rapid Association Rule Mining et MVR	Température de l'air, température du sol, humidité relative, humidité des feuilles et autres données météorologiques	Il traite les données manquantes et la qualité des données (sélection et réduction des données, élimination des valeurs nulles et bruitées)
[193]	Prédiction des parasites (taille de la population et sa dynamique)	RF, Analyse des séries chronologiques	Données des parasites	Il traite les données manquantes (en utilisant l'algorithme KNN)
[54]	Détection et cartographie des risques de parasites et de maladies	Régression linéaire et logistique	Données satellitaires météorologiques (données de surface terrestre) et données météorologiques in situ	Il traite les données manquantes et bruitées (en utilisant la régression linéaire)
[138]	Détection des plantes de betteraves sucrières / mauvaises herbes	RF, RMF	Images de deux champs différents capturées par un robot	-
[244]	Détection des mauvaises herbes	CNN	Images prises à différents endroits	Redimensionnement des images
[24]	Identification des mauvaises herbes dans les betteraves, les épinards et les champs d'haricots	CNN combiné avec l'algorithme 'crops lines algorithm'	Images de légumes prises par des drones	Suppression du sol et de l'ombre (segmentation de l'arrière-plan)
[139]	Extraction d'informations sur les terres cultivées	CNN, transfert d'apprentissage et ECLE	Images d'UAV	Orthorectification, correspondance d'images et mosaïquage
[215]	Prédire le statut des ravageurs / maladies	Gaussian Naive Bayes, Rapid Association Rule Mining et MVR	Données météorologiques (température, humidité de l'air, humidité des feuilles, humidité du sol) et autres paramètres de surveillance	Il traite les données manquantes (en utilisant l'algorithme Expectation – Maximization (EM) [129]) et la qualité des données (sélection et réduction des données, élimination des valeurs nulles et bruitées)
[38]	Détection et classification des cultures et des mauvaises herbes	SVM	Images de plantes	-
[164]	Suggestion de la quantité d'engrais et contrôle de l'irrigation	L' algorithme Naive Bayes	Données de capteur (température, humidité et pH) et données de prévisions météorologiques	-
[14]	Gestion du processus de croissance des cultures d'une ferme hydroponique	BN	Données de capteurs (conductivité électrique, niveau de pH, intensité lumineuse, humidité et température de l'eau)	Omission des valeurs aberrantes
[55]	Regroupement de la surface cultivée	K-means, NN	Données spatiales, données spatio-temporelles	Normalisation des données

### 2.3.4 Prédiction du rendement des cultures et des effets du changement climatique sur la productivité

Une des préoccupations des agriculteurs est la prédiction des rendements des cultures ainsi que l'influence de certains paramètres environnementaux (climatiques) sur la productivité, qui influent principalement sur la politique des marchés. Le problème de prédiction de rendement rencontré peut être résolu à l'aide de Data Mining. Ainsi, beaucoup de recherches ont été faites dans ce sens.

Compte tenu du rôle essentiel que jouent une estimation précise du rendement et une gestion optimisée de l'azote dans l'agriculture de précision, [47] a fait une étude consacrée aux techniques de Data Mining permettant de prédire avec précision le rendement des cultures et d'estimer l'état de l'azote.

Veenadhari et al. [225] ont utilisé l'arbre de décision pour prédire l'influence des paramètres climatiques (humidité relative, évapotranspiration, température, précipitations) sur la productivité du soja. Les auteurs ont utilisé l'algorithme ID3, et ils ont conclu que l'humidité est le paramètre le plus déterminant pour le rendement, suivi de la température et des précipitations, et que l'arbre de décision a été exécuté rapidement.

Les mêmes auteurs ont présenté dans [226] un outil logiciel appelé «crop advisor» pour prédire l'influence des paramètres climatiques (fréquence des jours de pluie, évapotranspiration potentielle, température maximale et minimale, précipitations, couverture nuageuse) sur le rendement des cultures. Le paramètre climatique qui influence le rendement des cultures sélectionnées est trouvé en appliquant l'algorithme C4.5.

Dans [224], différents capteurs, notamment des capteurs de température et d'humidité et d'autres données invariantes, sont utilisés pour suggérer la culture appropriée et effectuer une irrigation intelligente en prévoyant les paramètres suivants : température de l'air, température du sol et humidité. L'article propose divers modèles d'apprentissage automatique, à savoir FFNN (Feed Forward Neural Network), LSTM (Long Short-Term Memory) et GRU (Gated Recurrent Unit). Sur la base de leurs résultats, le modèle LSTM est celui qui convient le mieux.

[200] a utilisé l'analyse de régression pour l'analyse des paramètres environnementaux (tels que la précipitation annuelle, la superficie sous culture et l'indice des prix des aliments influençant le rendement des cultures), et l'établissement de la relation entre eux. L'analyse a été effectuée sur la culture du riz en Inde.

Dans [143], l'humidité du sol à Cojocna est prédite à l'aide de neuf algorithmes, à savoir : KNN, SVM, régression logistique, NN, induction de règles (RI : Rule Induction), DT, FLM (Fast Large Margin), RF et régression linéaire. Cette prédiction peut aider l'agriculteur à prendre les bonnes mesures pour éviter les dommages aux cultures et il a été prouvé que la précision de la prédiction était très élevée.

Suresh et Krishna Priya [209] ont développé un modèle statistique pour prévoir le rendement de la canne à sucre. L'article utilise des paramètres météorologiques tels que la température, l'hu-

midité relative et la quantité de pluie collectées à partir de la station météorologique. Les résultats du modèle dépendent de la précision des données d'entrée. Le rendement de la canne à sucre avant la récolte est estimé avec succès avant la récolte effective.

Dans [19], le modèle ARIMA (Auto Regression Integrated Moving Average)(1, 2, 2) est sélectionné comme étant le meilleur des modèles de séries chronologiques mis au point pour prédire la production de blé au Pakistan. Les auteurs prévoient que la production doublera en 2060 par rapport à 2010 sur la base du modèle sélectionné.

Le modèle de prévision de production présenté dans [27] est basé sur l'analyse des séries chronologiques. Deux techniques existantes sont utilisées : SVM, Naive Bayes et deux autres proposées : AdaSVM et AdaNAIVE, pour la prédiction de séries chronologiques. Les expériences ont montré de bons résultats pour les deux techniques proposées en termes d'exactitude et d'erreur de classification.

Ruß et Brenning [194] présentent une nouvelle technique de validation croisée spatiale pour la prédiction du rendement. Ils fournissent également une étude comparative de cette nouvelle approche avec d'autres approches (arbres de régression, réseaux de neurones et machines à vecteurs de support), soulignant la différence importante entre les données spatiales et les données non spatiales. Les auteurs ont conclu qu'il était important d'examiner de plus près les relations spatiales inhérentes aux ensembles de données et que leur nouvelle méthode améliorerait la tâche de prédiction du rendement.

Cai et al. [42] prédisent le rendement du blé en se basant sur les méthodes d'apprentissage automatique (RF, SVM, NN) et la méthode de régression LASSO (Least Absolute Shrinkage and Selection Operator). Différentes données sont utilisées, notamment des données climatiques et satellitaires. L'analyse exploratoire des données est appliquée avant d'utiliser les méthodes d'apprentissage automatique. Les meilleures performances sont montrées en intégrant les données climatiques et satellitaires.

[50] présente une étude visant à comprendre l'interaction entre le sol, les géotypes, le climat et les pratiques de gestion afin de montrer les effets à long terme de différentes approches de gestion des sols sur le rendement des cultures, le stockage du carbone organique du sol et les émissions de gaz à effet de serre. Sur la base de l'analyse de régression, l'article confirme que, dans des conditions climatiques réelles, les systèmes de travail du sol de conservation sont utiles pour atténuer les pertes de carbone organique du sol.

Afin de maximiser le rendement des cultures, [124] a proposé un système de gestion de la canne à sucre basé sur divers capteurs pour analyser avec précision des paramètres tels que la température, l'humidité de l'air et l'humidité du sol responsables de la croissance saine de la culture. Afin de suggérer diverses mesures pouvant être utiles pour optimiser le rendement, deux techniques à savoir KNN et SVM ont été utilisées.

Dans [184], un modèle qui fournit une analyse de la meilleure séquence de cultures, de la chaîne culture à cultiver pour une meilleure production et de la production totale dans une zone

d'intérêt a été proposé. Grâce à l'échantillonnage du sol en temps réel, l'agriculteur sera en mesure d'obtenir les besoins en engrais actuels pour les cultures afin d'améliorer la production agricole tout en réduisant le coût des besoins en engrais. Pour trouver des informations utiles à partir des données d'entrée, l'algorithme de classification C5.0 et les règles d'association sont utilisés.

[33] utilise la méthode PDCA (Plan, Do, Check, Act) afin de prédire le rendement d'abricot. La prédiction est effectuée à l'aide de réseaux de neurones artificiels qui prennent en entrée : la longueur des pousses, le poids des fruits, l'épaisseur des pousses, la quantité d'engrais et le début de la récolte. Les résultats de la prédiction de rendement sont affichés par une application Web.

La cartographie des types de cultures est importante dans différentes applications agricoles. Elle fournit des informations qui peuvent être utiles, notamment pour l'estimation du rendement. Dans [241], une méthode basée sur une grille est proposée pour la classification des cultures basée sur des images à haute et moyenne résolution. Le succès de la méthode proposée est évalué à l'aide d'images GF-1 WFV et d'images sentinelles-2 avec deux classificateurs (RF, SVM). Des résultats intéressants ont été montrés en utilisant les images GF-1 WFV et la méthode forêt aléatoire pour la province du Heilongjiang dans 3 ans.

Wang et al. [232] ont proposé la méthode forêt aléatoire et deux méthodes non supervisées (K-means, modèle de mélange gaussien (Gaussian Mixture Model (GMM))), guidées par des statistiques sur les superficies cultivées pour la cartographie des types de cultures. Différentes données sont utilisées dans l'étude (données satellitaires et statistiques). Les caractéristiques introduites dans les méthodes d'apprentissage automatique sont extraites à l'aide d'une régression harmonique. L'attribution d'étiquettes de cultures dans les méthodes non supervisées est assurée par des statistiques agrégées. Leur méthodologie est validée à l'aide des données sur les terres cultivées.

Dans [254], un CNN est développé pour cartographier une rizière à partir de données satellitaires (MODIS et Landsat 8). Le classificateur CNN est comparé à deux autres classificateurs : SVM et RF. La méthode proposée permet d'obtenir la meilleure performance avec une précision globale de 97,06% par rapport aux autres méthodes. La forte corrélation entre la superficie rizicole dérivée de Landsat et les données statistiques gouvernementales confirme les résultats.

Dans [253], la zone de distribution du maïs est extraite à partir du classificateur SVM et des données satellitaires. Les résultats ont montré qu'il était possible de surveiller la superficie des cultures en utilisant la télédétection. La précision obtenue était de 74%.

Pour identifier la zone de culture du maïs dans la province du Hebei, Tian et al.[213] ont proposé une approche basée sur la technologie de l'informatique en nuage et des images haute résolution (à une résolution de 10 m). Les composites métriques pour les images sentinelle 1 et sentinelle 2 sont calculées et utilisées comme entrée pour le classificateur RF. Le résultat de l'algorithme RF est intégré à la segmentation des objets et une précision globale de 89,89% est atteinte.

Pour améliorer la productivité des champs et donc le rendement des cultures, les agriculteurs peuvent également mettre en œuvre des plans de rotation des cultures annuels ou saisonniers. En effet, une bonne sélection des cultures contribue à l'amélioration de la productivité. La prise de

décision implique la comparaison de différentes données sur la biologie du sol et les besoins des cultures. Dans ce contexte, plusieurs techniques ont été utilisées pour fournir toutes les informations nécessaires à la conception de plans de séquences de cultures à mettre en œuvre sans compromettre la santé du sol ni la productivité à long terme. Ravichandran et Koteeshwari [186] ont proposé un système proposant aux agriculteurs de choisir les bonnes cultures pour leurs terres. Ce système de prédiction est basé sur le réseau de neurones artificiels, qui prend en compte sept paramètres qui déterminent la productivité, à savoir PH, phosphate, potassium, azote, profondeur, température et précipitations, et fournit la culture adaptée. Ce système fournit également le statut de productivité pour la culture souhaitée et les engrais nécessaires. Les résultats ont montré une précision de 92%.

Dans [81], la technique d'analyse d'association est appliquée à une base de données sur la science du sol en Inde. L'article a mis en œuvre l'algorithme Apriori pour prédire des relations significatives.

Le Tableau 2.5 résume certains articles traitant de la prédiction du rendement des cultures et des effets du changement climatique sur la productivité.

**Tableau 2.5 – Résumé de certains articles traitant de la prédiction du rendement des cultures et des effets du changement climatique sur la productivité**

Référence	Objectif	Techniques	Données utilisées	Prétraitement des données
[27]	Prédiction du rendement de 5 cultures	SVM, Naive Bayes, AdaSVM, AdaNaive Bayes	Données historiques sur la production agricole et données climatiques	Les caractéristiques moins importantes sont ignorées
[231]	Prédiction de la rouille de la bande de blé	SVM	Trois ensembles de données sont utilisés (données sur les maladies et données météorologiques)	-
[226]	Prédiction de l'influence des paramètres climatiques sur le rendement des cultures sélectionnées	DT	Paramètres climatiques	Il gère les données manquantes
[225]	Prédiction de l'influence des paramètres climatiques sur le rendement du soja	DT	Paramètres climatiques : précipitations, température, évaporation, humidité relative	-
[19]	Prédiction de la production de blé	Techniques de séries chronologiques	Données recueillies à partir des publications gouvernementales	-
[224]	Prédiction de la température de l'air, de la température du sol et de l'humidité	Les réseaux de neurones (FFNN, LSTM, GRU)	Données de température et d'humidité et autres données invariantes	Réduction de données
[200]	Analyse des paramètres environnementaux et établissement de la relation entre eux	Analyse de régression	Paramètres environnementaux (superficie cultivée, précipitations annuelles et indice des prix des denrées alimentaires)	-
[209]	Prédiction du rendement de la canne à sucre	Modèle de régression	Données météorologiques (précipitations, température et humidité relative)	-
[143]	Prédiction de l'humidité du sol	KNN, SVM, régression logistique, NN, RI, DT, FLM, RF et régression linéaire	Données des stations météo	Réduction de données

[194]	Prédiction de rendement	Arbres de régression, NN and SVM, Bagging	Enregistrements de données spatiales et non spatiales de deux champs	-
[42]	Prédiction du rendement du blé	RF, SVM, NN et LASSO	Données climatiques et satellitaires, données sur le rendement du blé, superficie cultivée.	-
[50]	Montrer les effets à long terme de différentes approches de gestion des sols sur le rendement des cultures, le stockage du carbone organique du sol et les émissions de gaz à effet de serre	Analyse de régression	Historique des données météorologiques, cartes de rendement et autres données sur le sol	-
[124]	Détection et identification des infections et des maladies et surveillance des paramètres de croissance saine	KNN, SVM	Images et données de capteurs (température, humidité de l'air et humidité du sol)	Images en niveaux de gris
[184]	Prédire le rendement et décider de la meilleure séquence de cultures	DT et règles d'association	Séquence précédente des cultures et informations sur les éléments nutritifs du sol	Élimination des valeurs bruitées
[33]	prédiction du rendement en abricot	ANN	longueur des pousses, poids des fruits, épaisseur des pousses, quantité d'engrais et début de la récolte	-
[186]	Prédire la culture à semer	ANN	PH, phosphate, potassium, azote, profondeur, température et précipitations	-
[81]	Prédiction de relations significatives dans une base de données sur la science du sol	l'algorithme Apriori	Données scientifiques du sol	-

## 2.4 Discussion

L'étude en cours a montré l'utilité de l'exploration de données pour résoudre les problèmes réels de l'agriculture intelligente. Sur la base de cette étude, plusieurs conclusions et quelques remarques peuvent être tirées sur l'application de Data Mining dans l'agriculture intelligente.

- D'après les articles examinés, la plupart des systèmes proposés basés sur l'exploration de données pour la gestion des cultures se limitent à des cultures spécifiques, contrairement aux systèmes basés sur l'apprentissage en profondeur qui sont appliqués à de nombreuses cultures afin de détecter plusieurs maladies. En outre, la plupart des systèmes de détection et de classification des maladies s'appliquent à un organe spécifique de la plante (feuille ou fruit).
- Un autre point intéressant est que, pour les systèmes basés sur le traitement d'images, les résultats des techniques de Data Mining dépendent fortement de la qualité de l'image traitée (position de capture, résolution de l'appareil, conditions météorologiques) et qui nécessite souvent beaucoup de traitements. De plus, les images prises pour l'apprentissage peuvent ne pas refléter les conditions réelles de la maladie. Pour surmonter ces limitations, il est nécessaire de combiner l'analyse d'images avec d'autres données capturées en temps réel. Le traitement de nombreuses sources d'information hétérogènes est essentiel pour une prise de

décision fiable.

- La plupart des articles examinés traitent d'images d'arrière-plan simples et ce n'est pas le cas en réalité. Ainsi, la conception et le développement de systèmes tenant compte de la complexité de l'arrière-plan de l'image sont nécessaires.
- Différentes techniques d'exploration de données ont été étudiées et comparées en respectant l'analyse de données agricoles. Ainsi, selon notre étude, il a été constaté que les techniques d'apprentissage en profondeur donnent de bons résultats en termes de précision dans la détection et la classification.
- Etant donné l'influence des réglages de paramètres sur la précision de l'exploration de données, il est nécessaire de concevoir et de développer des algorithmes pour obtenir des réglages de paramètres optimaux, en particulier pour les algorithmes évolutifs et d'apprentissage en profondeur. La formation préalable peut améliorer les modèles, comme indiqué dans [39]. La visualisation des caractéristiques apprises peut jouer un rôle clé, en particulier pour comprendre le fonctionnement du classificateur, ce qui permet aux chercheurs de développer et d'améliorer des architectures.
- La technologie mobile offre aux agriculteurs la possibilité de devenir des acteurs de la collecte et du traitement des données de terrain en temps réel à l'aide d'appareils mobiles. Il est donc important de développer des modèles moins gourmands en temps, en espace et offrant une bonne précision à intégrer dans les appareils mobiles. La technologie d'exploration de données distribuée peut aussi être efficace.
- Dans l'agriculture intelligente, les données sont tellement importantes. Elles constituent un outil précieux pour prendre les bonnes décisions, au bon moment et au bon endroit. Les techniques d'exploration de données manipulent différents types de données (images, données de capteurs), y compris des données historiques et des données capturées en temps réel. Les données utilisées dans les articles examinés proviennent de différentes sources telles que Plant-Village, les données de télédétection, les données collectées par le gouvernement, les données des stations météorologiques, etc.
- Les techniques de Data Mining sont largement utilisées dans diverses tâches agricoles. Il n'existe pas de méthode unique d'exploration de données pour résoudre les problèmes de l'agriculture. En outre, la précision des méthodes d'exploration de données varie en fonction des caractéristiques et de la taille de l'ensemble de données. Pour obtenir la plus grande précision, qui est indispensable dans le diagnostic et la prise de décision, en tenant compte des caractéristiques des données prises en charge, il est important de concevoir des modèles hybrides susceptibles d'améliorer l'analyse, la prédiction et la décision.
- Dans l'agriculture intelligente, divers capteurs et différents robots, tels que : drones et rovers, génèrent de nombreux et différents types de données telles que des images et des vidéos.

Compte tenu de la complexité des paramètres à surveiller et à gérer dans un environnement agricole ainsi que de la difficulté à collecter des mesures de certains paramètres importants en raison des dommages éventuels causés aux capteurs, la combinaison de Data Mining avec des techniques améliorées d'intelligence artificielle peut jouer un rôle majeur dans l'amélioration de la précision dans la résolution des problèmes liés au secteur agricole.

- La notion de temps réel en agriculture est essentielle. Avoir de meilleures informations en temps réel offre un grand potentiel pour améliorer l'efficacité et la productivité. Dans un système temps réel, la contrainte à respecter est que la tâche doit être terminée dans les délais prédéfinis. Certains articles prétendent mettre au point des systèmes répondant aux besoins de l'agriculteur en temps réel, basés sur différentes technologies, notamment des capteurs (PH, azote, phosphore, potassium, humidité, profondeur, température, humidité des feuilles, vitesse du vent, direction du vent, appareil photo numérique et capteurs optiques) et les technologies de transmission de données (la plus répandue est ZigBee), toutefois, le traitement des données à l'aide des techniques d'exploration de données adopte la méthode traditionnelle de traitement des données après le transfert des données d'origine au centre de traitement. Pour concevoir des systèmes en temps réel, il est important d'avoir un degré d'intégration élevé et un système de collecte d'informations agricoles offrant un service tout-en-un pour l'acquisition et le traitement des données.

## 2.5 Défis de Data Mining dans l'agriculture intelligente

L'application de Data Mining pour améliorer la productivité agricole implique un certain nombre de défis pour la collecte, le traitement et l'exploitation des données. L'un des plus grands défis de l'agriculture intelligente est de savoir comment garantir la sécurité et la confidentialité des données. Un autre problème clé des systèmes agricoles est que la disponibilité et la qualité des données sont souvent médiocres et doivent être garanties avant leur exploitation.

Ce problème est plus difficile avec de grandes quantités de données en temps réel. L'évolutivité et l'intégration des données spatiales et sémantiques dans l'exploration de données posent également des problèmes dans ce domaine.

### 2.5.1 Confidentialité

Dans les applications agricoles, les données collectées contiennent des informations personnelles sur les agriculteurs (identité, localisation géographique, informations économiques, connaissances entrepreneuriales, etc.). Beaucoup d'agriculteurs dont leurs données sont accessibles via des plateformes numériques ne sauront pas ce qui se passe avec leurs données. En effet, les agriculteurs ne savent pas que leurs informations personnelles sont collectées et, surtout, à quoi vont servir leurs données ?

L'exploration de données permet aux entreprises d'accumuler et d'analyser une grande quantité d'informations sur les agriculteurs, suffisante éventuellement pour créer des profils de personnalité. Il est même possible que ces informations puissent être utilisées à l'encontre de l'agriculteur

ou simplement révélées à d'autres personnes, ce qui peut avoir un effet négatif sur toutes ses activités. Les agriculteurs doivent être rassurés que leurs données seront analysées afin de produire des idées productives sans que ces mêmes données ne soient mises à la disposition de leurs concurrents ou exploitées à des fins lucratives.

Les informations tirées des données à l'aide de l'exploration de données constituent une ressource extrêmement précieuse, mais peuvent poser un risque pour la vie privée des agriculteurs. Des efforts considérables doivent donc être investis dans les protocoles de confidentialité.

### **2.5.2 Qualité et précision des données**

Dans l'agriculture intelligente, l'information est si importante qu'une grande quantité de données devrait être collectée et transformée en information. Cette masse de données brutes devrait être énorme dans l'ombre de ce développement technologique que les outils d'analyse traditionnels ne peuvent pas analyser avec succès. L'application de Data Mining dans ce domaine aide considérablement l'agriculteur à mieux gérer ses cultures en fournissant une bonne analyse et une meilleure décision.

L'obtention d'informations fiables et utiles à partir d'un processus de DM requiert une bonne qualité des données. Les données en agriculture sont généralement collectées à partir de sources différentes avec des configurations de base de données et des modèles totalement différents, ce qui les rend ambiguës, complexes, avec beaucoup de données manquantes et des normes de codage différentes pour les mêmes domaines. Le traitement de ce type de données n'est pas simple et nécessite un effort considérable de nettoyage et de prétraitement avant d'être utilisé dans le processus d'exploration de données.

La fourniture de données pour les modèles se heurte à un certain nombre d'obstacles et de défis tels que l'incertitude, le manque et l'incapacité et nécessite donc un traitement (géo-spatial et temporel) et une transformation pour assurer la qualité, la cohérence et la compatibilité avec d'autres sources.

Avec l'augmentation de la disponibilité des données, il est nécessaire de renforcer les exigences afin de garantir l'interopérabilité en alignant syntaxe (formats) et sémantique (définitions). L'amélioration de l'interopérabilité des données offre de nouvelles opportunités pour tous les types d'analyse et de développement de nouveaux produits.

### **2.5.3 Importance des données spatiales**

L'objectif de la gestion de l'agriculture intelligente est d'effectuer et d'estimer le traitement des cultures spécifique au site afin de minimiser les dommages environnementaux et de maximiser les profits. Les approches largement utilisées de l'exploration de données se concentrent généralement sur des bases de données relationnelles et les informations spatiales provenant de sites physiquement dispersés ne sont souvent pas exploitées. De nouveaux algorithmes d'exploration de données adaptés aux ensembles de données spatiales sont nécessaires dans l'agriculture intelligente pour prendre en compte les relations dans les données, telles que l'auto-corrélation spatiale et temporelle des données.

### **2.5.4 Intégration des connaissances du domaine agricole dans l'exploration de données**

L'agriculture est un domaine largement interdisciplinaire comprenant les sciences du climat et de l'environnement, l'économie agricole, les sols, la sélection végétale, etc.

Dans l'agriculture intelligente, les données nécessaires à une tâche donnée de Data Mining peuvent provenir de différentes sources. Cette mutation, qui s'appuie en particulier sur les capteurs et les données à haut débit, pose de nouveaux défis : comment décrire les données produites de manière à ce que leur signification soit décrite avec précision, préservée dans le temps et exploitable automatiquement avec une interopérabilité sémantique plutôt que par des experts humains ?

L'intégration de la connaissance du domaine est un problème complexe dans l'exploration de données. Le Data Mining sémantique peut donc être envisagé comme une solution à l'intégration des connaissances du domaine agricole au Data Mining et reste un domaine de recherche important.

### **2.5.5 Évolutivité des algorithmes de Data Mining**

La révolution de l'agriculture intelligente a entraîné une énorme quantité de données générée par l'utilisation croissante d'un certain nombre de périphériques. En particulier, les capteurs à distance produisent en permanence des données très hétérogènes. Avoir une grande quantité de données est un avantage car elles montreront probablement les relations existantes. Toutefois, pour prendre des décisions en temps réel dans le secteur agricole, où les agriculteurs doivent prendre d'innombrables décisions chaque jour, une question se pose naturellement : les algorithmes de Data Mining peuvent-ils gérer la quantité de données collectées dans un délai raisonnable ?

Il est donc essentiel que les algorithmes d'exploration de données gérant ces ensembles de données volumineux soient évolutifs en utilisant des stratégies de recherche spéciales pour traiter les problèmes de recherche exponentiels. Le développement d'algorithmes parallèles et distribués est un problème essentiel pour l'évolutivité dans le domaine de l'exploitation de données en agriculture.

## **2.6 Conclusion**

Ce chapitre a été consacré pour le Data Mining dans lequel une étude sur l'application de ses techniques dans l'agriculture intelligente est présentée et discutée. Ces techniques sont utilisées pour résoudre divers problèmes agricoles apparaissant dans les systèmes de contrôle des cultures tels que la gestion des maladies et des ravageurs, la planification des intrants (eau, pesticides) et la prédiction du rendement. Les avantages des techniques de Data Mining par rapport aux méthodes statistiques traditionnelles et les caractéristiques particulières des données agricoles rendent l'exploration de données très nécessaire pour l'analyse des données agricoles.

Les données analysées à l'aide des techniques de Data Mining peuvent provenir de plusieurs sources, plus particulièrement des drones. Ces derniers sont une technologie puissante qui tend à

faire une révolution dans l'agriculture. Ils permettent non seulement la collecte de données de haute précision, mais ils peuvent également servir à faire plusieurs tâches comme la pulvérisation avec des produits phytosanitaires. Cependant, leur placement (déploiement) influence considérablement la qualité des services qui seront fournis par ces drones.

Le chapitre suivant aborde le problème du déploiement d'un RCSF et également d'un essaim de drones et passe en revue les différentes approches de résolution présentées dans la littérature.

# Chapitre 3

## Techniques de déploiement dans les réseaux (RCSF, UAVs)

### 3.1 Introduction

Les drones sont devenus un sujet d'actualité. Ils ont vu ces dernières années une utilisation accrue dans différents domaines grâce aux avantages qu'ils procurent, en particulier dans l'agriculture. Il est vrai que cette technologie offre de nouvelles alternatives à l'agriculture où des images à haute résolution spatiale délivrées en temps réel sont nécessaires mais elle fait face à des problèmes très difficiles tels que le problème de déploiement, qui a poussé les chercheurs à redoubler d'efforts pour trouver des solutions satisfaisantes. Le déploiement optimal, qu'il soit statique ou dynamique, déterministe ou aléatoire, a été l'un des sujets cruciaux qui a été largement étudié dans les RCSF et récemment abordé dans les drones. Ce chapitre est consacré aux techniques de déploiement dédiées aux RCSF et aux drones. Commencant par quelques concepts théoriques, le chapitre détaille plus ou moins toutes les approches de déploiement utilisées dans le cas d'environnements statiques et mobiles. De plus, les avantages et les inconvénients de chaque stratégie proposée sont présentés.

### 3.2 Préliminaire / Concepts de Base

#### 3.2.1 Réseaux de capteurs et réseaux de drones

Grâce à la technologie de réseaux, les capteurs ou les drones ne sont plus utilisés de manière unique mais sont organisés en réseaux et profitent des avancées des communications sans fil pour mettre en œuvre des réseaux plus grands et plus efficaces. Ces réseaux permettent une surveillance d'une plus grande zone à l'aide de nœuds collaboratifs. Les réseaux de drones ou de capteurs sans fil présentent une amélioration significative par rapport aux systèmes de surveillance traditionnels. Leur introduction a attiré de nombreux chercheurs et est devenue l'un des domaines de recherche les plus étudiés au cours des dernières années.

Le déploiement des nœuds (capteurs ou drones) est leur placement physique dans une zone de surveillance. De nombreux travaux de recherche se sont focalisés sur le processus de déploiement et on retrouve dans la littérature plusieurs formes de déploiement. Le déploiement peut se faire de

manière aléatoire ou déterministe, mais aussi en une seule activité ou en continu tout au long de la durée de vie du réseau. Quelle que soit la stratégie de déploiement choisie, les emplacements des nœuds sont d'une grande importance car cela affectera directement la qualité de la surveillance. Avant de présenter les différentes techniques largement discutées dans la littérature, quelques notions de base liées au déploiement sont introduites ci-dessous :

- **Déploiement aléatoire Vs déploiement déterministe** : Dans le déploiement déterministe, les positions des nœuds sont prédéterminées. Ce type de déploiement est généralement utilisé dans des applications où des nœuds sont chers ou leur mission est vraiment affectée par les positions. Par contre, dans le déploiement aléatoire, les nœuds sont distribués aléatoirement et il peut être essentiel dans beaucoup de cas pratique notamment pour le RCSF, en raison de l'inaccessibilité de la région d'intérêt ou de la grande échelle du RCSF.
- **Déploiement homogène Vs déploiement hétérogène** : Un réseau peut être hétérogène ou homogène. Un réseau est homogène si tous les nœuds du réseau ont les mêmes capacités (détection, traitement, communication, etc.). En revanche, dans les réseaux hétérogènes, les nœuds sont dotés de capacités différentes, pouvant servir à différentes applications. En règle générale, certains nœuds auront plus de ressources disponibles, telles que la puissance de traitement et l'énergie, que le reste des nœuds [154].
- **Déploiement statique Vs déploiement dynamique** : Dans le déploiement statique, les nœuds ne peuvent pas se déplacer une fois déployés. Par contre dans le déploiement mobile, en plus des fonctions des nœuds fixes, les nœuds ont la capacité de se déplacer.

### 3.2.2 Les modèles de couverture

Un modèle de couverture de nœuds est mathématiquement défini comme une fonction qui accepte des entrées (par exemple des distances entre des nœuds et un point dans l'espace) et donne en sortie la mesure de couverture [154]. La couverture d'un nœud est généralement représentée par deux types de modèles décrits ci-dessous :

**Le modèle de couverture binaire** : c'est le modèle le plus simple dans lequel le facteur d'incertitude n'est pas pris en compte. Dans ce modèle chaque nœud est capable de détecter tous les points qui se trouvent dans sa portée de détection ( $r$ ). La FIGURE 3.1 montre ce modèle.

La probabilité de détection d'un point ( $p$ ) par un nœud ( $s$ ) est représentée par l'équation suivante :

$$P_{s/p} = \begin{cases} 1, & \text{si } d(s, p) < r \\ 0, & \text{sinon} \end{cases} \quad (1)$$

**Le modèle de couverture probabiliste :** ce modèle prend en compte la nature stochastique de la détection. Un exemple d'une formulation mathématique de ce modèle [157] est donné par l'équation (2).

$$P_{s/p} = \begin{cases} 0, & \text{si } r + r_e < d(s, p) \\ e^{-\lambda\alpha^\beta}, & \text{si } r < d(s, p) \leq r + r_e \\ 1, & \text{si } d(s, p) \leq r \end{cases} \quad (2)$$

où  $\alpha = d(s, p) - r$ .  $r_e$  est la plage d'incertitude.  $\beta, \lambda$  sont deux constantes qui dépendent des caractéristiques du matériel. Ce modèle est représenté dans la FIGURE 3.1. D'autres formulations mathématiques de ce modèle sont représentées dans [154].

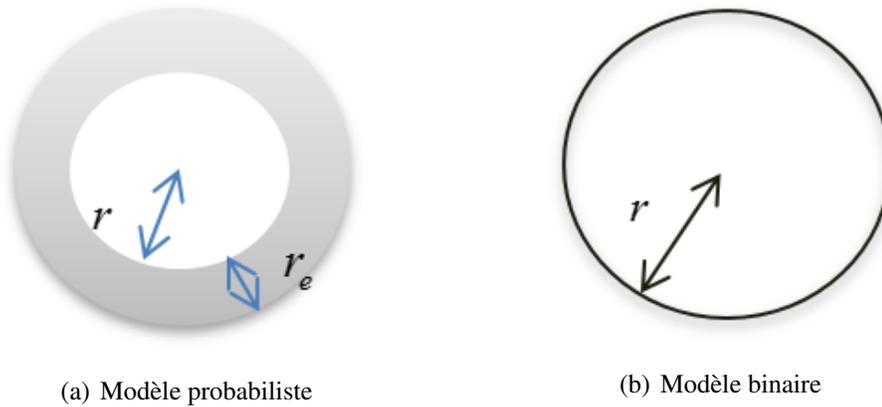


FIGURE 3.1 Les modèles de couverture

### 3.2.3 Les modèles de communication

Un modèle de communication de nœud est un modèle mathématique qui quantifie la connectivité directe entre les nœuds [154]. Le modèle de connectivité de disque est le modèle de communication communément supposé. Dans ce modèle, un nœud peut communiquer avec d'autres nœuds situés dans la portée de communication ( $R_c$ ). En d'autres termes, deux nœuds  $a$  et  $b$  sont connectés si leur distance euclidienne  $d(a, b) \leq R_c$ .

## 3.3 Objectifs et contraintes d'une stratégie de déploiement

Dans les stratégies de déploiement de nœuds, un ou plusieurs objectifs à optimiser doivent être pris en compte selon les besoins de l'application sous une ou plusieurs contraintes. Les objectifs et les contraintes généralement considérés dans la littérature seront introduits dans cette section.

### 3.3.1 Couverture

Elle représente l'objectif principal de tout déploiement. Une couverture totale d'une zone d'intérêt dépend étroitement de déploiement des nœuds. Selon l'exigence de l'application, différents types de couverture peuvent être implémentés. Ces types incluent la couverture d'une zone, la couverture de points, et la couverture de barrière (voir FIGURE 3.2).

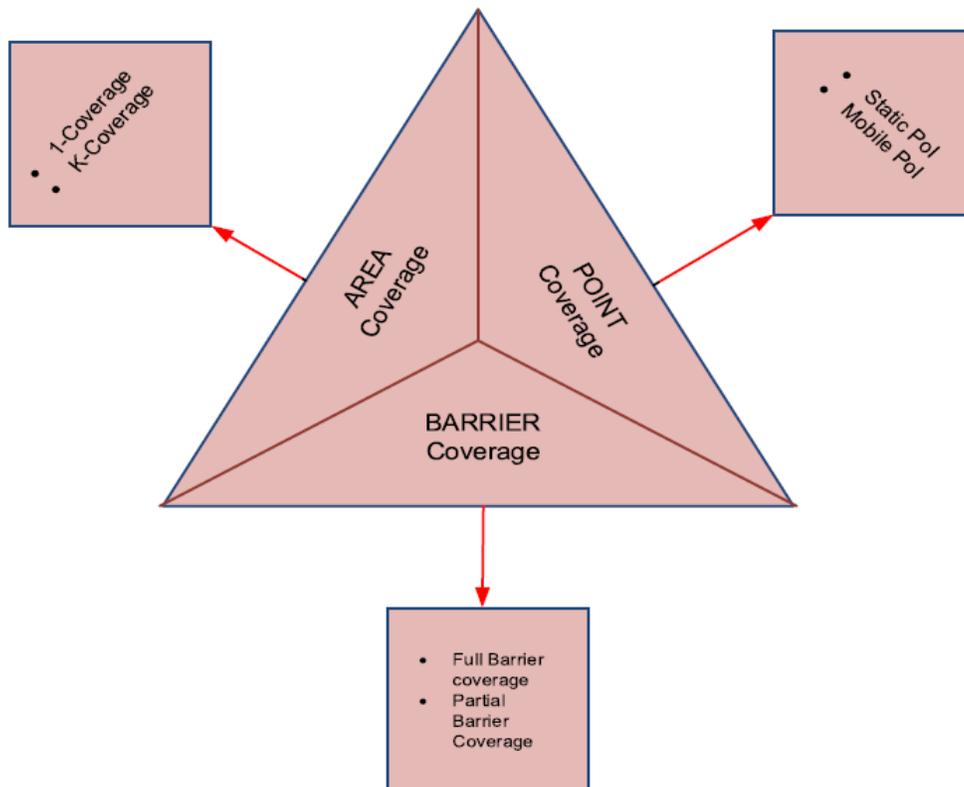


FIGURE 3.2 Les différents types de couverture [178]

La couverture d'une zone/champ vise à couvrir ou à observer toute la zone avec un ensemble de nœuds dont chaque point de la zone est couvert au moins par un nœud. Certaines applications comme les applications militaires nécessitent que chaque point de la zone soit couvert au moins par  $K$  nœuds où  $K$  est une valeur prédéfinie.

La couverture de points ou de cibles vise à surveiller toute cible qui peut être stationnaire ou mobile dans la région d'intérêt.

Le troisième type de couverture qui est la couverture de barrière vise à surveiller les brèches ou les infiltrations éventuelles le long de la barrière [178]. Ce type est encore divisé en deux sous types, la couverture de barrière faible et la couverture de barrière forte. Dans le premier, en raison de la présence de trous de couverture, la cible peut ne pas être détectée. Par contre, dans le second, le balayage précis des cibles par les nœuds sera garanti.

### 3.3.2 Connectivité

Le maintien de la connectivité dans les réseaux de capteurs ou de drones constitue un autre problème de déploiement. En effet, une couverture seule sans connectivité diminuera la qualité du réseau puisque les données transmises par un nœud n'arriveront jamais aux centres de contrôle. Dans certaines applications, à côté de la couverture, la connectivité du réseau est obligatoire.

D'une manière générale, on dit qu'un réseau est connecté si chaque nœud peut communiquer directement ou indirectement (à l'aide de nœuds intermédiaires) avec n'importe quel autre nœud du réseau. La connectivité d'un réseau peut être vérifiée par un graphe de connectivité. Si le graphe n'est pas entièrement connecté, le réseau est supposé non connecté ou partitionné; sinon il est connecté. La FIGURE 3.3 représente un exemple d'un réseau connecté et non connecté.

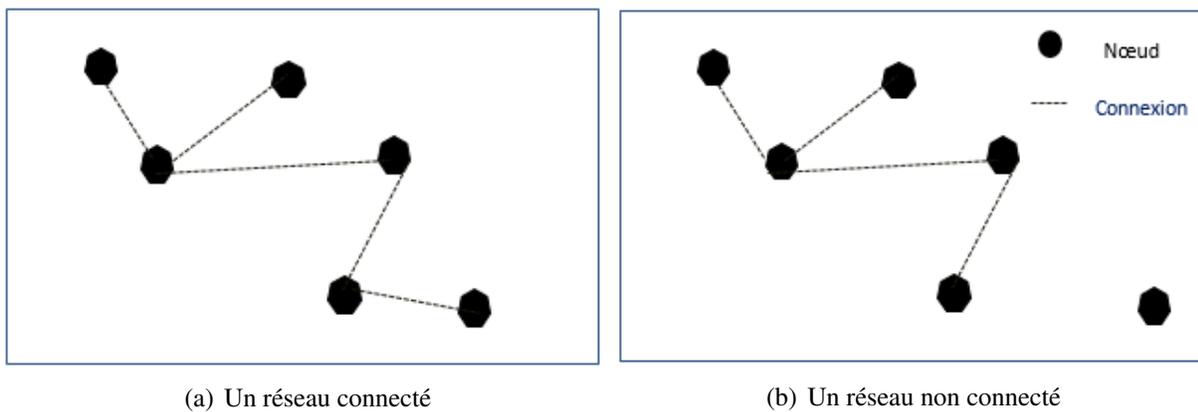


FIGURE 3.3 Exemple d'un réseau représenté par un graphe connecté et non connecté

En outre, tout comme la couverture, la connectivité peut également être simple (unique) ou multiple (K-connectivité). Dans le premier cas, un seul chemin existe entre deux nœuds. En revanche, dans le second cas, de nombreux chemins existent entre deux nœuds du réseau [157].

### 3.3.3 Efficacité énergétique

Ce paramètre est souvent utilisé de manière interchangeable avec la durée de vie du réseau [154, 157]. Étant donné que chaque nœud dans le réseau de capteurs ou de drones a une ressource d'énergie limitée, des mécanismes permettant d'optimiser l'utilisation d'énergie des nœuds sont nécessaires afin de prolonger la durée de vie du réseau. Les techniques de conservation d'énergie dans les RCSF peuvent généralement être classées en deux catégories [45]. La première catégorie consiste à programmer les capteurs pour alterner entre le mode actif et le mode veille et la deuxième catégorie consiste à régler la portée de transmission ou de détection des nœuds. En se basant sur cette dernière, l'énergie peut également être conservée dans les réseaux de drones.

### 3.3.4 La durée de vie du réseau

Comme on l'a mentionné préalablement, le fonctionnement global du réseau peut être affecté par la limitation des ressources énergétiques. Après le déploiement du réseau, il est difficile ou impossible de remplacer la batterie des nœuds ou de la recharger. Donc, c'est important d'améliorer la consommation d'énergie pour maximiser la durée de vie globale. Cette dernière peut être décrite comme l'intervalle de temps entre le déploiement et l'instant où une partition réseau se produit. Elle peut aussi être vue comme l'intervalle de temps entre le déploiement et l'instant où le pourcentage des nœuds actifs tombe en dessous d'un seuil spécifique [4]. D'autres définitions sont présentées dans [159].

La durée de vie du réseau est aussi reliée à deux paramètres qui sont le nombre de nœuds et leurs positions. En effet un petit nombre de nœuds ne garantit pas une durée de vie longue et un grand nombre de nœuds ne signifie pas une grande durée de vie. La durée de vie du réseau peut être améliorée en adoptant différentes solutions comme le déploiement incrémentiel et la relocalisation des nœuds.

### 3.3.5 Le coût de déploiement et le nombre de nœuds

Le coût de déploiement est l'une des mesures essentielles à prendre en compte dans le processus de déploiement d'un réseau. En effet, l'utilisation d'un grand nombre de nœuds implique un coût élevé de déploiement. Autrement dit, plus le nombre de nœuds est élevé plus le coût de déploiement est élevé. Le coût de déploiement varie en fonction de l'application et des équipements de détection (les capteurs intégrés dans le nœud).

### 3.3.6 Fiabilité des données

La crédibilité des données collectées constitue un autre facteur important à prendre en compte lors de la conception et du déploiement des réseaux. Les lectures de différents capteurs indépendants souvent hétérogènes permettent de fournir une évaluation des phénomènes détectés. La fusion de ces données augmente la fiabilité des incidents signalés et diminue la probabilité de fausses alarmes. La précision des données fusionnées peut être améliorée en augmentant le nombre de nœuds déployés dans la région d'intérêt. Cependant, la redondance des nœuds entraînera des coûts supplémentaires.

### 3.3.7 Tolérance aux pannes et équilibrage énergétique

Dans un réseau, différents types de pannes (par exemple pannes logicielles et matérielles) peuvent survenir et résulter l'interruption de bon fonctionnement du réseau. Par conséquent, un degré de tolérance aux pannes doit être assuré. Ce degré de tolérance aux pannes varie d'une application à une autre selon ses exigences qui peuvent être différentes [48].

Il existe diverses techniques de tolérances aux pannes permettant d'éviter les pannes et parmi elles on cite l'équilibrage de la consommation d'énergie entre les nœuds et la redondance des nœuds.

### 3.4 Les stratégies de déploiement des RCSF

Le placement des capteurs pour la plupart des RCSF constitue une étape importante dans leur conception. Ce problème est connu dans la littérature sous plusieurs appellations comme le problème de placement, de couverture ou de déploiement des RCSF. La stratégie de déploiement a une influence sur la couverture, la consommation d'énergie et aussi sur la durée de vie du réseau, faisant du déploiement un problème très critique dans les RCSF. Une bonne stratégie de déploiement ne réduit pas seulement le coût de déploiement mais aussi augmente la durée de vie du réseau [4]. Elle requiert d'autres objectifs qui sont souvent contradictoires.

Les capteurs déployés dans les RCSF pourraient être homogènes et / ou hétérogènes, statiques, mobiles. Selon les exigences de l'environnement et de l'application, les capteurs peuvent être déployés de manière aléatoire ou déterministe.

Plusieurs classifications des stratégies de déploiement sont fournies dans la littérature [4, 154, 159, 178, 157, 76]. Dans cette section, la classification proposée par [76] est adoptée (voir FIGURE 3.4).

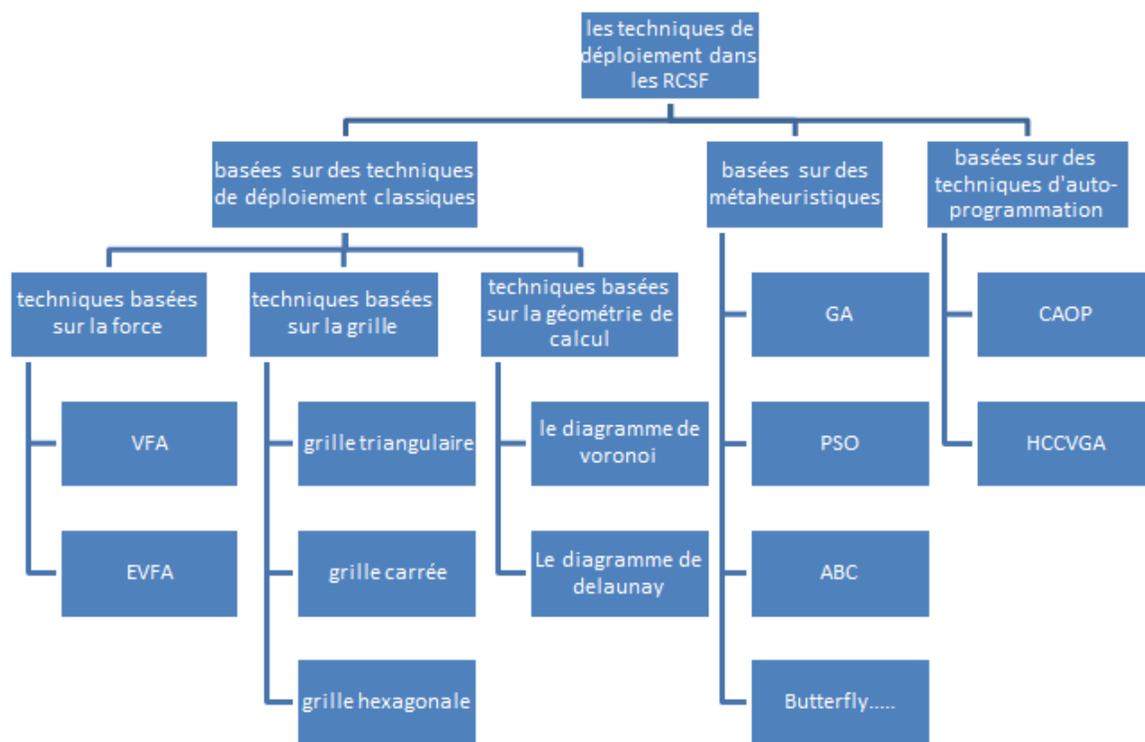


FIGURE 3.4 La classification des différentes techniques de déploiement [76]

### 3.4.1 Déploiement basé sur des techniques de déploiement classiques

Comme on l'a mentionné précédemment, selon les applications du RCSF, le succès de son fonctionnement dépend étroitement de la position de ses capteurs. Compte tenu du déploiement aléatoire des capteurs adopté par une variété d'applications, dans lequel les capteurs sont répartis de manière irrégulière dans la région d'intérêt, certaines zones contiennent un grand nombre de capteurs déployés tandis que d'autres zones peuvent contenir un nombre de nœuds qui peut ne pas être suffisant [178, 76]. Par conséquent, l'obtention d'une couverture complète ne peut pas être assurée à cause de l'apparition probable de trous de couverture.

Les stratégies de déploiement classiques sont catégorisées en trois classes, à savoir les techniques basées sur la force, les techniques basées sur la grille et les techniques basées sur la géométrie de calcul.

#### 3.4.1.1 Techniques basées sur la force

Cette catégorie de techniques est basée sur des forces virtuelles qui peuvent être attractives, répulsives, ou nulles. La distance entre une paire de nœuds adjacents permet de définir la nature de la force exercée. Lorsque la distance séparant deux capteurs voisins est inférieure à une valeur de seuil prédéfinie, une force de nature répulsive est exercée. Tandis que, si cette distance est supérieure au seuil, une force d'attraction est exercée. Sinon (c.à.d. que la distance est égale au seuil) une force nulle est exercée. La FIGURE 3.5 montre un exemple de trois forces différentes exercées sur un nœud.

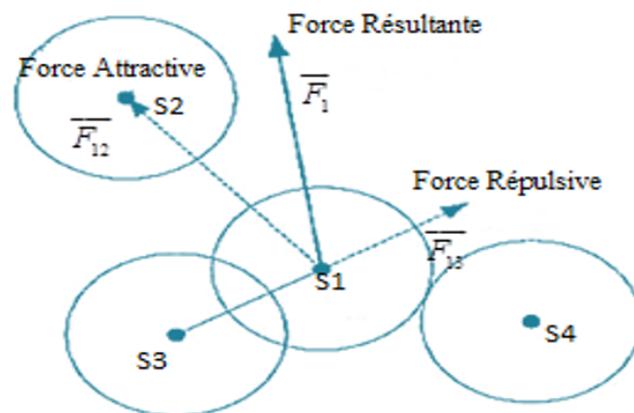


FIGURE 3.5 Un exemple de trois forces différentes sur un nœud [154]

Les techniques basées sur la force sont utilisées afin de déplacer les capteurs vers leur prochaine position. Dans la littérature, plusieurs travaux ont adopté ce type de stratégies. Rout et Roy [191] ont proposé un schéma de déploiement autonome des nœuds basé sur la force appelé OAVFA

(Obstacle Avoidance Virtual Force Algorithm). Cette stratégie a été introduite pour le déploiement de capteurs mobiles homogènes et hétérogènes dans une zone de détection carrée contenant différentes formes d'obstacles. Les auteurs visent à améliorer la couverture et garantir la connectivité en présence d'obstacles. L'algorithme proposé dans cet article est un algorithme localisé puisque il est exécuté sur chaque capteur. Différentes sources de forces sont considérées à savoir les nœuds voisins, les obstacles et les limites du champ. Selon la force totale calculée, un nœud peut se déplacer vers sa nouvelle position. Leur approche a fourni des meilleurs résultats par rapport à IVFA (Improved Virtual Force Algorithm) et EVFA (Exponential Virtual Force Algorithm) pour le déploiement des nœuds hétérogènes et homogènes avec et sans obstacles.

Dans [131], une amélioration de l'algorithme de force virtuelle pour le déploiement d'un réseau de capteurs mobiles est proposée. La force d'équilibre et la gravitation centrale sont également incluses dans l'équation de force présentée. Les résultats de simulation ont montré l'efficacité de l'amélioration proposée en se basant sur deux métriques qui sont le taux de couverture et le degré d'homogénéité permettant de mesurer l'uniformité de la distribution des capteurs.

Un algorithme d'auto-déploiement de nœuds basé sur la force virtuelle pour les nœuds mobiles est présenté dans [245]. Dans l'approche proposée, la triangulation de Delaunay est introduite dans l'algorithme de force virtuelle d'origine pour limiter la source de force et pour définir la relation d'adjacence. La relation d'adjacence des nœuds est définie si deux nœuds sont connectés dans le diagramme de Delaunay. Les résultats de la simulation ont montré que l'approche proposée a un temps de convergence plus rapide et un taux de couverture plus élevé que l'algorithme de force virtuelle original.

Un algorithme de redéploiement distribué basé sur des forces virtuelles a été introduit dans [152] (DVFA : Distributed Virtual Force Algorithm). Cet algorithme permet de redéployer les nœuds initialement déployés d'une manière aléatoire pour garantir la connectivité et la couverture tout en minimisant la consommation d'énergie en minimisant la distance parcourue par chaque capteur. Les résultats de simulation ont montré que l'algorithme distribué surpasse un autre algorithme centralisé (CVFA : Centralized Virtual Force Algorithm).

Dans [246], une approche de déploiement de capteurs basée sur la force de Van Der Waals est proposée. La triangulation de Delaunay est utilisée afin de définir la relation d'adjacence des nœuds. De plus, la force du frottement est considérée dans le calcul de la force permettant de produire une accélération des nœuds pour qu'ils se déplacent. Les auteurs ont introduit une métrique pour évaluer l'uniformité de la distribution des nœuds appelée "la fonction de corrélation de paires". Les résultats de la simulation ont montré que leur stratégie a une configuration plus uniforme, un taux de couverture plus élevé et converge rapidement par rapport à quelques autres algorithmes basés sur la force.

Van Le et al. [222] ont adressé le problème de couverture d'un phénomène en mouvement dans une zone inconnue et ouverte. Ils ont proposé une approche qui est basée sur les forces virtuelles appelée VirFID (Virtual Force (VF)-based Interest-Driven moving phenomenon monitoring). Dans cette dernière, les capteurs se déplacent vers les positions où des données de détection plus intéressantes peuvent être obtenues en utilisant la force virtuelle calculée selon de la distance entre les

capteurs et des valeurs détectées dans la zone d'intérêt. Trois variantes de VirFID ont été proposées à savoir VirFID-GHL (Global High and Lowest), VirFID-LIB (Local Information-Based) et VirFID-IBN (Interest at Boundary Nodes). En termes de l'efficacité de la couverture pondérée, les résultats de simulation ont montré que les algorithmes proposés surpassent certains autres algorithmes.

Dans [212], un déploiement d'un réseau de capteurs mobiles a été abordé afin de former une topologie hexagonale qui a été prouvée théoriquement qu'elle permet d'obtenir une couverture maximale en utilisant un nombre constant de capteurs. Pour atteindre cet objectif, les auteurs ont proposé un algorithme de force virtuelle basé sur les lois physiques dans un système de plasma poussiéreux (VFA-DP (Dusty Plasma system)). En se basant sur la métrique de performance introduite qui est la fonction de corrélation de paires dans une structure cristalline, la conformité de la distribution des nœuds à un hexagone est évaluée et aussi la rapidité de la convergence de l'algorithme proposé. Les résultats de simulation ont montré que l'algorithme proposé surpasse un algorithme de force virtuelle basé sur le potentiel de Lennard-Jones (VFA-LJ).

Le Tableau 3.1 résume les objectifs atteints et les limites des travaux de recherches cités précédemment avec les stratégies de déploiement utilisées basées sur la force.

**Tableau 3.1** – Résumé des objectifs atteints et les limites des articles cités précédemment avec les différentes stratégies utilisées basées sur la force

Article   Stratégie utilisée	Les objectifs atteints	Limites
[191]   OAVFA	<ul style="list-style-type: none"> <li>• Une bonne couverture et maintenance de la connectivité</li> <li>• Traitement des obstacles</li> </ul>	<ul style="list-style-type: none"> <li>• Présence de trous de couverture</li> </ul>
[131]	<ul style="list-style-type: none"> <li>• Une bonne couverture</li> </ul>	<ul style="list-style-type: none"> <li>• Les obstacles, la consommation d'énergie et la connectivité ne sont pas considérés</li> </ul>
[245]	<ul style="list-style-type: none"> <li>• Un taux de couverture élevé</li> </ul>	<ul style="list-style-type: none"> <li>• Les obstacles, la consommation d'énergie et la connectivité ne sont pas considérés</li> </ul>
[152]   DVFA	<ul style="list-style-type: none"> <li>• Assurance de la couverture et de la connectivité</li> <li>• Consommation d'énergie réduite par nœud</li> </ul>	<ul style="list-style-type: none"> <li>• Les obstacles ne sont pas considérés</li> </ul>

[246] | basée sur la force de Van der Waals

- Un taux de couverture élevé
- Temps de convergence acceptable
- La consommation d'énergie et la connectivité ne sont prises en compte
- Terrain irrégulier et les obstacles ne sont pas considérés

[222] | VirFID

- Une couverture acceptable
- Les coûts de communication élevés

[212] | VFA-DP

- Couverture totale est assurée
- Les coûts d'énergie et de la communication et les obstacles ne sont pas considérés

### 3.4.1.2 Techniques basées sur la grille

Ce type de stratégies est considéré comme un déploiement déterministe où les positions des nœuds sont déterminées selon un modèle de grille spécial tel que le treillis triangulaire, la grille carrée ou la grille hexagonale [114] (voir FIGURE 3.6). Selon l'algorithme de déploiement utilisé, la zone d'intérêt est divisée en petites cellules et les nœuds de capteurs sont situés soit dans les sommets des cellules ou au centre des cellules. Etant donné que toutes les cellules de la grille générées ont la même taille et la même forme, ce modèle de déploiement est considéré régulier.

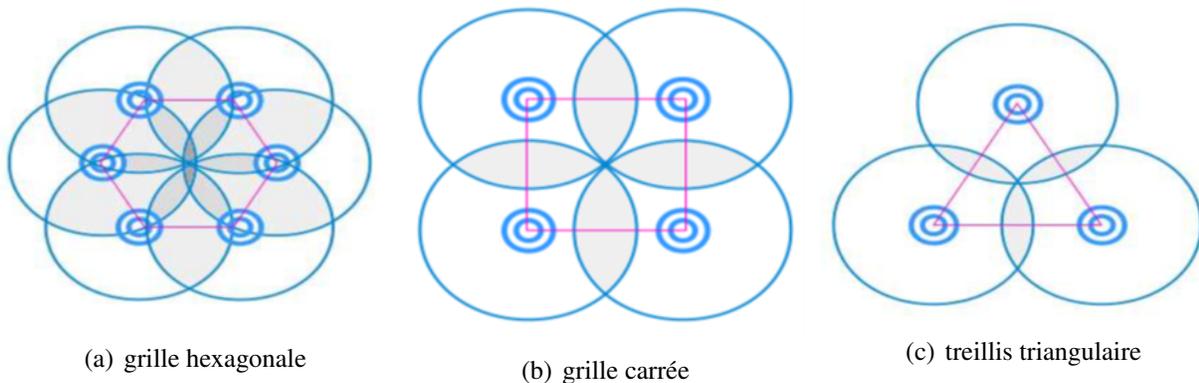


FIGURE 3.6 Les modèles les plus utilisés [154]

Des modèles de déploiement réguliers ont été étudiés dans [115] afin d'obtenir une p-couverture et q-connectivité dans les RCSF. En particulier, les auteurs ont proposé trois modèles de déploiement qui sont la grille carrée, la grille hexagonale et le treillis triangulaire. Tous les modèles fournis ont été comparés en termes de nombre de capteurs nécessaire pour avoir p-couverture et q-connectivité.

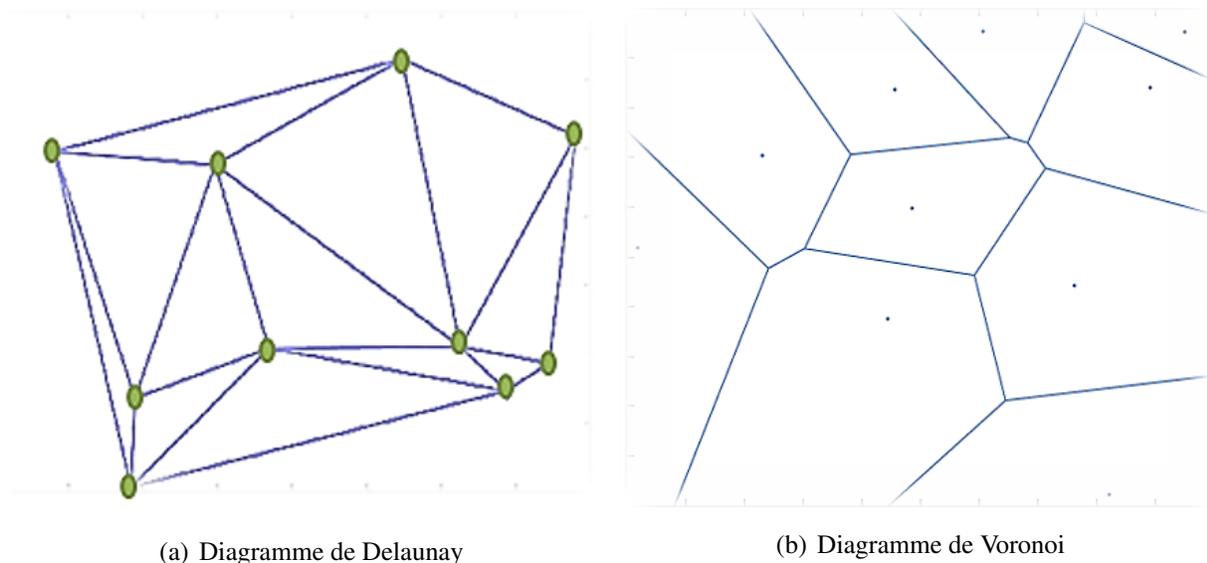
Un algorithme de couverture basé sur une grille carrée virtuelle (VSGCA : Virtual Square Grid-based Coverage Algorithm) est présenté par Liu et al. [137] pour satisfaire les exigences de couverture, de connectivité, de nombre minimum de nœuds actifs et de faible coût de calcul. Dans leur algorithme, la région de détection d'un nœud est divisée en grilles carrées. Si toutes les grilles sont couvertes par des nœuds voisins, le nœud est un nœud redondant. Cet algorithme peut garantir la couverture et la connectivité dans l'ensemble du réseau avec un nombre minimum de capteurs actifs et une complexité de calcul moindre.

Différentes stratégies de déploiement des nœuds sont étudiées dans [141] pour atteindre une couverture complète avec un nombre minimum de capteurs. Ces stratégies incluent le placement triangulaire, grille et bande. Ils ont constaté que la stratégie de déploiement de bande représente le meilleur cas parmi d'autres si le nombre de capteurs n'est pas critique, dans le cas contraire le schéma de déploiement basé sur la grille est le meilleur choix.

Le Tableau 3.2 résume les objectifs atteints et les limites des articles cités précédemment avec les stratégies de déploiement utilisées basées sur la grille.

### 3.4.1.3 Techniques basées sur la géométrie de calcul

La stratégie de géométrie de calcul est utilisée pour résoudre des problèmes basés sur des objets géométriques : points, polygones, segments de ligne, etc [114]. La triangulation de Delaunay et le diagramme de Voronoi sont les deux méthodes de géométrie de calcul les plus utilisées pour résoudre le problème de déploiement dans les RCSF [76] (voir FIGURE 3.7).



**FIGURE 3.7** Exemple d'un diagramme de Voronoi et d'une triangulation de Delaunay [76]

Le diagramme de Voronoi est défini dans [114] comme une méthode de partitionnement de la zone d'intérêt en un certain nombre de polygones en fonction des distances entre les nœuds des

capteurs. Chaque nœud occupe un seul polygone et est plus proche de n'importe quel point de ce polygone que tout autre nœud des polygones voisins. Par contre, la triangulation de Delaunay est le graphe dual de diagramme de Voronoi et il est construit en connectant tous les deux points adjacents dans le diagramme de Voronoi dont les polygones partagent une arête commune.

Dans [211], des algorithmes basés sur la triangulation de Delaunay ont été conçus afin de découvrir les trous de couverture dans la région d'intérêt et de les couvrir d'une manière efficace. La méthode proposée traite des régions régulières, irrégulières et des obstacles. Par rapport à d'autres méthodes, celle proposée peut atteindre une couverture complète tout en économisant considérablement les capteurs, en particulier lorsque la région et les obstacles sont irréguliers.

Une approche de couverture dynamique d'une longue région basée sur Voronoi est proposée dans [2]. Dans l'approche proposée, les agents s'adaptent à la zone de couverture variable en imposant la dynamique des frontières à leur loi de contrôle respective. Cette dernière garantit que les agents se déplacent vers le centroïde de leur cellule de Voronoi respective tout en tenant compte de l'effet des limites mobiles. La méthode de couverture proposée assure un partitionnement (local) optimal pour la zone de couverture mobile. Les auteurs ont examiné la performance de leur méthode en utilisant des sections de la rivière Ohio et une zone tampon frontalière.

Wang et al. [230] ont utilisé le diagramme de Voronoi pour détecter les trous de couverture. Ils ont proposé trois algorithmes à savoir VEC (VECTor based algorithm), VOR (Voronoi based algorithm) et Minimax pour calculer les positions cibles des capteurs en cas d'existence des trous de couverture. Dans le premier algorithme, les capteurs se déplacent des zones denses vers des zones moins denses. Par contre dans l'algorithme VOR, les capteurs migrent vers les trous et dans Minimax, ils se déplacent également vers les trous, mais de manière plus conservatrice pour ne pas générer de nouveaux trous.

Un résumé des objectifs atteints et des limites de ces articles avec les stratégies de déploiement basées sur la géométrie de calcul utilisées est présenté dans le Tableau 3.2.

**Tableau 3.2** – Résumé des objectifs atteints et les limites des articles cités précédemment avec les différentes stratégies utilisées basées sur la grille et sur la géométrie de calcul

Article   Stratégie utilisée	Les objectifs atteints	Limites
[115]   basée sur la grille carrée, la grille hexagonale et le treillis triangulaire	<ul style="list-style-type: none"> <li>• Une couverture multiple et une connectivité multiple</li> </ul>	<ul style="list-style-type: none"> <li>• La couverture dépend de différents rapports de <math>R_c</math> (la portée de communication) et <math>R_s</math> (la portée de détection)</li> </ul>

[137]] basée sur la grille carrée	<ul style="list-style-type: none"> <li>• Assurance de la couverture et de la connectivité avec un nombre minimum de nœuds</li> <li>• Une consommation d'énergie et complexité de calcul minimales</li> </ul>	<ul style="list-style-type: none"> <li>• Il ne traite pas la mobilité des nœuds et les obstacles ne sont pas considérés</li> </ul>
[141]] basée sur le treillis triangulaire, la grille et bande	<ul style="list-style-type: none"> <li>• Une couverture complète avec un nombre minimum de capteurs</li> </ul>	<ul style="list-style-type: none"> <li>• La connectivité n'est pas prise en compte</li> </ul>
[2]] basée sur Voronoï	<ul style="list-style-type: none"> <li>• Couverture dynamique</li> </ul>	<ul style="list-style-type: none"> <li>• Approche centralisée</li> <li>• L'hétérogénéité des nœuds n'est pas considérée</li> </ul>
[211]] basée sur la triangulation de Delaunay	<ul style="list-style-type: none"> <li>• Couverture totale en présence d'obstacles</li> </ul>	<ul style="list-style-type: none"> <li>• Approche centralisée</li> <li>• L'hétérogénéité et la mobilité des nœuds ne sont pas considérées</li> </ul>
[230]] basée sur Voronoï	<ul style="list-style-type: none"> <li>• Réduction des trous de couverture</li> <li>• Assurance de la connectivité</li> </ul>	<ul style="list-style-type: none"> <li>• La portée de communication est faible</li> <li>• Le nombre de nœuds n'est pas considéré</li> </ul>

### 3.4.2 Techniques basées sur des méta-heuristiques

Plusieurs méta-heuristiques ont été utilisées dans la littérature pour résoudre le problème de couverture dans les RCSF. Ce problème a été traité par plusieurs articles comme un problème d'optimisation mono-objectif ou multi-objectif. Dans [162], une méthode basée sur la colonie d'abeilles artificielles (ABC : Artificial Bee Colony) a été proposé pour résoudre le problème de déploiement dynamique des RCSF. L'algorithme ABC a été comparé avec l'algorithme d'optimisation par essaim de particules (PSO) et les résultats ont montré que l'algorithme ABC permettait un meilleur déploiement que l'algorithme PSO. Le même algorithme est aussi utilisé dans [163] pour augmenter le taux de couverture.

Yu et al. [248] ont proposé une version modifiée de l'algorithme ABC pour optimiser le déploiement des capteurs. La version originale de cet algorithme a subi des modifications au niveau de la phase d'abeille spectatrice et de la phase d'abeille éclairceuse. L'objectif principal était d'avoir une convergence rapide et une bonne couverture.

Li et Lei [133] ont proposé une optimisation par essaim de particules améliorée en optimisant l'emplacement des nœuds capteurs et en augmentant le taux de couverture. Une comparaison a été effectuée avec un algorithme de force virtuelle et les résultats ont montré l'efficacité de l'algorithme amélioré.

Dans [88], deux algorithmes pour le placement des nœuds relais qui fournissent la k-connectivité des nœuds capteurs ont été proposés. Le premier est basé sur un algorithme génétique et le second est basé sur un algorithme gourmand (Greedy algorithm). Les résultats de leurs expérimentations ont montré que l'algorithme génétique est meilleur que l'algorithme gourmand. En outre, l'algorithme génétique a des performances plus élevées car il donne une solution optimale globale tandis que l'algorithme gourmand donne une solution optimale locale.

Liao et al. [134] ont proposé une stratégie basée sur l'algorithme de colonies de fourmis (ACO : Ant Colony Optimization) pour obtenir une couverture complète et maximiser la durée de vie du réseau. Le problème de déploiement des capteurs a été modélisé comme un problème des sacs à dos multiples. Les résultats de simulations ont montré que l'algorithme présenté peut prolonger la durée de vie du réseau dans les scénarios de déploiement considérés.

Un algorithme génétique modifié (MGA-RNP : Modified Genetic Algorithm for Relay Node Placement) a été proposé dans [83] pour résoudre le problème de placement des nœuds relais dans les RCSF. Les auteurs ont présenté une approche multi-objectif permettant de minimiser le nombre de nœuds relais tout en garantissant la connectivité entre les nœuds. Les résultats de simulations ont montré l'efficacité de l'algorithme présenté.

Dans [71], le recuit simulé (SA : Simulated Annealing) qui est une méthode de recherche probabiliste inspirée du processus de recuit en métallurgie [65] a été combinée avec la méthode du gradient pour un déploiement optimal du RCSF avec un nombre minimum de nœuds. La méthode proposée permet de résoudre à la fois les problèmes de couverture de zone et de barrière, pour les applications à 1-couverture et 1-connectivité.

Deux optimiseurs d'essaim de particules à savoir CCPSO2 (Cooperative Coevolutionary Particle Swarm Optimization 2) et CLPSO (Comprehensive Learning Particle Swarm Optimizer) ont été utilisés dans [44] afin de maximiser la couverture et de prolonger la durée de vie du réseau dans un espace industriel 3D avec des obstacles. Les résultats expérimentaux ont montré que lorsqu'un nombre croissant de nœuds relais est utilisé, CCPSO2 et CLPSO fournissent de meilleures performances en prolongeant la durée de vie du réseau.

Un algorithme de déploiement basé sur la recherche d'harmonie (HS : Harmony Search) a été proposé dans [150] pour maximiser la couverture et minimiser le nombre de drones dans le réseau. Les résultats de simulation ont démontré la capacité de l'algorithme proposé à atteindre une couverture complète avec un nombre minimum de nœuds. De plus, il a une convergence rapide par rapport à un algorithme génétique. Cependant, la connectivité réseau et la consommation d'énergie des nœuds n'ont pas été prises en compte.

Un résumé des objectifs atteints et des limites des stratégies de déploiement basées sur les

méta-heuristiques est présenté dans le Tableau 3.3.

**Tableau 3.3** – Résumé des objectifs atteints et les limites des articles cités précédemment avec les différentes stratégies utilisées

Article   Stratégie utilisée	Les objectifs atteints	Limites
[162]   ABC	Augmentation du taux de couverture	<ul style="list-style-type: none"> <li>• La connectivité n'est pas considérée</li> <li>• Une consommation d'énergie élevée</li> </ul>
[163]   ABC	Augmentation du taux de couverture	<ul style="list-style-type: none"> <li>• La connectivité n'est pas considérée</li> <li>• Une consommation d'énergie élevée</li> </ul>
[248]   ABC	Amélioration du taux de couverture	<ul style="list-style-type: none"> <li>• La connectivité et la consommation d'énergie ne sont pas considérées</li> </ul>
[133]   PSO	Augmentation du taux de couverture	<ul style="list-style-type: none"> <li>• La connectivité et la consommation d'énergie ne sont pas considérées</li> </ul>
[88]   GA, algorithme gourmand	Une K-connectivité avec un nombre minimum de nœuds relais	<ul style="list-style-type: none"> <li>• Une consommation d'énergie élevée</li> </ul>
[134]   ACO	Une couverture complète et une durée de vie maximale du réseau	<ul style="list-style-type: none"> <li>• La connectivité n'est pas considérée</li> </ul>
[83]   MGA-RNP	Minimisation du nombre de nœuds relais tout en garantissant la connectivité entre les nœuds	<ul style="list-style-type: none"> <li>• La consommation d'énergie n'est pas considérée</li> </ul>
[71]   SA combinée avec l'algorithme du gradient	Résolution des problèmes de couverture de zone et de barrière avec un nombre minimum de nœuds	<ul style="list-style-type: none"> <li>• La connectivité n'est pas assurée en cas de défaillance d'un nœud</li> </ul>

[44] PSO	Maximisation de la couverture, prolongation de la durée de vie du réseau et traitement des obstacles	<ul style="list-style-type: none"> <li>• La connectivité n'est pas garantie</li> </ul>
[150] HS	Une couverture complète avec un nombre minimum de nœuds	<ul style="list-style-type: none"> <li>• La connectivité et les obstacles ne sont pas considérés</li> </ul>

### 3.4.3 Techniques basées sur l'auto-programmation

L'idée principale de ce type de stratégies est de rendre les nœuds redondants dans un état de sommeil. Une approche efficace pour économiser l'énergie dans les RCSF consiste à planifier des intervalles de sommeil pour les nœuds redondants, tout en gardant les nœuds restants actifs pour maintenir la couverture et la connectivité du réseau.

Dans [96], un algorithme appelé CAOP (Coverage Aware scheduling for Optimal Placement) a été proposé pour maintenir la couverture de la zone et prolonger la durée de vie du réseau par une consommation énergétique efficace. À cet effet, le nombre minimum de nœuds pouvant couvrir toute la zone est d'abord calculé puis leurs meilleurs emplacements. Après cela, la zone est divisée en sous-zones dont chacune peut être couverte par un seul capteur. Ensuite, dans chaque sous-zone, une procédure de programmation locale planifie l'ordre d'activation du capteur. Les résultats de la simulation ont montré que l'algorithme proposé peut maximiser la durée de vie du réseau tout en conservant une couverture de zone complète.

Le problème de la couverture a également été abordé dans [13]. Les auteurs ont proposé une méthode pour planifier les nœuds capteurs de telle sorte que seul un nombre minimum de capteurs soit actif, répondant aux exigences de connectivité et de couverture et prolongeant la durée de vie du réseau. De plus, ils ont proposé une heuristique pour la couverture connectée sur une architecture de grille virtuelle (HCCVGA : Heuristic for Connected Coverage over Virtual Grid Architecture). Les résultats de la simulation ont montré l'efficacité de l'heuristique proposée.

## 3.5 Le déploiement des réseaux de drones

Le problème de couverture est l'un des problèmes les plus fondamentaux des drones. En effet, un mauvais déploiement entraînera une dégradation des services fournis en termes de qualité et entraînera un coût d'investissement supplémentaire.

Ce problème est généralement défini comme la capacité des nœuds UAVs à surveiller l'espace donné et à coopérer entre eux. Le problème de couverture de zone dans les réseaux de drones coopératifs est à la base de nombreuses applications.

Le problème de couverture de zone est un problème classique dans un réseau de capteurs sans fil où déjà, une étude approfondie a été effectuée et de nombreuses approches ont été avancées.

Les UAVs peuvent embarquer des capteurs sans fil lors de l'exécution de missions, de sorte que les solutions aux problèmes de couverture dans les réseaux de capteurs sans fil peuvent parfois également s'intégrer dans les réseaux d'UAVs. Cependant, un réseau d'UAVs a ses propres caractéristiques et cela conduit à de nombreuses nouvelles solutions. Contrairement aux capteurs, les réseaux de drones ne possèdent pas une forte densité de nœuds et ils se caractérisent par une longue portée et dans un contexte plus général par une grande liberté de mobilité et une vitesse plus élevée. Ajouté à ça, leur altitude opérationnelle apporte une grande probabilité pour que les communications soient en visée directe (LOS : Line-Of-Sight). Cependant les obstacles à l'intérieur de la zone de surveillance doivent être pris en compte.

Cela rend les approches de couverture dans les drones très compliquées par rapport aux approches de couverture existantes dans les réseaux de capteurs sans fil. Le tableau ci-dessous résume les principales différences entre un réseau de capteurs et un réseau de drones :

**Tableau 3.4** – Comparaison entre le RCSF et le réseau de drones coopératifs

Critère	Le réseau de capteurs sans fil	Le réseau de drones coopératifs
Le sens de données	Sens unique (vers le puits)	Vers tous les sens
La densité	Elevée	Faible par rapport au RCSF
Le rôle	Passif	Actif
Modèle de propagation radio	NLOS (Non-Line-Of-Sight)	LOS
La portée de communication	Courte	Longue
Mobilité	Statique ou faible	Moyenne ou haute
Contrainte d'énergie	Moyenne	Moyenne (5h Fixed wing UAV) – haute (15-30 min Rotary wing UAV)

Même si le problème de couverture a été étudié en profondeur dans les réseaux de capteurs sans fil, la recherche du même problème dans les réseaux de drones est toujours en cours et d'actualité. Avant d'entamer les travaux de recherche abordant le problème de couverture propre aux UAVs, parlons d'abord de deux scénarios de couverture que les UAVs peuvent fournir.

### 3.5.1 Déploiement de plusieurs drones et planification de la trajectoire d'un seul drone

Dans les problèmes de surveillance aérienne par drones, deux scénarios différents peuvent exister : avec un drone ou avec plusieurs drones. Pour un drone, la méthode de couverture de la zone doit être la planification de la trajectoire. Le drone doit voler suivant un chemin prédéfini. La zone de surveillance est entièrement surveillée lorsque le drone vole le long de cette trajectoire et les informations sont enregistrées avec les capteurs montés sur l'UAV. Pour plusieurs drones, les drones doivent se déplacer vers des points prédéfinis pour la surveillance et la prise de courtes vidéos ou

d'images ou autres paramètres en temps réel dans toute la zone de surveillance. Les drones peuvent être déployés à partir d'une position de départ et atteindre les positions optimisées. Cela peut améliorer la précision du déploiement. Ainsi, la méthode de déploiement affecte la couverture de la zone de surveillance.

Bien que la planification de trajectoire d'un drone et le déploiement de plusieurs drones soient deux scénarios, ils partagent le même principe. Maza et al. [144] montrent que la couverture dans le cas de plusieurs drones peut être reformulée en utilisant un seul drone pour sonder une zone. Dans les deux cas, les positions que doivent prendre les drones sont déterminées par plusieurs facteurs, notamment la position initiale et la méthode d'optimisation. Cette thèse se focalise sur une solution pour trouver les meilleurs points de déploiement et explique comment cela pourrait fonctionner pour le cas de plusieurs drones.

### 3.5.2 Les stratégies de déploiement dans les drones

Dans un réseau de drones, le placement optimal a été étudié sous de nombreux points de vue, mais en particulier pour la couverture des cibles [258, 179, 257, 43, 189]. Zorbas et al. [258] ont visé la minimisation de la consommation totale d'énergie d'un ensemble de drones lorsque la couverture de tous les événements est requise. Cependant, les auteurs supposent que le nombre de drones est infini ; ils ne peuvent donc pas optimiser le coût de déploiement. Pugliese et al. [179] ont abordé le même problème mais avec une énergie initiale limitée pour les drones. L'article minimise le nombre de drones tout en couvrant toutes les cibles. Cependant, la connectivité réseau n'est pas prise en compte.

Zorbas et al. [257] abordent le problème de la minimisation des coûts de déploiement pour assurer la couverture des cibles statiques ou mobiles. Ils ont distinguées deux sous problèmes de minimisation, à savoir la minimisation du nombre de drones et la minimisation de la consommation totale d'énergie. La complexité des modèles mathématiques proposés (modèles d'optimisation linéaire en nombres entiers et non linéaire mixte en nombres entiers) a motivé les auteurs à utiliser des heuristiques qui sont Centralized-SDLP pour le déploiement statique, Centralized-MDLP et Localized-MDLP pour le déploiement mobile. Les stratégies discutées se concentrent uniquement sur le nombre de drones et la consommation d'énergie. Néanmoins, le problème de connectivité réseau n'est pas résolu.

Caillouet et Razafindralambo [43] ont considéré la connectivité comme l'un des objectifs les plus importants. Ils ont assuré la communication entre les drones et la station de base située au sol pour la collecte et l'envoi d'informations aux cibles. Ils ont fourni une solution pour minimiser l'altitude maximale avec un petit nombre d'UAVs supplémentaires en utilisant un modèle linéaire.

Le problème de déploiement a été traité dans [189] dans le but de déployer une infrastructure réseau pour servir les victimes en cas de catastrophe. Dans l'article, deux phases de déploiement sont distinguées. Dans la première, les drones sont placés en fonction d'informations partielles sur le scénario de la catastrophe à l'aide d'un algorithme génétique. Dans la seconde phase, les drones se déplacent selon l'algorithme de recherche local HCA (Hill Climbing Algorithm) afin d'explorer les victimes non couvertes. L'article était axé sur la communication et ne tenait pas compte du

nombre de drones et de la consommation d'énergie de l'ensemble des drones.

Dans [180], un modèle d'optimisation appelé BCLP-CC (Backup Coverage Location Problem with Complementary Coverage) a été développé pour trouver les emplacements optimaux pour un ensemble de sites de lancement de drones médicaux. Il a visé à maximiser deux critères qui sont la couverture primaire (qui peut être partielle) de la demande distribuée en continu (représentée sous forme de polygones) qui nécessite des services médicaux d'urgence et sa couverture secondaire (pour répondre à un autre événement dans la zone de service). Cependant, l'article n'a pas pris en compte le chevauchement et la communication.

Ruan et al. [192] ont proposé une approche basée sur la théorie des jeux pour résoudre le problème de couverture tout en considérant la connectivité entre les drones. L'algorithme de déploiement appelé MUECD-SAP (Spatial Adaptive Play based Multi-UAV Energy-efficient Coverage Deployment algorithm) est conçu pour atteindre le point d'équilibre de Nash de la méthode proposée. Les résultats de la simulation ont montré l'efficacité de leur approche. Cependant, le nombre de drones n'est pas optimisé.

Pham et al. [172] ont proposé un algorithme d'apprentissage par renforcement multi-agents (MARL : Multi-Agent Reinforcement Learning algorithm) pour fournir une couverture complète tout en minimisant les parties qui se chevauchent. Cet algorithme se focalise uniquement sur la couverture et le chevauchement et il ne prend pas en compte d'autres paramètres comme la consommation d'énergie et la connectivité qui sont vraiment importants dans un réseau de drones.

Savkin et Huang [198] ont proposé un algorithme basé sur des triangulations pour la surveillance d'une région. L'algorithme proposé vise à minimiser le nombre de drones et à trouver leurs positions. Cependant, l'article n'a pas pris en compte la consommation d'énergie, le chevauchement et la communication.

Ullmert et Fröhlich [220] ont proposé un algorithme heuristique pour résoudre le problème du déploiement statique des drones en prenant en compte le nombre de drones et la consommation d'énergie. L'algorithme proposé est basé sur deux phases. Dans la première, le nombre de drones est limité à un nombre fini. Cela a conduit à une solution non optimale. Dans la deuxième phase, la solution de la première phase est ré-optimisée en déplaçant les drones sans modifier l'ensemble des cibles observées. L'algorithme a été comparé à une méthode d'optimisation linéaire et les résultats ont montré qu'il donne de bons résultats beaucoup plus rapidement que la méthode exacte. Par contre, la connectivité n'est pas considérée.

Le problème de placement des drones a été résolu par Strumberger et al. [207] en utilisant l'algorithme EHO (Elephant Herding Optimization). Les auteurs ont visé à surveiller des cibles avec le moins de drones possible. De plus, ils ont supposé que les drones sont à la même altitude. Cependant, la consommation d'énergie et la connectivité de tous les drones ne sont pas prises en compte.

Afin de résoudre le problème de couverture d'une flotte de drones utilisée comme puits de collecte de données mobiles d'un réseau de capteurs sans fil, Tuba et al. [219] ont proposé un algorithme récent appelé Bare Bones Fireworks algorithm (BBFWA). L'objectif est de couvrir un

nombre maximum de capteurs avec un nombre minimum de drones avec une consommation de carburant minimale. Cependant, la connectivité n'a pas été abordée.

Les drones peuvent également être utilisés comme un support pour les réseaux cellulaires. [16] a proposé une méthode modélisée mathématiquement pour résoudre le problème de placement d'une station de base (BS : Base Station) montée sur un drone (drone-BS) afin de maximiser le nombre d'utilisateurs couverts avec une puissance de transmission minimale. Kalantari et al. [103] ont étudié le placement 3D optimal d'un drone-BS sur une zone urbaine en tenant compte des exigences et des limites de la liaison backhaul sans fil. Le problème a été résolu à l'aide de deux approches à savoir l'approche centrée sur le réseau et l'approche centrée sur l'utilisateur. Alzenad et al. [17] ont proposé deux algorithmes appelés respectivement Maximal Weighted Area algorithm (MWA) et Exhaustive Search algorithm (ES) pour optimiser le nombre d'utilisateurs couverts demandant différentes exigences de qualité de service. Cependant, les trois articles ont considéré un seul drone.

[64] a optimisé la densité de déploiement des DSC (Drone Small Cells<sup>1</sup>) pour avoir des performances de couverture maximales tout en tenant compte des interférences cumulées entre les cellules. Il a d'abord calculé l'interférence entre les cellules en utilisant la méthode EUDPE (Equivalent Uniform Density Plane-Entity). Ensuite, il a proposé un algorithme pour calculer la densité optimale. Plusieurs résultats ont été présentés pour valider l'analyse théorique mais il a ignoré la consommation d'énergie.

Le déploiement de drones a été également abordé dans [202] pour améliorer la couverture et la capacité des réseaux sans fil hétérogènes. Les auteurs ont proposé des solutions qui utilisent la dominance par priorité et les approches d'entropie pour résoudre deux problèmes qui sont le problème de décision de la macro station de base (qui doit décider où placer les drones) et le problème d'allocation d'UAVs coopératifs (qui se concentre sur la formation d'un réseau coopératif). Les résultats ont montré une amélioration des paramètres réseau considérés, mais l'énergie n'a pas été prise en compte.

Dans [104], un algorithme heuristique basé sur l'optimisation par essaim de particules est proposé pour trouver le nombre minimum d'UAV-BS et leurs emplacements 3D afin de servir les utilisateurs. Il a supposé qu'il n'y avait pas de BS au niveau du sol. En revanche, la consommation d'énergie de l'ensemble de drones n'est pas prise en compte.

Le Tableau 3.5 présente un résumé des stratégies de déploiement précédentes.

---

1. Les DSC sont des stations de base sans fil montées sur des drones pour fournir des services sans fil aux utilisateurs au sol.

**Tableau 3.5** – Un résumé des stratégies de déploiement des drones

Référence	Année	Application	Approche adoptée	Paramètres considérés
[207]	2017	Suivi des cibles	L'algorithme EHO	<ul style="list-style-type: none"> <li>• Nombre de drones</li> </ul>
[258]	2013	Suivi des cibles	Modèle mathématique et un algorithme localisé (Localized Altitude Scheduler)	<ul style="list-style-type: none"> <li>• Consommation d'énergie</li> </ul>
[179]	2016	Couverture de la cible	Modèles mathématiques (des modèles d'optimisation non linéaire mixte en nombres entiers et des procédures heuristiques basées sur une formulation à programmation mixte en nombres entiers)	<ul style="list-style-type: none"> <li>• Nombre de drones</li> <li>• Consommation d'énergie</li> </ul>
[257]	2016	Couverture de la cible	Modèles mathématiques (modèles d'optimisation linéaire en nombres entiers et non linéaire mixte en nombres entiers) et quelques heuristiques (Centralized-SDLP, Centralized-MDLP et Localized-MDLP)	<ul style="list-style-type: none"> <li>• Nombre de drones</li> <li>• Consommation énergétique totale</li> </ul>
[43]	2017	Couverture de la cible	Modèle linéaire	<ul style="list-style-type: none"> <li>• Connectivité</li> <li>• Nombre de drones</li> </ul>
[189]	2018	Scénario de catastrophe (au service des victimes)	GA et HCA	<ul style="list-style-type: none"> <li>• Connectivité</li> </ul>
[17]	2017	Réseaux cellulaires sans fil (pour servir les utilisateurs)	L'algorithme MWA et algorithme ES	<ul style="list-style-type: none"> <li>• Nombre d'utilisateurs couverts avec différentes exigences de qualité de service</li> </ul>
[16]	2017	Réseaux cellulaires sans fil (pour servir les utilisateurs)	Modélisé mathématiquement	<ul style="list-style-type: none"> <li>• Consommation d'énergie</li> </ul>
[103]	2017	Réseaux cellulaires sans fil (pour servir les utilisateurs)	Approches centrées sur le réseau et centrées sur l'utilisateur	<ul style="list-style-type: none"> <li>• Nombre d'utilisateurs servis</li> <li>• Somme-taux d'utilisateurs servis</li> </ul>
[104]	2016	Réseaux cellulaires sans fil (pour servir les utilisateurs)	L'algorithme PSO	<ul style="list-style-type: none"> <li>• Nombre d'UAV-BSs</li> </ul>
[219]	2018	Couverture des réseaux de capteurs sans fil (pour couvrir un nombre maximum de capteurs)	L'algorithme BBFWA	<ul style="list-style-type: none"> <li>• Nombre de drones</li> </ul>
[180]	2018	À des fins médicales (pour servir les victimes)	BCLP-CC	<ul style="list-style-type: none"> <li>• Couverture primaire et secondaire</li> </ul>

[192]	2018	Couverture de zone	Approche de la théorie des jeux	<ul style="list-style-type: none"> <li>• Communication économe en énergie</li> </ul>
[64]	2018	Réseaux cellulaires sans fil	la méthode EUDPE et un algorithme calculant la densité de déploiement optimale	<ul style="list-style-type: none"> <li>• Interférence cumulée entre les cellules</li> </ul>
[202]	2017	Environnement de communication hétérogène 5G	Une approche utilisant la dominance par priorité et les approches d'entropie	<ul style="list-style-type: none"> <li>• Couverture</li> </ul>
[172]	2018	Couverture de zone	Un algorithme d'apprentissage par renforcement multi-agents	<ul style="list-style-type: none"> <li>• Sections se chevauchant</li> </ul>
[198]	2019	Couverture de zone	Algorithme basé sur des triangulations	<ul style="list-style-type: none"> <li>• Nombre de drones</li> </ul>
[220]	2019	Couverture de la cible	Un programme linéaire et une heuristique	<ul style="list-style-type: none"> <li>• Nombre de drones</li> <li>• Consommation énergétique totale</li> </ul>

Du Tableau 3.5, nous pouvons dire que la plupart des travaux effectués dans ce domaine se concentrent sur la façon de couvrir le maximum de cibles en tenant compte d'au plus deux paramètres (énergie, nombre de drones, nombre de drones et énergie, nombre de drones et connectivité). Cependant, sans une prise en compte exhaustive de toutes les contraintes de déploiement (coût, couverture, connectivité, chevauchement, énergie), les performances ne peuvent pas être entièrement optimisées. Compte tenu de la limitation susmentionnée, la combinaison de tous ces paramètres est considérée dans ce travail.

Une nouvelle approche de déploiement de drones [11] est proposée, basée sur des méta-heuristiques, contrairement aux approches exactes telles que la programmation linéaire [257, 43] et la programmation non linéaire mixte en nombres entiers [179, 257] qui sont limitées par la complexité de calcul. Elle vise à trouver la topologie de déploiement optimale en tenant compte de quatre objectifs : minimiser la consommation d'énergie, maximiser la couverture totale, éviter les chevauchements et maintenir la connectivité.

## 3.6 Conclusion

Le déploiement des RCSF ou de drones constitue une étape nécessaire et critique dans le développement d'une application réelle. Dans ce chapitre, nous avons présenté quelques concepts de base liés au déploiement ainsi que les différents objectifs généralement considérés dans la littérature et les différentes stratégies de déploiement des RCSF et des drones. Plus précisément, nous avons catégorisé les techniques de déploiement des RCSF en suivant la classification fournie par farsi et al. [76] en trois classes qui sont : les techniques classiques, les techniques basées sur des méta-heuristiques, et les techniques basées sur l'auto-programmation. Puis, les techniques adoptées

pour le déploiement des drones sont présentées, ainsi que leurs inconvénients et avantages.

La solution la plus efficace consiste en général à utiliser le moins de nœuds (capteurs ou drones) pour couvrir une plus grande surface. Diverses solutions du problème de déploiement ont été étudiées dans la littérature. Dans la plupart, le déploiement est déterminé par une méthode déterministe ou exacte. En tant que tel, le problème ne peut généralement pas être résolu par des méthodes déterministes. Des solutions basées sur des méta-heuristiques ont été proposées mais elles ne prennent en compte que deux paramètres au maximum. Dans la suite de cette thèse, nous essaierons de donner une formulation mathématique à ce problème et de voir la possibilité d'obtenir un meilleur déploiement de plusieurs drones avec des méthodes d'optimisation méta-heuristiques en se basant sur une combinaison de quatre paramètres de performance.

# **Deuxième partie**

## **Contribution**

# Chapitre 4

## Une approche de déploiement d'une flotte de drones basée sur des méta-heuristiques

### 4.1 Introduction

Le problème de déploiement vise à identifier les emplacements optimaux pour un ensemble de nœuds afin de produire la plus grande zone de couverture possible. Les chercheurs ont étudié ce problème sous différents angles ; ainsi, il existe différents objectifs et philosophies de solution lors de sa définition et de sa résolution. Néanmoins, aucune stratégie de déploiement ne peut être adaptée à tous les scénarios d'application et exigences d'utilisation, car de nombreux objectifs sont généralement conflictuels. Le problème est généralement connu pour être NP-difficile en complexité. Il entre dans le domaine de l'optimisation multi-objectif.

L'optimisation multi-objectif est un domaine de recherche et développement important, son importance a permis le développement de multiples heuristiques et méta-heuristiques. À cet égard, de nombreuses recherches ont porté sur la conception de solutions heuristiques de déploiement de drones pour trouver des solutions sous-optimales dans un délai raisonnable. Cependant, selon la littérature actuelle, ces algorithmes s'attaquent à un certain nombre de paramètres tels que l'énergie et le nombre de drones mais prêtent peu d'attention au chevauchement et à la connectivité. Non seulement les nœuds doivent surveiller de près un certain nombre de phénomènes physiques (ou cibles), éventuellement dispersés sur une grande région géographique, mais également le réseau formé par ces nœuds doit être entièrement connecté. De plus, l'énergie dissipée par l'ensemble de nœuds doit également être prise en compte pour l'accomplissement de la mission. Par conséquent, notre approche tente à optimiser le placement des drones tout en tenant en compte quatre paramètres de déploiement, tout en considérant 3 scénarios avec un nombre de drones différent.

Dans ce chapitre nous proposons cinq méta-heuristiques à savoir MOABC, MOPSO, SPEA II, NSGA II, NSGA III pour résoudre le problème de placement des drones. Nous présentons dans un premier lieu les concepts de base, en particulier les problèmes d'optimisation multi-objectif et le concept de Pareto. Puis nous définissons le problème et nous finissons par la description des différentes méta-heuristiques utilisées.

## 4.2 Concepts de base

### 4.2.1 Problème d'optimisation multi-objectif

L'optimisation multi-objectif (également connue sous le nom d'optimisation vectorielle, optimisation multi-critère, optimisation multi-attribut ou optimisation de Pareto) est un domaine de prise de décision multi-critère, concernant des problèmes d'optimisation mathématique impliquant plus d'une fonction objectif à optimiser simultanément.

L'optimisation multi-objectif a été appliquée à de nombreux domaines de la science et de l'ingénierie, où des décisions optimales doivent être prises en présence de compromis entre deux ou plusieurs objectifs contradictoires. Dans ces cas, une étude d'optimisation multi-objectif doit être réalisée, qui fournit de multiples solutions représentant les compromis entre les fonctions objectif.

Dans [51], un problème d'optimisation multi-objectif ou multi-critère est défini comme *la recherche d'un vecteur de variables de décision qui satisfait les contraintes et optimise une fonction vectorielle dont les éléments représentent les fonctions objectif. Ces fonctions forment une description mathématique de critères de performance qui sont généralement en conflit les uns avec les autres. Par conséquent, le terme «optimiser» signifie trouver une telle solution qui donnerait les valeurs de toutes les fonctions objectif acceptables au décideur.*

Un problèmes multi-objectif implique plus d'une fonction objectif qui doit être minimisée ou maximisée. La réponse est un ensemble de solutions qui définissent le meilleur compromis entre des objectifs contradictoires.

En général, un problème d'optimisation multi-objectif permettant d'optimiser un nombre d'objectifs est écrit comme suit [118] :

$$\text{Optimiser } F(x) = (f_1(x), f_2(x), \dots, f_m(x))$$

$$\text{avec } g_k(x) \leq c_k \quad k = 1, 2, \dots, K$$

$$x_i^{(L)} \leq x_i \leq x_i^{(U)} \quad i = 1, 2, \dots, n$$

Où

$x = (x_1, x_2, \dots, x_n)$  : est le vecteur de solution dans  $X$ , qui est l'ensemble de solutions réalisables associées aux contraintes  $g_k(x)$  ainsi qu'à des limites explicites  $x_i^{(L)}$  inférieures et  $x_i^{(U)}$  supérieures dans l'espace de décision(D).

$F(x)$  présente le vecteur de fonctions objectif.

Dans un problème d'optimisation mono-objectif, la supériorité d'une solution sur les autres solutions est facilement déterminée en comparant leurs valeurs de fonction objectif. Dans le problème d'optimisation multi-objectif, même pour un problème d'optimisation multi-objectif trivial, il est peu probable qu'il existe une solution unique qui optimise simultanément chaque objectif. Dans de nombreux cas, les fonctions objectif seraient conflictuelles, et il existe un nombre (peut-être infini) de solutions optimales et la qualité d'une solution est déterminée par la notion de dominance.

Une solution est appelée non dominée si aucune des fonctions objectif ne peut être améliorée en valeur sans dégrader certaines des autres valeurs de fonctions objectif. De telles solutions sont appelées optimales de Pareto. Sans informations de préférence supplémentaires, toutes les solutions optimales Pareto sont considérées comme tout aussi bonnes.

### 4.2.2 Concept de Pareto

La plupart des problèmes d'optimisation multi-objectif du monde réel impliquent des objectifs conflictuels, il n'est peut-être pas possible d'obtenir une solution optimale unique, au lieu de cela, un ensemble de solutions qui fournit un compromis entre les fonctions objectif est obtenu où aucune solution n'est dominée par une autre solution. De telles solutions sont connues sous le nom de Pareto optimal ou solutions non dominées. L'ensemble des vecteurs correspondants dans l'espace de décision est généralement appelé ensemble non dominé ou ensemble Pareto optimal. La FIGURE 4.1 présente le concept de Pareto pour un problème d'optimisation à deux objectifs. Les notions liées aux concepts de Pareto sont définies ci-dessous :

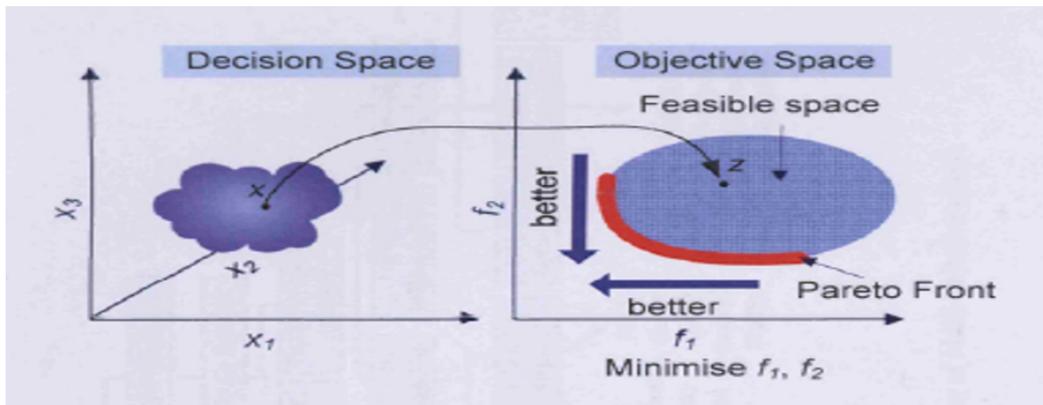


FIGURE 4.1 Le concept de Pareto pour un problème d'optimisation à deux objectifs [127]

**Définition 1 (Le concept de dominance)** On dit qu'une solution  $X$  domine une solution  $Y$  (noté  $X \preceq Y$ ) si et seulement si :  $\forall i \in [1, \dots, N] \quad x_i \leq y_i$  Et  $\exists i \in [1, \dots, N] \quad /x_i < y_i$  (À noter que cette définition pour la minimisation des objectifs)

**Définition 2 (Optimalité de Pareto)** Une solution  $x \in D$  est Pareto optimale si et seulement si il n'existe pas une solution  $x' \in D$  pour laquelle  $F(x')$  domine  $F(x)$

**Définition 3 (L'ensemble de Pareto optimal)** Pour un problème d'optimisation multi-objectif donné, l'ensemble de Pareto optimal ( $P^*$ ) est défini comme :

$$(P^*) = \{x \in D / \neg \exists x' \in D : F(x') \preceq F(x)\}$$

**Définition 4 (Le Front de Pareto)** L'ensemble de solutions dans l'espace objectif correspondant aux points de solutions dans l'ensemble optimal de Pareto est appelé front de Pareto. Le front de Pareto  $PF$  est défini comme suit :  $PF = \{\vec{u} = F(x) / x \in P^*\}$

### 4.3 Définition du problème de déploiement

Dans le domaine agricole, les cultures peuvent être infectées par des ravageurs. Ainsi, leur traitement avec des produits phytosanitaires est nécessaire pour améliorer les rendements. En raison de la vitesse, de la précision et de l'efficacité de la procédure de pulvérisation, l'utilisation de drones où un système de pulvérisation est monté sur le drone devient une demande inévitable pour le développement de l'agriculture moderne. L'intégration des drones avec un système de pulvérisation permet de fournir une plate-forme pour la lutte antiparasitaire. Il s'agit d'une application spécifique au site pour un grand champ de cultures.

Les méthodes de pulvérisation traditionnelles nécessitent non seulement d'énormes quantités de pesticides, mais ont également un impact sur l'environnement et la santé humaine. En effet, les agriculteurs pulvérisent des pesticides dans un grand champ de cultures avec des sacs de pulvérisation portés sur leurs dos. Ajouté à cela, la difficulté à les distribuer d'une manière uniforme dans tout le champ. L'application de la technologie de pulvérisation de précision sur drone devient alors une demande inévitable pour le développement de l'agriculture moderne.

Les partisans de la pulvérisation par drone citent aussi la capacité d'un drone à accéder aux zones où la topographie est un problème, comme les pentes, où la productivité de l'application manuelle est beaucoup plus faible, ou les zones basses où l'humidité du sol empêche les véhicules terrestres. L'exposition de l'opérateur est aussi réduite par rapport à une application portable.

Cependant, des zones du champ de cultures non correctement couvertes pendant la pulvérisation ou des zones de culture se chevauchent, sont des facteurs qui réduisent la qualité des pulvérisations. L'utilisation aléatoire et irrationnelle de produits phytosanitaires entraîne une toxicité en cas de surdosage et une carence dans le cas contraire.

Pour remédier à ces problèmes, le déploiement des drones doit être optimisé en couvrant toute la zone pour éviter leur non traitement tout en évitant les chevauchements de pulvérisation (c.-à-d. éviter le surdosage). Pour cela, il est supposé que la zone couverte par un drone est l'espace de pulvérisation. La communication entre les drones et même entre les drones et l'infrastructure terrestre est un autre critère à considérer. Pour assurer les communications entre les drones, le maintien de la connectivité réseau est de la plus haute importance. En effet, la détermination de la dose à appliquer sur le champ revient aux données qui devraient être collectées par l'infrastructure terrestre ou bien par les drones eux même. Ces données devraient être communiquées à un centre de contrôle ayant la capacité de déterminer la dose à appliquer. Ce dernier communique la commande aux drones qui peuvent être responsables de son exécution ou la communiquent à l'infrastructure terrestre. Un autre critère important est la consommation énergétique de l'ensemble de drones. En effet, la minimisation de la consommation réduira les coûts de déploiement et garantira le succès de la mission. La FIGURE 4.2 montre un exemple d'illustration d'un ensemble de drones couvrant une zone agricole.

Le but dans cette thèse est de déployer un ensemble de drones d'une manière optimale et de trouver les positions 3D permettant de :

- Maximiser le taux de couverture d'une zone ( $CR$ )

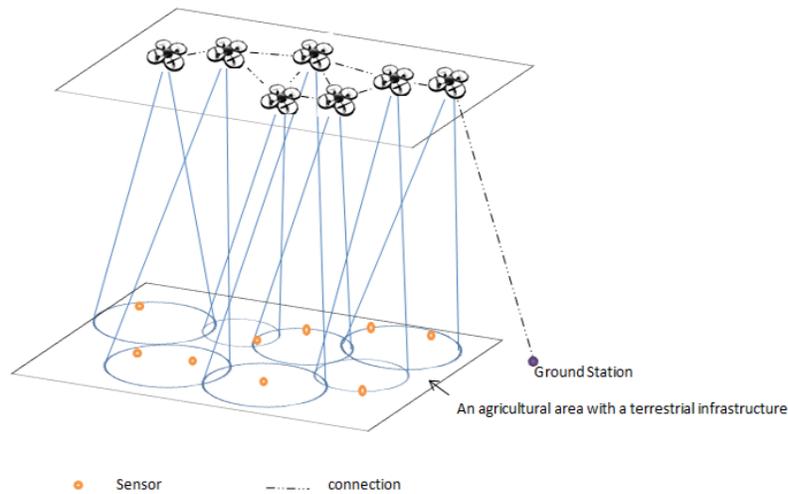


FIGURE 4.2 Un exemple de déploiement de drones dans une zone agricole

- Minimiser la consommation d'énergie sachant que cette dernière dépend de l'altitude ( $C_{sol}$ )
- Minimiser les chevauchements (assurer une qualité de la pulvérisation) ( $CH$ )
- Maintenir la connectivité et maximiser le nombre de liens de communication entre les drones ( $COM$ )

## 4.4 Description et formulation mathématique

Avant de passer à la formulation mathématique du problème de déploiement, on donne d'abord sa description comme suit :

Un ensemble de drones  $U$  est utilisé pour couvrir une zone agricole qui est un terrain plat à deux dimensions ( $x_{max} \times y_{max}$ ). Les drones sont équipés d'une caméra permettant de surveiller les plantes du champ, de modules de communication et d'un système de pulvérisation.

Chaque drone  $u \in U$  est localisé dans un espace à trois dimensions :  $P = (x_u, y_u, h_u)$  sachant que  $(x_u, y_u)$  est la position de  $u$  dans l'espace bidimensionnel (2D), et  $h_u$  son altitude. Plus l'altitude est haute plus la consommation énergétique est grande plus le rayon de couverture est large (Et vice versa). Les drones vont voler à des hauteurs comprises entre  $h_{min}$  et  $h_{max}$ .

### 4.4.1 Couverture

La couverture est souvent considérée comme étant une mesure de performance très importante. Elle reflète la façon dont une zone donnée est surveillée (contrôlée), c'est-à-dire comment chaque point de la zone de surveillance est observée et suivie par l'ensemble des nœuds.

En effet, certaines applications peuvent exiger un fort degré de couverture afin de remplir pleinement leurs missions. Par exemple, c'est le cas de l'agriculture intelligente. Comme nous l'avons

mentionné au niveau de la partie état de l'art, il existe plusieurs types de couverture : couverture de points, couverture de région et couverture de barrière. Cependant, nous considérons dans notre étude la couverture de région. Selon [76], on dit qu'une région est totalement couverte si tous les points intérieurs sont couverts.

La maximisation de la couverture est l'objectif principal de toutes les stratégies de déploiement. La zone de couverture d'un drone dans l'espace bidimensionnel est fournie par l'angle  $\theta$  et l'altitude  $h_u$ .  $\theta$  est la visibilité du drone représentée par un cercle dans l'espace  $(x, y)$  avec  $r_u$  (rayon d'observation) qui dépend de  $h_u$ .

$$r_u = h_u \times \tan\left(\frac{\theta}{2}\right) \quad (1)$$

On dit qu'un drone  $u$  couvre un point  $p(x, y)$  si la distance euclidienne entre sa projection dans le plan 2D  $(x_u, y_u)$  et ce point  $(x, y)$  est inférieure au rayon d'observation de ce drone.

$$c(u) = \begin{cases} 1, & \text{si } d(u, p) < r_u \\ 0, & \text{sinon} \end{cases} \quad (2)$$

Où

$$d(u, p) = \sqrt{(x_u - x_p)^2 + (y_u - y_p)^2}$$

Le taux de couverture du réseau de drones est calculé par l'équation suivante :

$$CR = \cup c_u / A, u \in U \quad (3)$$

$c_u$  : c'est la couverture d'un drone  $u$ .  $A$  : c'est la taille totale de la zone d'intérêt.

#### 4.4.2 Consommation d'énergie

L'énergie consommée par un réseau de drones est considérée comme un paramètre très important lors d'un déploiement. Ici, notre objectif est de minimiser l'énergie totale consommée par le réseau. La dissipation de l'énergie dans ce type de réseaux peut être vue selon différents aspects notamment l'énergie liée au fonctionnement du drone et la communication. En effet, l'énergie liée au fonctionnement du drone est nécessaire pour maintenir le drone en altitude et soutenir sa mobilité. Elle est généralement beaucoup plus élevée que la consommation d'énergie de communication [250]. Par conséquent, par rapport à l'énergie liée au fonctionnement du drone, l'énergie de communication peut être pratiquement ignorée [250]. Dans cette thèse, seule la consommation d'énergie liée au fonctionnement de l'UAV est considérée. Elle est déterminée par l'altitude de l'UAV, car cette dernière influence grandement sa consommation.

La minimisation de la consommation d'énergie de l'ensemble de drones est importante afin de garantir le succès de la mission. La consommation d'énergie d'une solution est égale à la consommation totale de tous les drones :

$$C_{\text{sol}} = \sum_{u=1}^{|U|} \text{con}_u \quad (4)$$

où la consommation énergétique d'un drone  $con_u$  dépend de son altitude  $h_u$ . [258]

$$con_u = m \times g \times h_u \quad (5)$$

$m$  : c'est la masse du drone et  $g$  : c'est l'accélération de la gravité.

### 4.4.3 Connectivité

Le taux de couverture d'une zone par des UAVs est non seulement déterminé par le nombre d'UAVs, mais il est également étroitement lié à la connectivité entre les drones.

Dans le monde des réseaux, deux types de connectivité peuvent être distingués : la connectivité simple (1-connectivité) et la connectivité multiple (k-connectivité). Une connectivité est dite simple s'il existe un seul chemin entre chaque deux nœuds et elle est dite multiple s'il existe plusieurs chemins distincts entre deux nœuds différents du réseau [229].

Selon les stratégies de placement des nœuds dans la zone de surveillance et selon les caractéristiques de l'application, une connectivité simple (ou multiple) peut être assurée lors de la phase de placement des nœuds.

Afin d'assurer la connectivité entre les drones et une station de base fixe pour former un réseau connecté et collecter efficacement des informations depuis le sol, et afin d'améliorer la résilience du réseau, nous avons considéré la k-connectivité. En effet, la surveillance en temps réel du champ générera une grande quantité de données à envoyer pour analyse, et leur envoi prendra beaucoup de temps et d'énergie en raison du fort trafic de transmission. Ainsi, l'établissement de multiples liens de connexion est vraiment intéressant pour assurer des échanges d'informations ininterrompus. Par conséquent, le temps de transmission sera réduit et potentiellement l'énergie.

Une flotte de drones forme un réseau connecté si chaque drone ( $u_i \forall u_i \in U$ ) est accessible depuis n'importe quel autre drone du réseau. C'est à dire l'existence d'un chemin entre un nœud (drone) et tout autre nœud est nécessaire pour s'assurer que tous les drones forment un réseau connecté.

Un drone ( $u_i$ ) peut communiquer avec un autre drone ( $u_j$ ) si leur distance dans le plan tridimensionnel est inférieure ou égale à la portée de communication du drone  $u_i$  ( $Rc_{u_i}$ ).

$$Com(u_i, u_j) = \begin{cases} 1, & \text{si } d(u_i, u_j) \leq Rc_{u_i} \\ 0, & \text{sinon} \end{cases} \quad (6)$$

où

$$d(u_i, u_j) = \sqrt{(x_{u_i} - x_{u_j})^2 + (y_{u_i} - y_{u_j})^2 + (h_{u_i} - h_{u_j})^2} \quad (7)$$

Le nombre de connexions est maximisé à l'aide de la formule suivante :

$$COM = \sum_{i=1}^{|U|-1} \sum_{j=2}^{|U|} Com(u_i, u_j) \quad (8)$$

Pour assurer la communication des drones avec une infrastructure terrestre, la hauteur maximale ne doit pas dépasser leur portée de communication c.-à-d. :

$$h_{\max_u} < Rc_u, \forall u \in U \quad (9)$$

#### 4.4.4 Chevauchement

Un chevauchement est défini comme une redondance de couverture dans une zone d'intérêt, c'est-à-dire, plusieurs nœuds peuvent couvrir en même temps la même surface d'une zone donnée.

Le déploiement aléatoire des drones peut induire beaucoup de chevauchements entre les nœuds. Ces chevauchements représentent une caractéristique qui est à la fois rentable et pénalisante. Ils peuvent être utilisés pour améliorer la fiabilité de la détection ou l'exactitude de la reconstruction des images collectées, par exemple, c'est le cas de la surveillance aérienne par drones. En effet, certaines applications peuvent exiger un fort degré de chevauchement des images afin d'extraire d'une manière précise des indices de végétation qui permettent aux agriculteurs de surveiller en permanence les cultures. Par contre, ils engendrent des quantités de produits agrochimiques plus importantes et impliquent une toxicité supplémentaire des cultures dans le cas des applications de pulvérisation aérienne.

Par conséquent, la nécessité de prise en compte des chevauchements pour la couverture d'une zone donnée varie en fonction des besoins applicatifs et doit être prise en considération dans la phase de conception et de déploiement. Ce constat nous motive à préciser que notre exemple applicatif dans cette thèse est la pulvérisation à base d'UAVs permettant une utilisation efficace des produits agrochimiques. Une telle application peut non seulement minimiser les quantités de produits agrochimiques mais également contribuer à la protection de l'environnement.

Notre objectif est de minimiser ce paramètre (chevauchement), ce qui améliore la qualité de la couverture en réduisant les effets indésirables d'une surdose de pesticides.

La formule (10) donne le taux de chevauchement.

$$CH = \frac{U(A_i \cap A_j)}{A}, i \neq j \quad (10)$$

où  $(A_i \cap A_j)$  est la zone d'intersection entre la zone  $(A_i)$  couverte par le drone  $i$  et la zone  $(A_j)$  couverte par le drone  $j$ .  $A$  est la taille totale de la zone d'intérêt.

La formulation mathématique finale du problème de déploiement optimal pour l'ensemble de drones est définie comme suit :

$$\left\{ \begin{array}{l} \text{Max} \quad CR = \cup c_u/A, u \in U \\ \text{Max} \quad COM = \sum_{i=1}^{|U|-1} \sum_{j=2}^{|U|} Com(u_i, u_j) \\ \text{Min} \quad C_{sol} = \sum_{u=1}^{|U|} con_u \\ \text{Min} \quad CH = \frac{\cup(A_i \cap A_j)}{A}, i \neq j \end{array} \right. \quad (11)$$

S.t.

$$\forall (x_u, y_u, h_u), 0 \leq x_u \leq x_{\max} \quad (12)$$

$$0 \leq y_u \leq y_{\max} \quad (13)$$

$$h_{\min} \leq h_u \leq h_{\max}, u \in U \quad (14)$$

$$c_u \in \{0, 1\}, Com(u_i, u_j) \in \{0, 1\}, \forall u \in U \quad (15)$$

$$h_{\max_u} < Rc_u, \forall u \in U \quad (16)$$

L'équation (11) optimise quatre fonctions objectif, elle maximise la couverture et le nombre de liaisons de communication et minimise la consommation d'énergie de tous les drones et les chevauchements. Les contraintes (12)- (15) spécifient le domaine des variables. L'équation (16) est une condition pour assurer la communication avec l'infrastructure terrestre.

## 4.5 Les méta-heuristiques utilisées

L'heuristique est l'aspect d'optimisation mathématique utilisé pour les algorithmes qui tentent de trouver une solution pour un problème donné parmi les possibles, mais ils ne garantissent pas que la meilleure sera trouvée, ils peuvent donc être considérés comme des solutions approximatives et non précises. Ces types d'algorithmes trouvent généralement une solution proche de la meilleure mais ils la trouvent rapidement. De plus, ces algorithmes ne sont pas basés sur des métriques particulières. L'heuristique est définie comme une technique qui recherche de bonnes solutions (quasi optimales) à un coût de calcul raisonnable sans pouvoir garantir l'optimalité [72]. Souvent, l'heuristique est spécifique au problème. Une méta-heuristique est définie de manière similaire, mais à un niveau d'abstraction plus élevé [90]. Autrement dit, elle est adaptable et applicable à un large ensemble de problèmes différents.

Les méta-heuristiques utilisées pour résoudre le problème multi-objectif considéré dans ce travail sont des algorithmes basés sur la population à savoir MOABC, MOPSO, SPEA II, NSGA II, et NSGA III.

### 4.5.1 MOABC

L'algorithme de colonie d'abeilles est l'une des méta-heuristiques les plus puissantes ; il est inspiré du comportement de recherche de la nourriture des essaims d'abeilles. Il a été introduit en 2005 par Karaboga [108]. La version de base traite des problèmes d'optimisation mono-objectif. Cet algorithme a été utilisé dans plusieurs domaines tels que l'optimisation des fonctions mathématiques, le clustering, la résolution de nombreux problèmes dans les RCSF, etc.

D'autres versions ont été améliorées dans le but de traiter des problèmes d'optimisation multi-objectif comme Non-dominated Sorting Artificial Bee Colony (NSABC) [118] et MOABC [259].

Dans ce travail, l'algorithme MOABC est utilisé pour optimiser le déploiement des drones. Il est basé sur la non dominance pour sélectionner les individus et sur une stratégie d'apprentissage globale pour obtenir des solutions diverses. Cet algorithme utilise une archive externe pour stocker les solutions non dominées. La distance de surpeuplement (ou d'encombrement) est également utilisée lorsque le nombre de solutions non dominées dépasse la taille de l'archive pour éliminer les individus surpeuplés. Les étapes principales de cet algorithme sont présentées dans l'Algorithme (1).

Dans cette méta-heuristique, la position de la source de nourriture représente la solution potentielle pour le problème d'optimisation et la quantité du nectar de la source de nourriture correspond aux fonctions objectif de la solution associée. Notre solution est représentée par un tableau de  $N$  éléments et chaque élément est la position d'un drone dans l'espace 3D (voir Tableau 4.1). Cet algorithme commence par une étape d'initialisation dans laquelle il évalue les positions de la source de nourriture puis il les trie en fonction de la non dominance. Les solutions non dominées sont sélectionnées et stockées dans l'archive externe (EA : External Archive). Ensuite, l'algorithme passe à une autre étape appelée phase des abeilles spectatrices. Dans cette phase une stratégie d'apprentissage globale est utilisée afin de générer des nouvelles solutions  $V_i$ . Chaque abeille apprend à partir d'une solution choisie aléatoirement dans l'EA en choisissant  $m$  dimensions. La nouvelle solution est générée en utilisant la formule suivante :

$$v_{i,f(m)} = x_{i,f(m)} + \phi(m)(EA_{k,f(m)} - x_{i,f(m)}) \quad (17)$$

Où  $K \in (1, 2, \dots, p)$  est l'indice choisi aléatoirement.  $p$  est le nombre de solutions dans l'EA.  $f(m)$  est les  $m$  premiers entiers d'une permutation aléatoire des entiers 1 à  $n$ , et  $f(m)$  définit quelles dimensions devraient apprendre à partir de l' $EA_k$ .  $\phi(m)$  produit  $m$  nombres aléatoires qui sont tous entre  $[0,2]$ . Les  $m$  nombres aléatoires correspondent aux  $m$  dimensions. Pour les dimensions restantes, d'autres solutions non dominées sont utilisées en utilisant la formule suivante :

$$v_{i,j} = x_{i,j} + \phi_{i,j}(EA_{l,j} - x_{i,j}) \quad (18)$$

où  $l \neq k, j \in (1, 2, 3, \dots, p)$  et  $j \notin f(m)$ .

---

**Algorithme 1: Les principales étapes de l'algorithme MOABC**

---

1. cycle = 1
2. Initialiser les positions des sources de nourriture (solutions)  $x_i, i = 1, \dots, SN$
3. Évaluer la quantité de nectar (fitness) des sources de nourriture
4. Les solutions initialisées sont triées en fonction de la non dominance
5. Stocker les solutions non dominées dans une archive externe EA
6. **répétez**
7. Phase des abeilles spectatrices
  - **Pour** chaque abeille spectatrice
    - Choisir au hasard une solution dans EA
    - Produire une nouvelle solution  $v_i$  en utilisant l'expression (17)
    - Vérifier  $V_i$ , pour rester dans la zone de couverture.
    - Calculer le  $fit_i$ .
    - Appliquer le mécanisme Greedy selection (l'Algorithme (2)) pour décider quelle solution entre dans EA
  - **Fin pour**
8. Les solutions qui sont dans l'EA sont triées sur la base de la non dominance
9. Garder les solutions non dominées dans l'EA
10. Si le nombre de solutions non dominées dépasse la taille allouée de EA, Utiliser la distance d'encombrement pour supprimer les membres surpeuplés
11. cycle = cycle + 1
12. **jusqu'au** cycle = nombre maximum de cycles

---

Une fois la nouvelle solution est produite, elle est évaluée et un mécanisme de sélection est appliqué (voir Algorithme (2)) pour sélectionner les solutions qui entrent à l'EA selon la non dominance.

En se basant sur la non dominance, les solutions de l'archive externe sont triées et celles qui ne sont pas dominées sont gardées dans l'EA. La distance d'encombrement est également utilisée lorsque le nombre de solutions non dominées dépasse la taille de l'archive pour supprimer les membres surpeuplés.

---

**Algorithme 2:** Algorithme de sélection glouton (Greedy selection)

---

- Si  $v_i$  domine  $x_i$ 
    - Mettez  $v_i$  dans EA
  - Sinon si  $x_i$  domine  $v_i$ 
    - Ne fais rien
  - Sinon si  $x_i$  et  $v_i$  ne sont pas dominés l'un par l'autre
    - Mettez  $v_i$  dans EA
    - Produire un nombre aléatoire  $r$  tiré d'une distribution uniforme sur l'intervalle unitaire
    - Si  $r < 0.5$ 
      - La solution originale est remplacée par une nouvelle solution en tant que nouvelle position de source d'alimentation.
      - Ce que signifie que  $x_i$  est remplacé par  $v_i$ .
    - Sinon
      - Ne fais rien
    - Finsi
  - Finsi
- 

## 4.5.2 MOPSO

MOPSO est une amélioration de l'algorithme PSO pour traiter les problèmes d'optimisation à plusieurs objectifs. Il est inspiré du comportement social des animaux (comme les oiseaux) au sein d'un troupeau. Cet algorithme utilisé dans cette étude est proposé par coello [52].

Le troupeau présente la population et la particule représente la solution. Chaque particule explore l'espace de recherche en se basant sur sa vitesse et sa position sachant que l'espace de recherche est affecté par son historique et d'autres particules. L'Algorithme (3) présente les étapes principales de cette méta-heuristique.

---

**Algorithme 3: Les grandes lignes de l'algorithme MOPSO**

---

1. Initialiser la population Pop
  2. Initialiser la vitesse de chaque particule
  3. Évaluer chaque particule de la population
  4. Stocker les positions des particules qui représentent les vecteurs non dominés dans le répertoire Rep
  5. Générer les hypercubes de l'espace de recherche exploré jusqu'à présent, et localiser les particules en utilisant ces hypercubes comme un système de coordonnées où les coordonnées de chaque particule sont définies en fonction des valeurs de ses fonctions objectif
  6. Initialiser la mémoire de chaque particule
  7. **Tant que** (*it*  $\neq$  *maxit*) **faire**
    - Calculer la vitesse de chaque particule en utilisant la formule (19)
    - Calculer les nouvelles positions des particules en ajoutant la vitesse (formule (21))
    - Vérifier que les particules sont dans l'espace de recherche
    - Évaluer chaque particule dans Pop
    - Mettre à jour le contenu du répertoire Rep et de la représentation géographique des particules dans les hypercubes.
    - Mettre à jour la mémoire des particules (selon la non dominance)
    - Incrémenter le compteur de la boucle
- Fin tant que**
- 

L'algorithme commence par la génération des particules de la population qui sont caractérisées par leurs positions et leurs vitesses. Puis ces particules sont évaluées en utilisant le concept de dominance. Les solutions non dominées sont stockées dans le répertoire externe Rep. Ce répertoire a pour but de sauvegarder les solutions non dominées trouvées tout au long de la recherche. En fait ce répertoire est basé sur 2 composants qui sont : le contrôle de l'archive et une grille. Le premier décide de la solution qui sera entrée ou non au répertoire selon le principe suivant : les solutions non dominées trouvées dans la population à chaque itération sont comparées (une par une) en respectant le contenu de l'archive Rep. Si la nouvelle solution est dominée par n'importe quelle solution de l'archive, donc elle sera ignorée sinon elle sera ajoutée. S'il existe des solutions dans Rep qui sont dominées par un nouveau individu, donc cette solution est supprimée de l'archive. Puisque la taille de l'archive est limitée, l'ajout fait appel à la grille.

La grille permet d'avoir des fronts de Pareto bien distribués. L'espace des objectifs est divisé en régions. Lorsqu'une nouvelle solution est trouvée en dehors de la grille actuelle, la grille sera mise à jour et les solutions qui sont à l'intérieur seront relocalisées. Elle est considérée comme un ensemble d'hypercubes dont chaque hypercube peut être vu comme une zone géographique contenant des individus ou vide.

La prochaine étape de l'algorithme après avoir initialisé la mémoire des particules qui sert à guider la recherche, est le calcul de la vitesse de chaque particule déplaçant vers  $Pbp$  et  $Lp$  en utilisant la formule suivante :

$$v_i^{t+1} = w_i * v_i^t + c_1 * r_1 * (Pbp_i^{t+1} - Ppos_i^t) + c_2 * r_2 * (Lp^{t+1} - Ppos_i^t) \quad (19)$$

où  $c_1, c_2$  : les constantes d'accélération,  $w_i$  : poids d'inertie d'une particule

$$w_i = 0.9 - \left(\frac{it}{MaxIt}\right) * 0.5 \quad (20)$$

$t$  : l'itération actuelle de l'algorithme,  
 $r_1, r_2$  : des nombres aléatoires dans l'intervalle [0..1],  
 $Pbp$  : la meilleure position d'une particule,  
 $Lp$  : la position du leader (globale),  
 $Lp$  est choisie d'une manière aléatoire dans un hypercube contenant moins d'individus en utilisant le mécanisme de sélection de roulette.

**Remarque :** Le calcul de la vitesse d'une particule signifie le calcul de la vitesse de chacune de ses dimensions.

Une fois la vitesse est calculée, la position est mise à jour selon la formule suivante :

$$X_i^{t+1} = X_i^t + V_i^{t+1} \quad (21)$$

Il se peut qu'une particule tombe en dehors de l'espace de recherche, donc il faut éviter de générer ce type de solutions en changeant la position de la particule pour la maintenir dans l'espace de recherche.

L'algorithme passe à l'évaluation des particules et la mise à jour de l'archive Rep et de la grille puis de la mémoire des particules en se basant sur le concept de dominance. Il finira par l'incrémentement du compteur et recommencement d'une autre itération si la condition d'arrêt n'est pas satisfaite.

**Note :** l'opérateur de mutation est aussi appliqué afin d'éviter que l'algorithme converge vers un faux front de Pareto (c'est équivalent à l'optimum local dans une optimisation globale). Le nombre de particules affecté par l'opérateur de mutation se réduit à chaque itération.

### 4.5.3 SPEA II

SPEA II est un algorithme élitiste proposé par Zitzler et al. [256]. Il utilise une stratégie d'attribution de fitness qui prend en compte le nombre d'individus dominés et dominants pour chaque individu, une méthode améliorée pour la troncature d'archives pour préserver les solutions limites et une technique de sélection basée sur la densité pour un guidage de recherche plus précis. L'archive enregistre l'ensemble non dominé et elle est séparée de la population contenant les solutions candidates. Cet algorithme est décrit ci-dessous.

---

**Algorithme 4: SPEA II**

---

1. Initialisation : Générer une population initiale et créer une archive vide
  2. L'assignement du fitness : Calculer la valeur de fitness pour les individus de la population et de l'archive
  3. Sélection environnementale :
    - Copier tous les individus non dominés de l'archive actuelle et de la population dans l'archive future
    - Si la taille de l'archive future  $>$  à la taille de l'archive ; réduire la taille de l'archive future
    - Si la taille de l'archive future  $<$  à la taille de l'archive ; compléter avec des individus dominés
  4. Terminaison : Si  $it > maxit$  ou un autre critère d'arrêt est satisfait, alors définissez l'ensemble des individus non dominés. Arrêtez.
  5. Sélection d'accouplement : Effectuer une sélection du tournoi binaire sur l'archive future afin de remplir le pool de reproduction.
  6. Variation : Appliquer des opérateurs de recombinaison et de mutation au pool d'accouplements et régler  $P_{t+1}$  sur la population résultante. Incrémentez le compteur de génération et passez à l'étape 2 .
- 

L'algorithme commence par la génération de la population initiale et la création de l'archive vide. Puis, pour chaque individu  $i$  de la population et de l'archive, une valeur de force ( $S_i$ ) est lui assignée (  $S_i$  : c'est le nombre de solutions qu'il domine). Sur la base de  $S$ , la valeur brute de fitness de chaque individu est calculé :

$$R(i) = \sum S(j) \quad (22)$$

(c'est la somme de  $S$  de toutes les solutions qui dominent l'individu  $i$ )

Lorsque la valeur de  $R = 0$  , ça correspond à un individu non dominé et lorsqu'il a une grande valeur signifie que l'individu est dominé par d'autres individus. Pour distinguer entre les individus qui ont le même  $R$ , la densité  $D(i)$  correspondant à  $i$  est utilisée et elle est définie comme suit :

$$D(i) = \frac{1}{\sigma_i^k + 2} \quad (23)$$

Pour chaque individu  $i$  , les distances (dans l'espace des objectifs) à tous les autres individus dans l'archive et la population sont calculées, puis triées en ordre croissant. Le  $k$  ème élément donne la distance recherchée, noté  $\sigma_i^k$  .

$$k = \sqrt{\text{taille de la population} + \text{taille de l'archive}} \quad (24)$$

Ensuite l'algorithme calcule l'adaptation (fitness) de l'individu  $i$  en additionnant  $R(i)$  avec  $D(i)$  :

$$F(i) = R(i) + D(i) \quad (25)$$

Dans la prochaine étape de l'algorithme, tous les individus non dominés (qui ont  $F < 1$ ) sont copiés de l'archive et de la population à l'archive de la prochaine génération. Si le nombre d'individus non dominés égal à la taille de l'archive, cette étape est terminée. Sinon si il est inférieur à la taille de l'archive, alors l'archive de la prochaine génération est complétée avec les meilleurs individus dominés de l'archive et de la population précédentes. Si il est supérieur à la taille de l'archive, alors une procédure de troncature d'archives qui supprime itérativement des individus jusqu'à ce que le nombre soit égal à la taille de l'archive est utilisée. (C'est-à-dire que l'individu qui a la distance minimale à un autre individu est choisi)

Si le critère d'arrêt n'est pas atteint, alors le sélecteur de tournoi binaire, la combinaison et la mutation sont utilisés pour produire des individus à ajouter à la population de la prochaine génération puis on incrémente le compteur de génération. Sinon on sort avec l'ensemble non dominé de l'archive de la prochaine génération.

#### 4.5.4 NSGA II

Cet algorithme a été proposé par [61] comme une version améliorée de NSGA [206]. Il est basé sur le tri non dominé et la distance d'encombrement. Les étapes de cet algorithme sont présentées dans l'Algorithme (5).

Dans cet algorithme, une population initiale est créée d'une manière aléatoire puis évaluée et triée selon le niveau ou le rang de non dominance (1 est le meilleur niveau, 2 est le prochain meilleur niveau, etc.) et la distance d'encombrement (c.à.d. que les solutions non dominées sont préférées aux solutions dominées, par contre entre deux solutions ayant le même rang, celle qui est dans la région moins peuplée est préférée).

Le sélecteur de tournoi binaire avec les opérateurs de recombinaison et de mutation sont ensuite utilisés pour créer une population de descendants Q. Cette dernière est combinée avec la population parent P et un tri non dominé est effectué. Par la suite, les solutions de meilleurs ensembles non dominés sont sélectionnées pour se propager à la génération suivante. Puisque la taille de la population est limitée et peut ne pas accueillir toutes les solutions d'un ensemble non dominé, une stratégie de surpeuplement est utilisée pour identifier les solutions qui résident dans une zone moins encombrée. La procédure de NSGA II est décrite schématiquement sur la FIGURE 4.3.

---

**Algorithme 5: NSGA II**

---

1. Générer une population initiale et l'évaluer
2. Appliquer le tri non dominé sur la population
3. Calculer la distance d'encombrement
4. Trier la population selon l'ordre décroissant de la distance d'encombrement et le rang
5. **Tant que** ( $it \neq maxit$ ) **faire**
  - Effectuer le croisement et la mutation et ajouter les individus à la population
  - Appliquer le tri non dominé
  - Calculer la distance d'encombrement
  - Trier la population
  - Tronquer la population pour garder Npop (taille prédéfinie de la population)
  - Appliquer le tri non dominé
  - Calculer la distance d'encombrement
  - Trier la population

**Fin tant que**

---

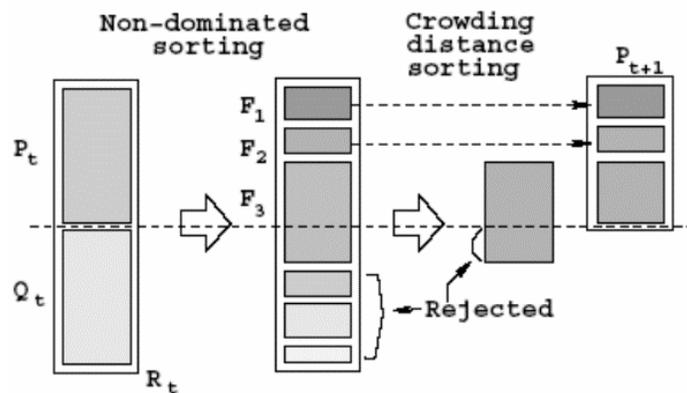


FIGURE 4.3 Un schéma décrivant la procédure NSGA II [61]

**La procédure de calcul de la distance d'encombrement pour toutes les solutions dans un ensemble non dominé :**

1. Appeler le nombre de solutions dans  $F$  comme  $l = |F|$
2. Pour chaque  $i$  de l'ensemble, attribuer d'abord la distance de surpeuplement,  $d_i = 0$ .
3. pour chaque objectif  $m = 1, 2, \dots, M$ , trier l'ensemble ou/et trouver le vecteur d'indices triés :

$$F_{\text{trié}}(m), Ind^m = \text{sort}(F(m)) \quad (26)$$

4. pour  $m = 1, 2, \dots, M$ , attribuer une grande distance aux solutions aux limites :

$d_{Ind_1^m} = d_{Ind_l^m} = \infty$  et pour toutes les autres solutions  $j = 2$  à  $(l - 1)$ , affectez :

$$d_{Ind_j^m} = d_{Ind_j^m} + \frac{F(m)^{Ind_{j+1}^m} - F(m)^{Ind_{j-1}^m}}{F(m)^{max} - F(m)^{min}} \quad (27)$$

#### 4.5.5 NSGA III

L'algorithme génétique de tri non dominé III est proposé par Deb et Jain [60] pour s'attaquer à des problèmes d'optimisation multi-objectif. Il s'agit d'un algorithme élitiste basé sur le Framework NSGA II avec des modifications dans son opérateur de sélection. Dans cet algorithme, la distance de surpeuplement n'est pas utilisée. Il utilise un certain nombre de points de référence bien répartis afin de maintenir la diversité dans les solutions obtenues. L'Algorithme (6) montre la procédure de NSGA III à une génération  $t$ .

---

#### Algorithme 6: NSGA III

---

**Entrée**  $P_0$  (population initiale);  $N$  (la taille de la population)

**Sortie**  $P_{t+1}$

1.  $S_t = \emptyset, i = 1$
2.  $Q_t = \text{croisement} + \text{mutation}$  de  $P_t$
3.  $R_t = P_t \cup Q_t$
4. Appliquer le tri non dominé sur  $R_t$  et trouver  $F_1, F_2, F_3, \dots$
5. **Tant que**  $|S_t| \leq N$  **do**

$$S_t = S_t \cup F_i$$

$$i = i + 1$$

**Fin tant que**

6. **si**  $|S_t| = N$  **faire**

$$P_{t+1} = S_t, \text{ Interruption}$$

**Sinon**

$$P_{t+1} = \cup_{j=1}^{l-1} F_j$$

Normaliser les objectifs et créer l'ensemble de référence

Associer chaque membre de  $S_t$  à un point de référence

Choisir  $N - |P_{t+1}|$  membres à partir de  $F_l$  en utilisant l'opérateur de conservation de niche

**Fin si**

7.  $t = t + 1$

8. Fin
-

A la deuxième étape de l'algorithme, les individus générés par croisement et mutation  $Q_t$  sont combinés avec la population parent  $P_t$ . Puis le tri basé sur la non dominance est appliqué pour classer  $R_t$  en différents niveaux de non dominance ( $F_1, F_2, \dots$ ).

Une fois l'algorithme a trouvé les différents niveaux de non dominance, il passe à la génération de la prochaine génération  $p_{t+1}$ . L'algorithme commence l'ajout des membres à partir de haut niveau de non dominance à  $S_t$  jusqu'à ce que sa taille soit égale à  $N$  ou la dépasse pour la première fois (par exemple le rang  $l$ ). Tous les individus des niveaux de non dominance supérieurs à  $l$  sont ignorés et les individus  $S_t \setminus F_l$  sont sélectionnés comme la prochaine génération  $p_{t+1}$ . Dans le cas où la taille de  $S_t = N$ , l'algorithme passe à la prochaine itération avec  $p_{t+1} = S_t$ . Dans le cas contraire,  $N - |p_{t+1}|$  membres sont choisis dans  $F_l$  en se basant sur les points de référence. Comme les objectifs peuvent être basés sur différentes échelles, ils sont normalisés et des points de référence sont générés dans l'espace normalisé. Puis chaque membre est affecté à un point de référence. Chaque membre de  $S_t \setminus F_l$  et  $F_l$  est ensuite affecté au point de référence le plus proche et les membres  $N - |p_{t+1}|$  restant dans  $F_l$  sont choisis de manière à ce que leur point de référence associé n'ait aucun membre associé dans  $S_t \setminus F_l$ .

► **Normalisation des objectifs :**

Chaque objectif est normalisé comme suit :

$$f_j^{norm} = \frac{f_j - z_j^{min}}{a_j}, \text{ Pour } j = 1, 2, \dots, M \quad (28)$$

$a_j$  : le point d'interception de l'objectif  $f_j$

$f_j^{norm}$  : l'objectif normalisé de  $f_j$

$z_j^{min}$  : la valeur minimale pour chaque objectif  $j = 1, 2, \dots, M$

► **Génération de points de référence :**

Les points de référence sont utilisés pour sélectionner des individus et maintenir la diversité de la population. Les points de référence peuvent être donnés par un expert ou générés systématiquement (dans ce travail, ils sont générés en utilisant la méthode de **Das et Dennis** [56], selon cette méthode ils sont placés dans un hyperplan normalisé où le nombre de points de référence dépend de nombre d'objectifs et de division).

Chaque membre de  $S_t$  est associé au point de référence qui a la distance euclidienne la plus proche de celui-ci.

► **L'opérateur de conservation de niche :**

Une fois chaque membre est associé au point de référence, certains points peuvent avoir un ou plusieurs membres associés et certains peuvent ne pas avoir de membres associés. Les membres de

$S_t \setminus F_l$  associés au point de référence  $j$  sont comptés et représentés par un compteur de niche  $P_j$ . Pour avoir une population diversifiée, la procédure suivante est utilisée : Premièrement les points de référence avec  $P_j$  minimal sont identifiés ; dans le cas où il y a plus d'un point de référence, un est choisi aléatoirement.

$P_j = 0$  indique que aucun membre de  $S_t \setminus F_l$  est associé avec le point de référence  $j$ . Dans ce cas on a deux situations : 1) il y a un ou plusieurs membre associés à ce point de référence dans  $F_l$ . Par conséquent, ajouter le membre ayant une distance perpendiculaire minimale à la prochaine génération  $p_{t+1}$  et incrémenter le compteur  $P_j$ . 2) il n'y a pas de membres associés à ce point de référence dans  $F_l$ . Donc, aller à un autre point de référence avec  $P_j$  minimal. Si  $P_j \geq 1$  le point de référence  $j$  a plus d'un membre associé et dans ce cas, si il y a un membre dans  $F_l$  associé au point de référence  $j$ , ajouter l'individu avec une distance perpendiculaire minimale à la génération suivante et incrémenter  $P_j$  de 1. Cette procédure est répétée jusqu'à ce que la taille de la population atteigne  $N$ .

## 4.6 Adaptation des méta-heuristiques

Dans la section précédente, nous avons décrit le principe général des algorithmes d'optimisation multi-objectif (MOABC, MOPSO, SPEA II, NSGA II, NSGA III). Dans cette section, nous allons procéder à leur adaptation pour résoudre le problème de déploiement abordé dans cette thèse. Lors de l'implémentation des algorithmes, le codage réel des solutions a été adopté.

### 4.6.1 Représentation d'une solution

Une solution potentielle dans l'espace  $X$  au problème de déploiement d'une flotte de drones est représentée par un individu ou un chromosome ou une source de nourriture. Elle est présentée par un tableau de taille fixe correspondant au nombre de drones ( $N$ ). Le tableau contient les positions des drones dans l'espace 3D et qui sont soumises à des limites explicites inférieures et supérieures décrites dans la partie modélisation (voir Tableau 4.1).

**Tableau 4.1** – Le tableau de la solution

1	2	...	N
$(x_1, y_1, h_1)$	$(x_2, y_2, h_2)$	...	$(x_n, y_n, h_n)$

Un point dans l'espace des objectifs correspondant à une solution  $x$  représente les valeurs des fonctions objectif de cette solution.

## 4.6.2 Initialisation de la population

Elle représente la première étape de différents algorithmes méta-heuristiques. Elle permet d'initialiser les différentes solutions potentielles à l'aide de l'opérateur d'initialisation aléatoire. Ce dernier attribue des positions aléatoirement dans l'espace à couvrir limité par des bornes inférieures et supérieures. Une fois qu'une solution est initialisée, elle est évaluée en calculant ses valeurs avec les fonctions objectif. Dans le cas où une solution viole une contrainte, elle sera pénalisée par une mauvaise valeur du critère considéré.

## 4.6.3 Les opérateurs de reproduction

Pour les algorithmes génétiques, les opérateurs génétiques suivants sont utilisés :

### 4.6.3.1 Le croisement

La population dans un algorithme génétique est intensifiée à l'aide de cet opérateur. Le croisement permet d'avoir plusieurs solutions en mélangeant les informations de deux chromosomes ou individus. Dans le cadre de notre problème, le croisement se fait en un seul point comme illustré dans la FIGURE 4.4. Le croisement en un seul point consiste à choisir aléatoirement un point de coupure qui divise les informations des deux parents en deux segments (vert et blanc). La génération de 2 nouveaux individus est effectuée en combinant le segment vert du parent 1 avec le segment vert du parent 2 et le segment blanc du parent 1 avec le segment blanc du parent 2.

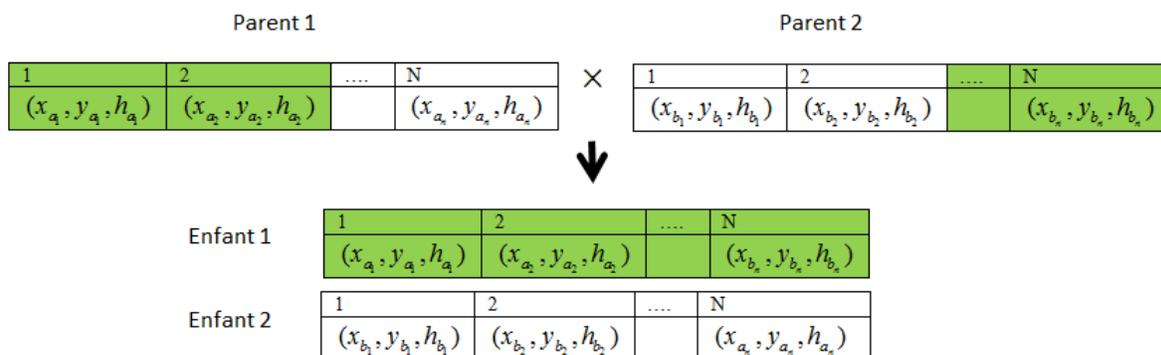


FIGURE 4.4 Étape de croisement

### 4.6.3.2 La mutation

Cet opérateur permet d'avoir des changements spontanés et aléatoires dans divers individus. L'obtention d'une mutation au sein d'un individu pourrait être effectuée par un moyen simple qui sert à modifier une ou plusieurs positions dans l'individu. Le nombre de positions concernées par le changement est en fonction du taux de mutation et de la taille de l'individu. Exemple d'une mutation à un point pour un individu (voir FIGURE 4.5).

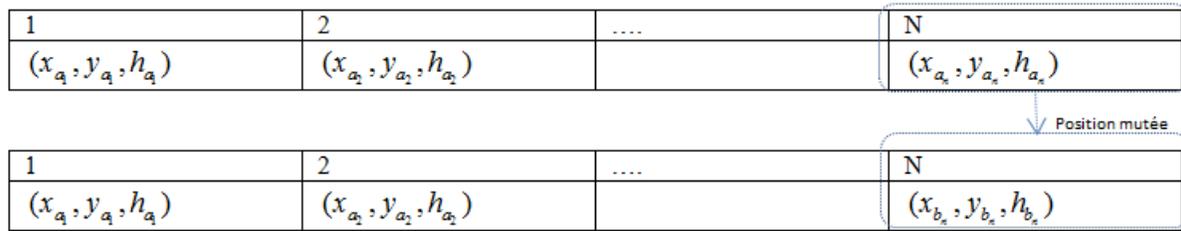


FIGURE 4.5 L'opérateur de mutation à un point

## 4.7 Conclusion

Dans ce chapitre, nous avons discuté l'optimisation multi-objectif, dans laquelle plus d'un objectif est optimisé simultanément. Nous avons commencé avec les concepts de base dont nous avons illustré les points essentiels de l'optimisation multi-objectif, en particulier le front de Pareto dans l'espace des critères. Ensuite, nous avons proposé une formalisation du problème de placement d'une flotte de drones coopératifs pour couvrir une zone agricole comme un problème d'optimisation multi-objectif qui permet d'optimiser quatre fonctions objectif. Nous avons également revu certaines méta-heuristiques que nous avons adaptées pour la résolution de notre problème. Différentes expérimentations ont été menées dans le chapitre suivant afin d'évaluer cette approche.

# Chapitre 5

## Évaluation des performances de l'approche proposée

### 5.1 Introduction

Dans le chapitre précédent nous avons présenté une approche basée sur des méta-heuristiques afin de résoudre le problème multi-objectif. Ce chapitre est entièrement dédié à l'évaluation et à l'exposition des différents résultats. Tout d'abord, trois scénarios ont été créés en variant le nombre de drones. Le premier avec 8 drones, le deuxième avec 12 drones et le troisième avec 15 drones. Puis, les approximations du front de Pareto obtenues par les différents algorithmes d'optimisation multi-objectif pour chaque scénario sont comparées en termes de la diversité et de la convergence en utilisant deux indicateurs de performance à savoir la métrique IGD et la métrique HV. Enfin, les coordonnées parallèles sont utilisées afin de visualiser les solutions fournies.

### 5.2 Description des scénarios

Un réseau de drones est utilisé pour couvrir une zone agricole de taille  $10\,000m^2$ . Chaque drone a une portée de communication de  $60\,m$ , une masse de  $10\,Kg$ , une gravitation  $g$  de  $10\,m/sec^2$ , une altitude comprise entre  $h_{min}$  et  $h_{max}$  qui sont égales à  $10\,m$  et  $45\,m$  respectivement. L'angle de pulvérisation de drone est égal à  $60^\circ$ . La portée de pulvérisation de chaque drone dépend de son altitude.

Dans ce travail, trois scénarios ont été créés. Le premier avec 8 drones, le deuxième avec 12 drones et le troisième avec 15 drones.

### 5.3 Paramètres de simulation

Les méta-heuristiques utilisées dans ce travail pour résoudre le problème multi-objectif sont des algorithmes basés sur la population. Ces algorithmes à savoir MOABC, MOPSO, NSGA II, SPEA II et NSGA III sont implémentés sous l'environnement MATLAB. Les simulations ont été

effectuées sur une machine Intel(R) Core(TM) i7, 2.7GHz, avec une RAM de 8 Go.

Pour tous les scénarios, les paramètres de contrôle sont les suivants : la taille des populations de MOABC, MOPSO et SPEA II est fixée à 50. La taille des populations de NSGA II et NSGA III, la taille des archives de MOABC et SPEA II et la taille du répertoire de MOPSO sont fixées à 200. Les autres réglages de paramètres sont définis tels qu'ils sont dans leurs études d'origine ou tel que recommandé dans leurs études d'origine.

**Tableau 5.1** – Configuration des paramètres des algorithmes

Algorithme	Paramètres
MOABC	Taille de la population = 50, taille de l'archive = 200, $it_{\max} = 5000$ .
MOPSO	Taille de la population = 50, taille du répertoire = 200 particules, $it_{\max} = 5000$ .
NSGA II	Taille de la population = 200, la probabilité de croisement = 0.9, la probabilité de mutation = $1/\text{la longueur d'un individu}$ , $it_{\max} = 5000$ .
NSGA III	Taille de la population = 200, la probabilité de croisement = 1, la probabilité de mutation = $1/\text{la longueur d'un individu}$ , $it_{\max} = 5000$ .
SPEA II	Taille de la population = 50, taille de l'archive = 200, la probabilité de croisement = 0.9, la probabilité de mutation = $1/\text{la longueur d'un individu}$ , $it_{\max} = 5000$ .

## 5.4 Les métriques d'évaluation

Le but des algorithmes d'optimisation multi-objectif visant à optimiser simultanément plusieurs objectifs conflictuels est de trouver un ensemble de solutions de compromis de haute qualité. Un tel ensemble est appelé approximation du front de Pareto. En raison de classement partiel fourni par la Pareto-optimalité entre les fronts d'approximation, leur qualité relative est évaluée en utilisant des métriques de qualité, également appelées indicateurs de qualité ou indicateurs de performance, permettant d'avoir un classement complet.

De nombreux indicateurs de performance ont été proposés dans la littérature pour mesurer différents critères souhaitables dans les approximations de haute qualité telles que la convergence et la diversité [46, 129]. Étant donné que chaque indicateur donne une importance différente aux diverses caractéristiques souhaitables des fronts d'approximation, plusieurs indicateurs de qualité peuvent être utilisés ensemble. La convergence et la diversité sont les deux objectifs visés en plus de l'optimalité de Pareto dans une approximation de qualité [46]. Le premier vise à découvrir des solutions aussi proches que possible des solutions Pareto-optimales, et le second vise à trouver des solutions aussi diverses que possible dans le front non dominé obtenu.

Par conséquent, pour comparer les différents algorithmes méta-heuristiques présentés dans cette thèse, deux indicateurs de performance à savoir l'hypervolume également appelé métrique S et la métrique de distance générationnelle inversée sont utilisés. Ces indicateurs permettent de mesurer simultanément la convergence et la diversité.

Le front de Pareto optimal du problème traité dans cette thèse est inconnu, et comme il est né-

cessaire pour le calcul des métriques de performance, un front de Pareto de référence est construit en sélectionnant des solutions non dominées parmi toutes les solutions obtenues par les différents algorithmes dans différentes exécutions avec différents paramètres.

**Distance générationnelle inversée**[101] : Soit  $P$  un ensemble de points uniformément répartis sur le front de Pareto ( $PF$ ) et  $S$  une approximation du front de Pareto. La distance générationnelle inversée est définie comme suit :

$$IGD(P, S) = \frac{(\sum_{i=1}^{|p|} d_i^q)^{\frac{1}{q}}}{|p|} \quad (1)$$

où

$d_i = \min_{\vec{s} \in S} \|F(\vec{p}_i) - F(\vec{s})\|$ ,  $\vec{p}_i \in P$ ,  $q = 2$ , et  $d_i$  est la plus petite distance de  $\vec{p} \in P$  aux solutions les plus proches dans  $S$ . Une petite valeur d'IGD indique une bonne convergence et diversité.

**Hypervolume** [30] : Cet indicateur estime l'hypervolume de la partie de l'espace objectif qui est dominé par un ensemble d'approximation. Plus la valeur HV est élevée, meilleur est le résultat. L'indicateur HV peut être exprimé comme suit :

$$HV = \bigcup_i vol_i \mid i \in PF \quad (2)$$

où  $vol_i$  correspond à une hyper zone délimitée par un point de référence prédéfini et une solution  $i$ . Le point de référence peut simplement être trouvé en construisant un vecteur des pires valeurs des fonctions objectif.

## 5.5 Résultats des simulations

Dans la première partie de cette section, nous essayons d'optimiser tous les objectifs définis dans le chapitre 4 pour voir les compromis entre eux. Deuxièmement, nous présentons la comparaison entre les cinq algorithmes méta-heuristiques.

### 5.5.1 Partie 1

Pour voir les compromis entre tous les objectifs, les expériences sont exécutées 10 fois pour tous les algorithmes et les valeurs moyennes de tous les objectifs sont calculées. Les résultats des expériences sont présentés dans la FIGURE 5.1.

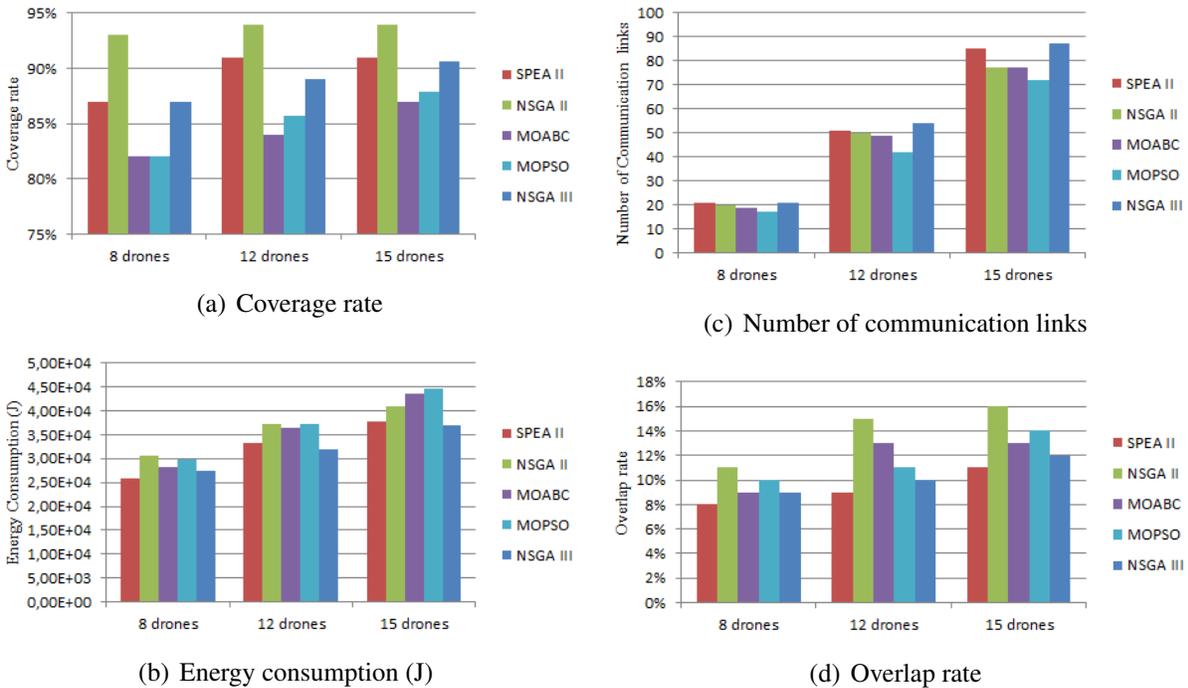


FIGURE 5.1 Valeurs des objectifs lors de l'optimisation de tous les objectifs ensemble

Pour voir l'influence du nombre de drones, des expériences ont été réalisées avec 3 scénarios (8, 12, 15 drones). La FIGURE 5.1(a) montre le taux de couverture après optimisation. De toute évidence, l'augmentation du nombre de drones peut augmenter la couverture. La couverture totale augmente pour les différentes méta-heuristiques. La FIGURE 5.1(a) montre que NSGA II, SPEA II et NSGA III peuvent augmenter la couverture totale de 4 à 10%, ce qui améliorera considérablement l'efficacité de la couverture. Par rapport à SPEA II et NSGA III, NSGA II a de meilleures performances dans tous les cas (8, 12, 15 drones). NSGA II peut fournir environ 93% de couverture même en utilisant 8 drones. Cependant, pour SPEA II, au moins 12 drones sont requis pour répondre à l'exigence de couverture de 91% et NSGA III nécessite au moins 15 drones pour répondre à cette exigence. Pour d'autres (MOABC, MOPSO) même avec 15 drones, seulement environ 87% de la surface est couverte, ce qui peut ne pas être suffisant.

Comme le montre la FIGURE 5.1(b), SPEA II et NSGA III consomment un peu moins d'énergie par rapport à MOABC et MOPSO malgré la grande couverture dans tous les scénarios. SPEA II et NSGA III ont pu obtenir une bonne couverture (12 et 15 drones) sous un contrôle énergétique optimal, ce qui est une caractéristique importante pour le déploiement d'une couverture économe en énergie.

La FIGURE 5.1(c) montre le nombre de liaisons de communication entre les drones. Comme nous pouvons le voir sur la FIGURE 5.1(c), le nombre de liaisons de communication est directement proportionnel au nombre de drones. Dans le cas de 8 drones, tous les algorithmes fournissent un nombre comparable de liaisons de communication. En revanche, dans le cas de 12 et 15 drones, NSGA III est meilleur. Après NSGA III, SPEA II optimise mieux ce paramètre que les autres

algorithmes (NSGA II, MOABC et MOPSO) dont NSGA II et MOABC fournissent des valeurs comparables et qui sont meilleures par rapport à la valeur fournie par MOPSO.

La FIGURE 5.1(d) montre le chevauchement réduit dans le cas de 8, 12 et 15 drones. Il est clair qu'une augmentation du nombre de drones étend la zone de couverture et génère donc plus de chevauchements. Par rapport aux algorithmes MOABC, SPEA II, NSGA III et MOPSO, l'algorithme NSGA II a un paramètre de performance de chevauchement pire, mais il est meilleur dans d'autres paramètres de performance.

Comme le montre la FIGURE 5.1(a), SPEA II pourrait fournir une bonne couverture dans les deux scénarios (12 et 15 drones). Il n'a pas pu atteindre le seuil de couverture souhaité ( $> 90\%$ ) avec 8 drones, mais il a été en mesure de fournir le minimum de zones de chevauchement (FIGURE 5.1(d)).

En résumé, NSGA II, NSGA III et SPEA II obtiennent de meilleures performances globales et font efficacement des compromis entre ces objectifs.

## 5.5.2 Partie 2

Dans cette deuxième partie, l'évaluation et la comparaison des performances de différents algorithmes d'optimisation multi-objectif sont effectuées, en termes de deux propriétés qui sont généralement requises : la convergence et la diversité à l'aide de la métrique d'hypervolume et de la métrique de distance générationnelle inversée présentées précédemment.

Les FIGURES 5.2, 5.3, 5.4, 5.5, 5.6, 5.7 montrent les valeurs moyennes d'hypervolume et de IGD dans 10 exécutions avec 8, 12, 15 drones pour différentes itérations pour les cinq algorithmes.

Pour le cas de 8 drones (FIGURE 5.2), on constate que les valeurs IGD obtenues par NSGA II sont inférieures aux valeurs obtenues par les autres algorithmes et que les valeurs d'hypervolume (FIGURE 5.3) obtenues par le même algorithme sont plus grandes par rapport aux autres à

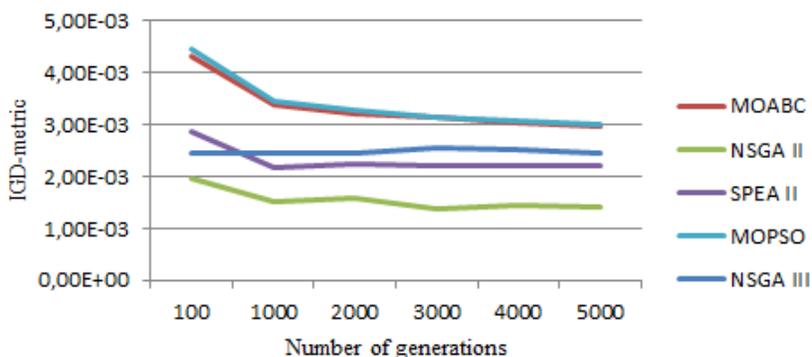


FIGURE 5.2 Evolution de la moyenne des valeurs IGD en fonction du nombre de générations pour 8 drones

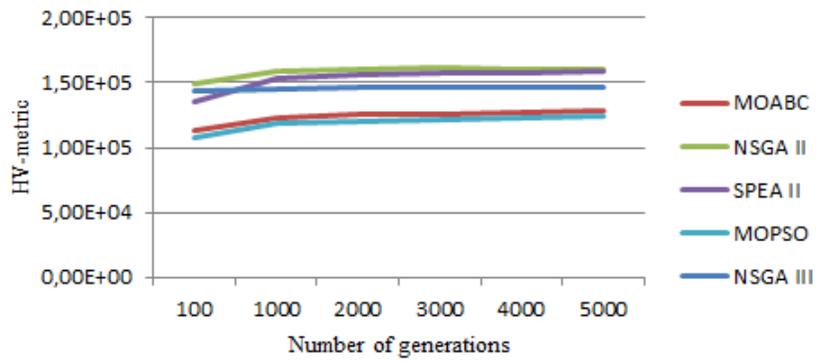


FIGURE 5.3 Evolution de la moyenne des valeurs HV en fonction du nombre de générations pour 8 drones

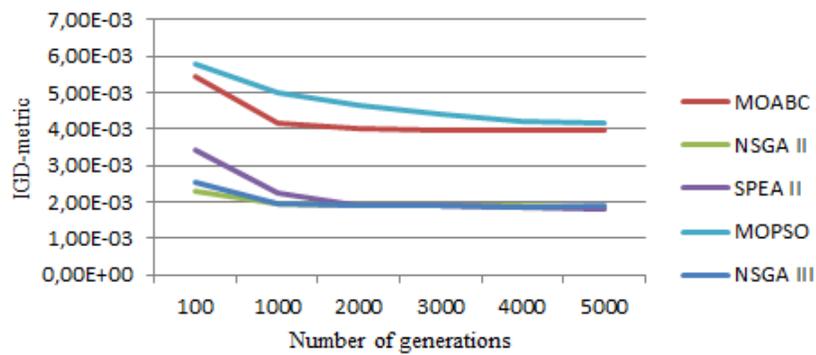


FIGURE 5.4 Evolution de la moyenne des valeurs IGD en fonction du nombre de générations pour 12 drones

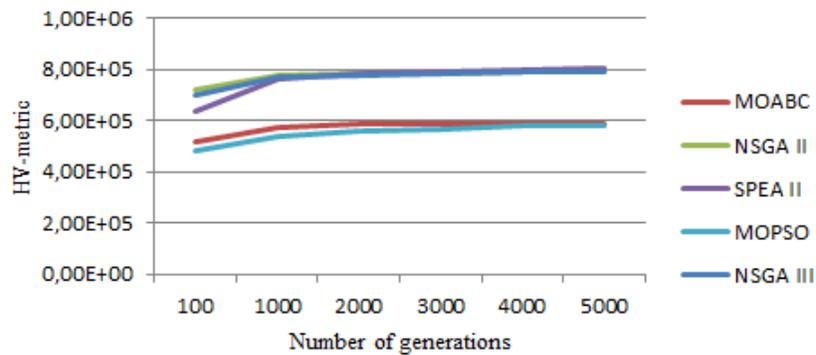
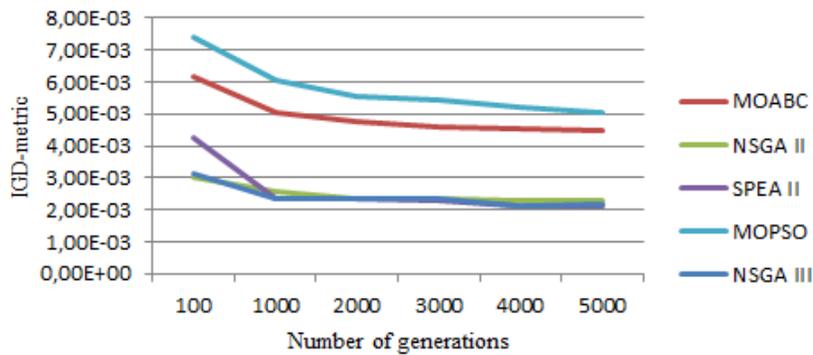
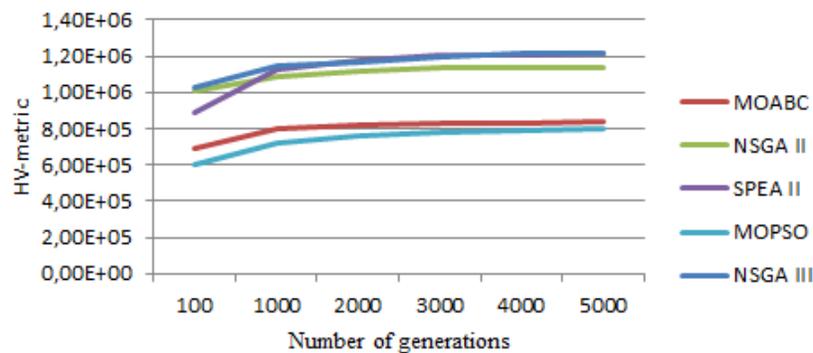


FIGURE 5.5 Evolution de la moyenne des valeurs HV en fonction du nombre de générations pour 12 drones



**FIGURE 5.6** Evolution de la moyenne des valeurs IGD en fonction du nombre de générations pour 15 drones



**FIGURE 5.7** Evolution de la moyenne des valeurs HV en fonction du nombre de générations pour 15 drones

travers les itérations. Cela signifie qu'il converge mieux que les autres. Pour les cas de 12 et 15 drones (FIGURES 5.4, 5.5, 5.6, 5.7), selon les valeurs IGD et de l'hypervolume, SPEA II surpasse légèrement NSGA III et NSGA II. La performance inférieure de NSGA II par rapport à SPEA II dans les deux derniers scénarios peut être due à la facilité de production de solutions dupliquées par NSGA II dans son processus itératif qui rend difficile la recherche de points isolés (et même la façon dont il réserve l'élitisme) et avec l'augmentation de la dimension du problème (12 et 15 drones), sa convergence est réduite.

Le tracé de coordonnées parallèles peut être utilisé pour visualiser les solutions afin de voir le compromis entre les différents objectifs, et même la distribution et la portée [94]. Les FIGURES 5.8, 5.9, 5.10, 5.11, 5.12 présentent un ensemble de solutions obtenues par les cinq algorithmes dans le deuxième scénario. Les valeurs des objectifs sont normalisées et présentées sur un tracé de chemin de valeurs. L'axe horizontal montre l'indice de l'objectif, où 1 présente la couverture, 2 la consommation d'énergie, 3 le nombre de liaisons de communication et 4 le chevauchement. L'axe vertical indique la valeur d'objectif correspondante.

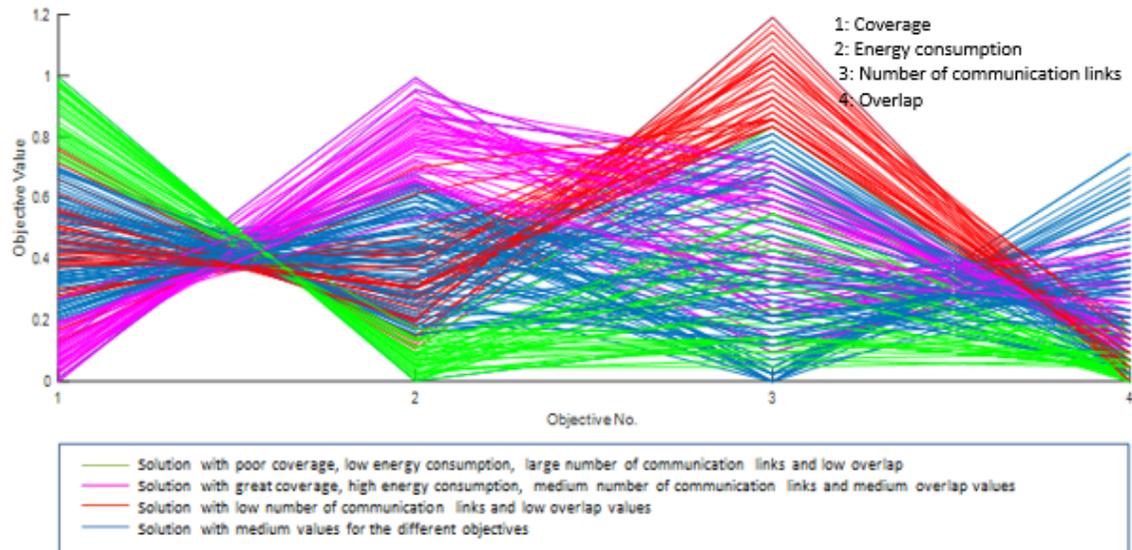


FIGURE 5.8 Coordonnées parallèles des solutions obtenues par l'algorithme NSGA II (IGD= $2.03e^{-3}$ )

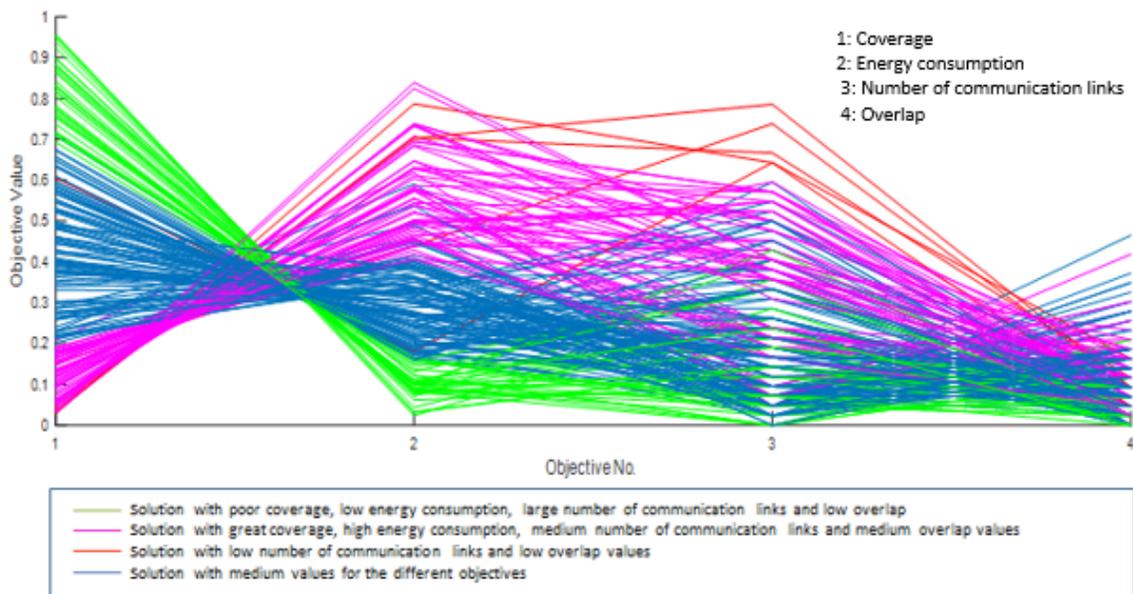


FIGURE 5.9 Coordonnées parallèles des solutions obtenues par l'algorithme SPEA II (IGD= $1.8e^{-3}$ )

Pour toutes les figures, les solutions obtenues par chaque algorithme sont regroupées en quatre couleurs. La couleur verte représente certaines solutions de l'ensemble de solutions qui ont une

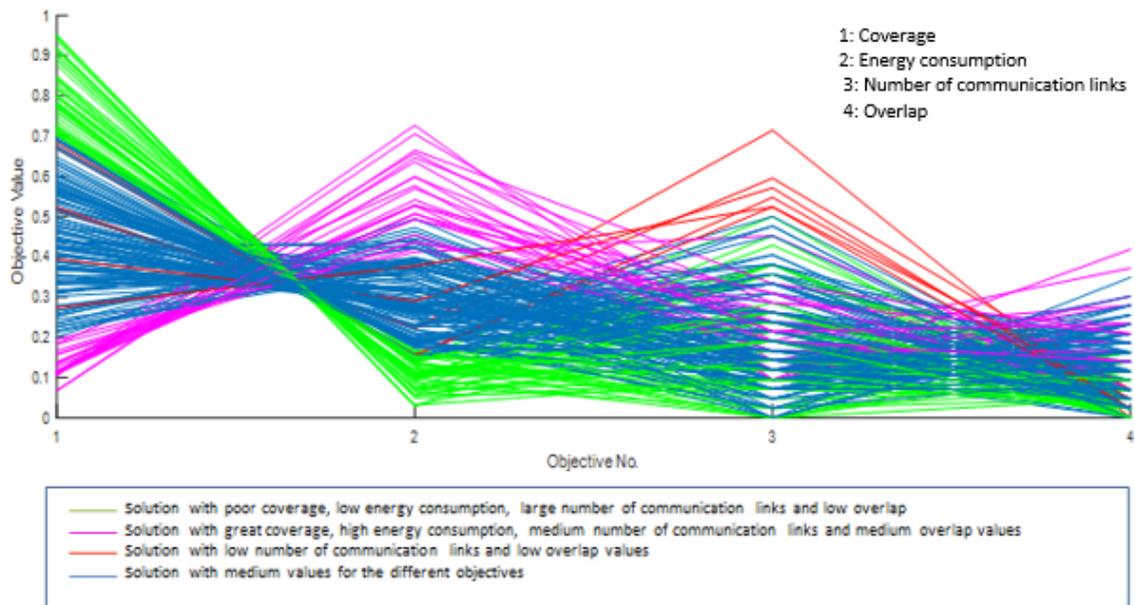


FIGURE 5.10 Coordonnées parallèles des solutions obtenues par l'algorithme NSGA III (IGD= $1.9e^{-3}$ )

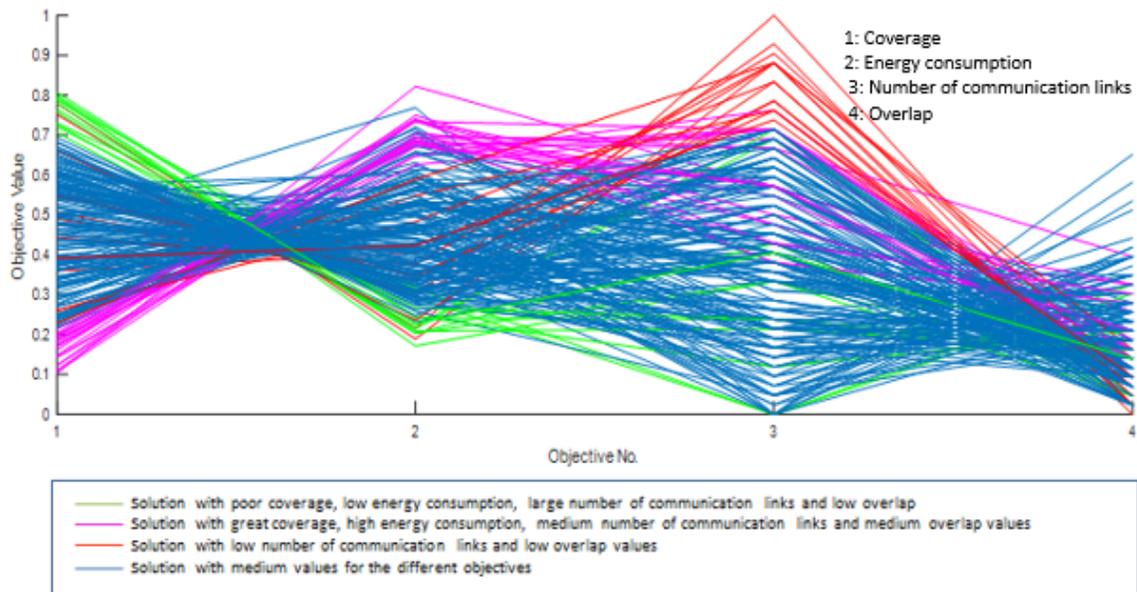


FIGURE 5.11 Coordonnées parallèles des solutions obtenues par l'algorithme MOPSO (IGD= $4.3e^{-3}$ )

faible consommation d'énergie avec une faible couverture et chevauchement, et un grand nombre de liaisons de communication. Pour les solutions qui sont en magenta, elles ont une grande couverture avec une consommation d'énergie élevée avec un nombre moyen de liaisons de communi-

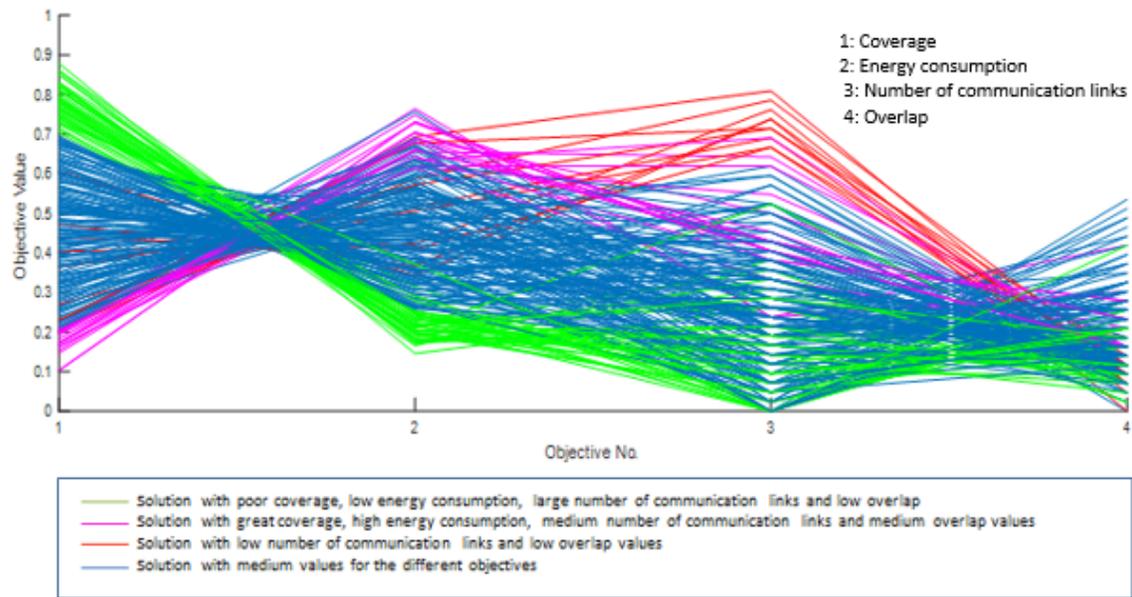


FIGURE 5.12 Coordonnées parallèles des solutions obtenues par l'algorithme MOABC (IGD=  $3.9e^{-3}$ )

cation et des valeurs moyennes de chevauchement. La couleur rouge représente les solutions avec un faible nombre de liaisons de communication et de faibles valeurs de chevauchement. La couleur bleue représente les solutions avec des valeurs moyennes pour les différents objectifs à l'exception du chevauchement. Ces couleurs sont utilisées pour mettre en évidence la relation entre les objectifs et pour aider le décideur à choisir une solution de déploiement en fonction de sa préférence. Par exemple, le choix d'une solution dans les solutions magenta permet à l'utilisateur d'avoir une grande couverture mais cette solution lui coûtera très cher en termes de consommation d'énergie.

De nombreuses lignes de croisement sont le signe d'une relation inverse ou d'un conflit d'objectifs. Par exemple, le conflit entre 1 et 2 est intense (plus la consommation d'énergie est élevée, plus le taux de couverture est élevé (et vice versa)). Le conflit entre 2 et 3 est légèrement intense. Le conflit entre 3 et 4 est plus intense (plus le nombre de liaisons de communication est faible, plus les valeurs de chevauchement sont faibles (et vice versa)) que le conflit entre 2 et 3. Comme nous pouvons le voir, le conflit entre 1 et 2 est plus intense que celui entre 2 et 3, 3 et 4.

À partir de ces figures, nous pouvons voir que l'ensemble des solutions de la FIGURE 5.8 qui représente NSGA II a plus de diversité et des valeurs d'objectifs plus grandes que les autres. Cependant, une grande diversité ne signifie pas toujours des valeurs d'hypervolume plus élevées [95]. Comme nous pouvons le voir, NSGA II a une faible convergence, avec son ensemble de solutions allant de 0 à 1,2 contrairement au front connu de Pareto allant de 0 à 1. SPEA II, MOABC, MOPSO et NSGA III sont tous dans la même plage et chacun peut converger différemment. Mais selon leurs valeurs IGD, SPEA II a une bonne convergence car il a la valeur la plus faible de la métrique de convergence. Les valeurs IGD indiquent que NSGA III est le deuxième en termes de convergence après SPEA II.

## 5.6 Conclusion

Dans ce chapitre, nous avons présenté les résultats de différentes expérimentations menées pour évaluer notre approche visant à obtenir un emplacement optimal des drones tout en satisfaisant certaines exigences mentionnées dans le chapitre précédent. Plus particulièrement, nous avons commencé par la description des scénarios. Ensuite, les différents paramètres de simulation sont présentés. De plus, les résultats de simulation ont été divisés en deux parties. La partie 1 permet de voir les différents compromis générés par les différents algorithmes pour trois scénarios différents. La partie 2 présente la comparaison entre les différents algorithmes en termes de la convergence et de la diversité.

L'étude comparative qui a été réalisée entre les cinq algorithmes méta-heuristiques (SPEA II, NSGA II, MOABC, MOPSO, NSGA III) a démontré que SPEA II, NSGA III et NSGA II permettent un meilleur compromis entre les objectifs d'optimisation.

Grâce à cet examen et aux résultats obtenus, les conclusions suivantes peuvent être tirées :

- Comme le montrent les résultats, dans le premier scénario, NSGA II a obtenu une meilleure convergence et diversité par rapport aux autres algorithmes. En revanche, dans les deux derniers, il trouve une difficulté à maintenir une bonne convergence par rapport à SPEA II et NSGA III qui la maintiennent avec une bonne diversité, et cela s'explique par son mécanisme d'élitisme et de recherche de solutions. A l'aide d'un ensemble de points de référence bien répartis, NSGA III a réussi à trouver un ensemble de solutions bien convergées et bien diversifiées. De même pour SPEA II, son mécanisme d'attribution de fitness et ses techniques de troncature d'archives et de sélection basée sur la densité lui ont permis d'avoir des meilleures valeurs de métriques de performance, en particulier dans les derniers scénarios. De plus, MOABC et MOPSO produisent des solutions moins diverses et qui convergent moins par rapport aux autres algorithmes, ce que signifie que ces derniers (NSGA II, NSGA III et SPEA II) exploitent mieux l'espace de recherche.
- La plupart des solutions de déploiement méta-heuristiques existantes souffrent du problème d'être sensible à l'initialisation des paramètres. Ces paramètres nécessitent un réglage fin avant que l'algorithme ne fonctionne de manière plus efficace.
- Bien qu'il y ait eu de nombreuses méta-heuristiques, aucune d'elles n'est parfaite et la sélection parmi elles dépend des exigences d'une situation de conception particulière.

# Conclusion générale

## Conclusion

Dans cette thèse, nous avons présenté et discuté les derniers travaux de recherche et les différentes technologies utilisées dans le domaine de l'agriculture intelligente pour relever les défis majeurs auxquels le secteur agricole est confronté, notamment la nécessité d'augmentation de la productivité pour satisfaire les exigences de la population mondiale croissante sous la pression des changements climatiques, la pénurie d'eau, etc.

Nous nous sommes intéressés également aux techniques d'exploration de données utilisées pour extraire des informations utiles à la gestion des exploitations agricoles. Ainsi, nous avons revu ces techniques pour concevoir et réaliser des solutions efficaces et pratiques aux problèmes agricoles. Cette contribution a mis en évidence la particularité des données agricoles, ce qui rend l'exploration de données très importante pour l'analyse des données agricoles. Elle démarque également les techniques utilisées pour découvrir et comprendre les relations entre les différents facteurs et leur influence sur le bien-être des plantes et leur rendement. Selon nos études, les travaux de recherche les plus récents ont commencé à se tourner surtout vers les méthodes hybrides pour tirer parti des différentes données.

Les travaux mentionnés dans cette thèse s'intéressent également à l'application des méthodes d'optimisation dans le cas d'un déploiement d'un nombre de drones et à l'application de méta-heuristiques d'optimisation pour résoudre le problème de positionnement et de déploiement de drones. En effet, les performances d'un réseau de drones sont fortement influencées par le processus de déploiement des nœuds. La question du déploiement et du positionnement des nœuds dans un réseau de drones est une stratégie qui est utilisée pour définir la topologie du réseau, le nombre et la position des nœuds. La qualité de la surveillance, la connectivité et la consommation d'énergie sont également directement affectées par la topologie du réseau. Le problème du placement optimal des nœuds est prouvé NP-difficile pour la plupart des formulations de déploiement.

L'utilisation des méta-heuristiques est justifiée par la nécessité de satisfaire différents objectifs (souvent contradictoires) comme l'optimisation de la consommation d'énergie, la maximisation de la couverture et la connectivité réseau et la minimisation du nombre de drones déployés. Etant donné l'espace de recherche énorme et la nécessité d'avoir une solution optimale mais dans un délai raisonnable, le véritable test des approches proposées est souvent impossible, nous avons recours à des simulations pour prouver l'efficacité de ces méthodes.

La performance de l'approche proposée est exposée dans le dernier chapitre en termes d'amélioration de quatre paramètres essentiels pour un meilleur déploiement tout en changeant le nombre de drones. A partir de nos résultats expérimentaux nous avons constaté qu'aucune méta-heuristique n'est meilleure qu'une autre et cela est fortement lié à plusieurs paramètres.

Les activités de déploiement peuvent être regroupées en deux phases principales. Une phase de pré-déploiement et de déploiement qui concerne le placement initial des nœuds. Une phase de post-déploiement qui est nécessaire si la topologie du réseau a évolué en raison d'un déplacement de nœuds ou d'un changement des conditions externes. Dans cette thèse nous nous sommes focalisés sur la première phase tandis que la seconde ouvre de nouveaux champs de recherche montrant des problèmes qui restent non résolus et qui pourront être abordés dans nos futurs travaux qui peuvent être poursuivis à court et à moyen termes.

## Perspectives

### A court terme :

Dans une recherche future, il serait intéressant d'améliorer notre approche en considérant :

- Un modèle qui prendra en compte les spécificités de l'application et des conditions plus particulières en milieu agricole (conditions météorologiques, évitant de couvrir des zones en dehors des limites de plantation).
- Un modèle de consommation d'énergie plus réaliste qui inclut aussi l'énergie de communication et de traitement de données, et d'autres paramètres importants pour un meilleur déploiement tels que la durée de vie des drones et l'évitement d'obstacles, ce qui améliorera également l'approche proposée.

### A Moyen terme :

- Dans une stratégie de déploiement statique, la décision sur les emplacements des drones est prise au moment de la configuration du réseau et ne tient pas compte des changements dynamiques dans sa phase d'exploitation. En effet, plusieurs situations peuvent se produire : l'usage d'une application peut varier dans le temps et les ressources réseau disponibles peuvent changer. Par conséquent, le redéploiement dynamique des drones pendant que le réseau est opérationnel est nécessaire pour faire face aux variations dynamiques des ressources du réseau et de l'environnement. Nous envisageons de concevoir un algorithme d'auto-déploiement efficace qui prend en compte les modifications dynamiques au cours de la mission pour répondre à la question : où déplacer et comment déplacer efficacement les drones mobiles afin que le déploiement final réponde aux objectifs de conception du réseau (couverture, connectivité, tolérance aux pannes, durée de vie, charge, etc.).
- A travers cette thèse, nous avons constaté que la performance des méta-heuristiques proposées dépend de l'ajustement de leurs paramètres qui sont difficilement réglables/contrôlables

et coûteux en temps d'exécution. Nous avons constaté aussi que le choix de la meilleure méta-heuristique pour le problème de déploiement n'est pas trivial et il est implicitement fait par l'expert du domaine qui utilise son expérience pour apprécier l'efficacité des différentes solutions. Par conséquent, nous envisageons de concevoir une hyper-heuristique, qui devrait produire des solutions de qualité acceptable, basées sur un ensemble de méta-heuristiques connues. Une hyper-heuristique est considérée comme une approche de haut niveau qui, compte tenu d'une instance de problème particulière et d'un certain nombre de méta-heuristiques de bas niveau, peut sélectionner et appliquer une bonne méta-heuristique ou une combinaison de méta-heuristiques appropriées à chaque point de décision. Elle permet aussi l'auto-réglage/auto-ajustement des valeurs des paramètres, selon l'exigence, au fur et à mesure du processus d'optimisation.

- Nous envisageons de développer un système d'aide à la décision qui assiste les agriculteurs dans l'exécution de leurs activités agricoles, leur évaluation et leur optimisation, ainsi que dans la planification stratégique, tactique et opérationnelle des opérations agricoles en se basant sur de nouvelles techniques de Data Mining en exploitant à la fois les données spatiales et temporelles et prendre des décisions sur la base d'une combinaison de données collectées par des drones (images, paramètres environnementaux, etc.).

# Publications

Ce travail de thèse a donné lieu à des publications dans des revues et conférences internationales :

## Articles dans une revue acceptée

1. H. Ait Issad, R. Aoudjit, J. J. Rodrigues. A comprehensive review of Data Mining techniques in smart agriculture. *Engineering in Agriculture, Environment and Food*, 12(4) :511-525, 2019. <https://doi.org/10.1016/j.eaef.2019.11.003>.
2. H. Ait Issad, R. Aoudjit, M. Belkadi, J. J. Rodrigues. Many-objective Optimization based Optimal drone deployment for agricultural zone. *International Journal of Communication Networks and Distributed Systems*, 26(1) :76-98, 2021.

## Conférences internationales

1. H. Ait Issad, R. Aoudjit, M. Belkadi, M. Lalam, M. Daoui. Deep Learning-based Crops and weeds classification in Smart Agriculture. In *International Conference on Environment and Life Science - (ICENLISC-19)*, pages 45-52, Dec 2019.

# Bibliographie

- [1] M. Abbasgholipour, M. Omid, A. Keyhani, and S. S. Mohtasebi. Color image segmentation with genetic algorithm in a raisin sorting system based on machine vision in variable conditions. *Expert Systems with Applications*, 38(4) :3671–3678, 2011.
- [2] F. Abbasi, A. Mesbahi, and J. M. Velni. A new voronoi-based blanket coverage control method for moving sensor networks. *IEEE Transactions on Control Systems Technology*, 27(1) :409–417, 2017.
- [3] M. Abdelhafidh, M. Fourati, L. C. Fourati, and A. Chouaya. Wireless sensor network monitoring system : architecture, applications and future directions. *International Journal of Communication Networks and Distributed Systems*, 23(4) :413–451, 2019.
- [4] S. Abdollahzadeh and N. J. Navimipour. Deployment strategies in the wireless sensor network : A comprehensive review. *Computer Communications*, 91 :1–16, 2016.
- [5] F. Aftab, A. Khan, and Z. Zhang. Bio-inspired clustering scheme for internet of drones application in industrial wireless sensor network. *International Journal of Distributed Sensor Networks*, 15(11), 2019.
- [6] C. C. Aggarwal. *Data Classification : Algorithms and Applications*. Chapman and Hall/CRC, first edition, 2014.
- [7] R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghavan. *Automatic subspace clustering of high dimensional data for data mining applications*. ACM, 1998.
- [8] R. Agrawal, R. Srikant, et al. Fast algorithms for mining association rules. In *20th International Conference very large data bases, VLDB*, pages 487–499, 1994.
- [9] W. AIOTI. Smart farming and food safety internet of things applications–challenges for large scale implementations. *Brussels : Alliance for the Internet of Things Innovation*, 2015.
- [10] H. Ait Issad, R. Aoudjit, M. Belkadi, M. Lalam, and M. Daoui. Deep learning-based crops and weeds classification in smart agriculture. In *International Conference on Environment and Life Science - (ICENLISC-19)*, pages 45–52, Dec 2019.
- [11] H. Ait Issad, R. Aoudjit, M. Belkadi, and J. J. Rodrigues. Many-objective optimization based optimal drone deployment for agricultural zone. *International Journal of Communication Networks and Distributed Systems*, 26(1) :76–98, 2021.
- [12] H. Ait Issad, R. Aoudjit, and J. J. Rodrigues. A comprehensive review of data mining techniques in smart agriculture. *Engineering in Agriculture, Environment and Food*, 12(4) :511 – 525, 2019.
- [13] J. N. Al-Karaki and A. Gawanmeh. The optimal deployment, coverage, and connectivity problems in wireless sensor networks : revisited. *IEEE Access*, 5 :18051–18065, 2017.

- [14] M. I. Alipio, A. E. M. D. Cruz, J. D. A. Doria, and R. M. S. Fruto. A smart hydroponics farming system using exact inference in bayesian network. In *2017 IEEE 6th Global Conference on Consumer Electronics (GCCE)*, pages 1–5. IEEE, 2017.
- [15] B. H. Y. Alsalam, K. Morton, D. Campbell, and F. Gonzalez. Autonomous uav with vision based on-board decision making for remote sensing and precision agriculture. In *Aerospace Conference*, pages 1–12. IEEE, 2017.
- [16] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu. 3-d placement of an unmanned aerial vehicle base station (uav-bs) for energy-efficient maximal coverage. *IEEE Wireless Communications Letters*, 6(4) :434–437, 2017.
- [17] M. Alzenad, A. El-Keyi, and H. Yanikomeroglu. 3d placement of an unmanned aerial vehicle base station for maximum coverage of users with different qos requirements. *IEEE Wireless Communications Letters*, 7(1) :38–41, 2017.
- [18] J. Amara, B. Bouaziz, A. Algergawy, et al. A deep learning-based approach for banana leaf diseases classification. In *BTW (Workshops)*, pages 79–88, 2017.
- [19] M. Amin, M. Aman Ullah, and A. Akbar. Time series modeling for forecasting wheat production of pakistan. *Journal of Animal and Plant Sciences*, 24 :1444–1451, 10 2014.
- [20] M. Ammad-Udin, A. Mansour, D. Le Jeune, E. H. M. Aggoune, and M. Ayaz. Uav routing protocol for crop health management. In *24th European Signal Processing Conference (EUSIPCO)*, pages 1818–1822. IEEE, 2016.
- [21] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander. Optics : ordering points to identify the clustering structure. In *Proc. ACM Sigmod record*, pages 49–60. ACM, 1999.
- [22] S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agricultural Engineering International : CIGR Journal*, 15(1) :211–217, 2013.
- [23] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.-H. M. Aggoune. Internet-of-things (iot)-based smart agriculture : Toward making the fields talk. *IEEE Access*, 7 :129551–129583, 2019.
- [24] M. D. Bah, E. Dericquebourg, A. Hafiane, and R. Canals. Deep learning based classification system for identifying weeds using high-resolution uav imagery. In *2018 Conference on Science and Information*, pages 176–187. Springer, 2018.
- [25] C. Bahr, D. Forristal, S. Fountas, E. Gil, G. Grenier, R. Hoerfarer, A. Jonsson, A. Jung, C. Kempenaar, K. Lokhorst, et al. Eip-agri focus group : Precision farming. Technical report, European Commission, 2015.
- [26] C. Balaceanu, I. Marcu, and G. Suci. Telemetry system for smart agriculture. In *International Conference on Business Information Systems*, pages 573–584. Springer, 2019.
- [27] N. Balakrishnan and G. Muthukumarasamy. Crop production-ensemble machine learning model for prediction. *International Journal of Computer Science and Software Engineering*, 5(7) :148, 2016.
- [28] D. Baquero, J. Molina, R. Gil, C. Bojacá, H. Franco, and F. Gómez. An image retrieval system for tomato disease assessment. In *2014 XIX Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, pages 1–5. IEEE, 2014.

- [29] A. Barbedo and J. Garcia. Digital image processing techniques for detecting, quantifying and classifying plant diseases. *SpringerPlus*, 2(1) :660, Dec 2013.
- [30] S. Bechikh, R. Datta, and A. Gupta. *Recent advances in evolutionary multi-objective optimization*, volume 20. Springer, 2016.
- [31] M. Bhangе and H. Hingoliwala. Smart farming : Pomegranate disease detection using image processing. *Procedia Computer Science*, 58 :280–288, 2015.
- [32] L. V. Bijuraj. Clustering and its applications. In *2013 National Conference on New Horizons in IT-NCNHIT*, page 169, 2013.
- [33] M. Blagojević, M. Blagojević, and V. Ličina. Web-based intelligent system for predicting apricot yields using artificial neural networks. *Scientia Horticulturae*, 213 :125–131, 2016.
- [34] S. Bonadies and S. A. Gadsden. An overview of autonomous crop row navigation strategies for unmanned ground vehicles. *Engineering in Agriculture, Environment and Food*, 12(1) :24–31, 2019.
- [35] S. Bonadies, A. Lefcourt, and S. A. Gadsden. A survey of unmanned ground vehicles with applications to agricultural and environmental sensing. In *Autonomous air and ground sensing systems for agricultural optimization and phenotyping*, volume 9866. International Society for Optics and Photonics, 2016.
- [36] P. Boniecki, K. Koszela, H. Piekarska-Boniecka, J. Weres, M. Zaborowicz, S. Kujawa, A. Majewski, and B. Raba. Neural identification of selected apple pests. *Computers and Electronics in Agriculture*, 110 :9–16, 2015.
- [37] G. Bontempi. Machine learning strategies for time series prediction. *European Business Intelligence Summer School, Hammamet, Lecture*, pages 1–128, 2013.
- [38] P. Bosilj, T. Duckett, and G. Cielniak. Connected attribute morphology for unified vegetation segmentation and classification in precision agriculture. *Computers in Industry*, 98 :226–240, 2018.
- [39] M. Brahimі, K. Boukhalfa, and A. Moussaoui. Deep learning for tomato diseases : classification and symptoms visualization. *Applied Artificial Intelligence*, 31(4) :299–315, 2017.
- [40] L. Breiman, J. Friedman, R. A. Olshen, and C. J. Stone. *Classification and regression trees*. CRC, 1984.
- [41] S. Brin, R. Motwani, J. D. Ullman, and S. Tsur. Dynamic itemset counting and implication rules for market basket data. *Acm Sigmod Record*, 26(2) :255–264, 1997.
- [42] Y. Cai, K. Guan, D. Lobell, A. B. Potgieter, S. Wang, J. Peng, T. Xu, S. Asseng, Y. Zhang, L. You, et al. Integrating satellite and climate data to predict wheat yield in australia using machine learning approaches. *Agricultural and Forest Meteorology*, 274 :144–159, 2019.
- [43] C. Caillouet and T. Razafindralambo. Efficient deployment of connected unmanned aerial vehicles for optimal target coverage. In *Global Information Infrastructure and Networking Symposium (GIIS)*, pages 1–8. IEEE, 2017.
- [44] B. Cao, J. Zhao, Z. Lv, X. Liu, X. Kang, and S. Yang. Deployment optimization for 3d industrial wireless sensor networks based on particle swarm optimizers with distributed parallelism. *Journal of Network and Computer Applications*, 103 :225–238, 2018.

- [45] M. Cardei, J. Wu, M. Lu, and M. O. Pervaiz. Maximum network lifetime in wireless sensor networks with adjustable sensing ranges. In *WiMob'2005, IEEE International Conference on Wireless And Mobile Computing, Networking And Communications*, volume 3, pages 438–445. IEEE, 2005.
- [46] S. Chand and M. Wagner. Evolutionary many-objective optimization : A quick-start guide. *Surveys in Operations Research and Management Science*, 20(2) :35–42, 2015.
- [47] A. Chlingaryan, S. Sukkarieh, and B. Whelan. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture : A review. *Computers and Electronics in Agriculture*, 151 :61–69, 2018.
- [48] S. Chouikhi. *Tolérance aux pannes dans un réseau de capteurs sans fil multi-canal*. PhD thesis, Paris Est, 2016.
- [49] A. Chowdhury and S. A. Raut. A survey study on internet of things resource management. *Journal of Network and Computer Applications*, 120 :42–60, 2018.
- [50] D. Cillis, B. Maestrini, A. Pezzuolo, F. Marinello, and L. Sartori. Modeling soil organic carbon and carbon dioxide emissions in different tillage systems supported by precision agriculture technologies under current climatic conditions. *Soil and Tillage Research*, 183 :51–59, 2018.
- [51] C. A. C. Coello, G. B. Lamont, D. A. Van Veldhuizen, et al. *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer, 2007.
- [52] C. C. Coello and M. S. Lechuga. Mopso : a proposal for multiple objective particle swarm optimization. In *Proceedings of the 2002 Congress on Evolutionary Computation*, volume 2, pages 1051–1056. IEEE, 2002.
- [53] F. G. Costa, J. Ueyama, T. Braun, G. Pessin, F. S. Osório, and P. A. Vargas. The use of unmanned aerial vehicles and wireless sensor network in agricultural applications. In *International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 5045–5048. IEEE, 2012.
- [54] J. M. Da Silva, C. V. Damásio, A. M. Sousa, L. Bugalho, L. Pessanha, and P. Quaresma. Agriculture pest and disease risk maps considering msg satellite data and land surface temperature. *International Journal of Applied Earth Observation and Geoinformation*, 38 :40–50, 2015.
- [55] S. D'Arpa, E. Barca, and V. F. Uricchio. Il data mining nello sviluppo di modelli previsionali per la gestione integrata dell'agricoltura sostenibile. *Journal of Agrometeorology*, 2(2011), 2011.
- [56] I. Das and J. E. Dennis. Normal-boundary intersection : A new method for generating the pareto surface in nonlinear multicriteria optimization problems. *SIAM journal on optimization*, 8(3) :631–657, 1998.
- [57] J. Das, G. Cross, C. Qu, A. Makineni, P. Tokekar, Y. Mulgaonkar, and V. Kumar. Devices, systems, and methods for automated monitoring enabling precision agriculture. In *2015 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 462–469. IEEE, 2015.
- [58] M. Díaz, C. Martín, and B. Rubio. State-of-the-art, challenges, and open issues in the integration of internet of things and cloud computing. *Journal of Network and Computer Applications*, 67 :99–117, 2016.

- [59] F. De Rango, N. Palmieri, A. F. Santamaria, and G. Potrino. A simulator for uavs management in agriculture domain. In *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, pages 1–8. IEEE, 2017.
- [60] K. Deb and H. Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i : Solving problems with box constraints. *IEEE Trans. Evolutionary Computation*, 18(4) :577–601, 2014.
- [61] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm : Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2) :182–197, 2002.
- [62] I. Dhau, E. Adam, K. K. Ayisi, and O. Mutanga. Detection and mapping of maize streak virus using rapideye satellite imagery. *Geocarto International*, 34(8) :856–866, 2019.
- [63] B. H. Dhivya, R. Manjula, S. Siva Bharathi, and R. Madhumathi. A survey on crop yield prediction based on agricultural data. *International Journal of Innovative Research in Science, Engineering and Technology*, 6 :4177–4183, March 2017.
- [64] C. Dong, J. Xie, H. Dai, Q. Wu, Z. Qin, and Z. Feng. Optimal deployment density for maximum coverage of drone small cells. *China Communications*, 15(5) :25–40, 2018.
- [65] K.-L. Du and M. Swamy. *Search and Optimization by Metaheuristics, Techniques and Algorithms Inspired by Nature*. Springer International Publishing, Switzerland, 2016.
- [66] B. Duan, S. Fang, R. Zhu, X. Wu, S. Wang, Y. Gong, and Y. Peng. Remote estimation of rice yield with unmanned aerial vehicle (uav) data and spectral mixture analysis. *Frontiers in plant science*, 10, 2019.
- [67] S. R. Dubey and A. S. Jalal. Detection and classification of apple fruit diseases using complete local binary patterns. In *2012 Third International Conference on Computer and Communication Technology*, pages 346–351, Nov 2012.
- [68] V. Duggal, M. Sukhwani, K. Bipin, G. S. Reddy, and K. M. Krishna. Plantation monitoring and yield estimation using autonomous quadcopter for precision agriculture. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5121–5127. IEEE, 2016.
- [69] H. Durmuş, E. O. Güneş, and M. Kırıcı. Disease detection on the leaves of the tomato plants by using deep learning. In *6th International Conference on Agro-Geoinformatics*, pages 1–5. IEEE, 2017.
- [70] J. Dyson, A. Mancini, E. Frontoni, and P. Zingaretti. Deep learning for soil and crop segmentation from remotely sensed data. *Remote Sensing*, 11(16) :1859, 2019.
- [71] Y. El Khamlichi, A. Tahiri, A. Abtoy, I. Medina-Bulo, and F. Palomo-Lozano. A hybrid algorithm for optimal wireless sensor network deployment with the minimum number of sensor nodes. *Algorithms*, 10(3) :80, 2017.
- [72] N. A. El-Sherbeny. Vehicle routing with time windows : An overview of exact, heuristic and metaheuristic methods. *Journal of King Saud University - Science*, 22(3) :123 – 131, 2010.
- [73] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. Kdd*, pages 226–231, 1996.
- [74] FAO. The state of food and agriculture. <http://www.fao.org/3/w1358e/w1358e.pdf>, 1996.

- [75] M. S. Farooq, S. Riaz, A. Abid, K. Abid, and M. A. Naeem. A survey on the role of iot in agriculture for the implementation of smart farming. *IEEE Access*, 7 :156237–156271, 2019.
- [76] M. Farsi, M. A. Elhosseini, M. Badawy, H. A. Ali, and H. Z. Eldin. Deployment techniques in wireless sensor networks, coverage and connectivity : a survey. *IEEE Access*, 7 :28940–28954, 2019.
- [77] K. P. Ferentinos. Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145 :311–318, 2018.
- [78] H. Finch, A. Samuel, and G. Lane. 10 - precision farming. In H. Finch, A. Samuel, and G. Lane, editors, *Lockhart Wiseman’s Crop Husbandry Including Grassland (Ninth Edition)*, pages 235 – 244. Woodhead Publishing, ninth edition, 2014.
- [79] D. H. Fisher. Knowledge acquisition via incremental conceptual clustering. *Machine learning*, 2(2) :139–172, 1987.
- [80] N. Gandhi and L. J. Armstrong. A review of the application of data mining techniques for decision making in agriculture. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pages 1–6, Dec 2016.
- [81] M. Geetha. Implementation of association rule mining for different soil types in agriculture. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(4) :520–522, 2015.
- [82] J. H. Gennari, P. Langley, and D. Fisher. Models of incremental concept formation. *Artificial intelligence*, 40(1) :11–61, 1989.
- [83] J. George and R. M. Sharma. Relay node placement in wireless sensor networks using modified genetic algorithm. In *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, pages 551–556. IEEE, 2016.
- [84] A. Goldstein, L. Fink, A. Meitin, S. Bohadana, O. Lutenberg, and G. Ravid. Applying machine learning on sensor data for irrigation recommendations : revealing the agronomists’ tacit knowledge. *Precision Agriculture*, 19(3) :421–444, 2018.
- [85] N. Gondchawar and R. Kawitkar. Iot based smart agriculture. *International Journal of advanced research in Computer and Communication Engineering*, 5(6) :838–842, 2016.
- [86] C. Goumopoulos, B. O’Flynn, and A. Kameas. Automated zone-specific irrigation with wireless sensor/actuator network and adaptable decision support. *Computers and electronics in agriculture*, 105 :20–33, 2014.
- [87] D. Grigg. Green revolution. In N. J. Smelser and P. B. Baltes, editors, *International Encyclopedia of the Social Behavioral Sciences*, pages 6389 – 6393. Pergamon, Oxford, 2001.
- [88] S. K. Gupta, P. Kuila, and P. K. Jana. Genetic algorithm for k-connected relay node placement in wireless sensor networks. In *Proceedings of the second international conference on computer and communication technologies*, pages 721–729. Springer, 2016.
- [89] J. Han, J. Pei, Y. Yin, and R. Mao. Mining frequent patterns without candidate generation : A frequent-pattern tree approach. *Data mining and knowledge discovery*, 8(1) :53–87, 2004.
- [90] J.-K. Hao, P. Galinier, and M. Habib. Métaheuristiques pour l’optimisation combinatoire et l’affectation sous contraintes. *Revue d’intelligence artificielle*, 13(2) :283–324, 1999.

- [91] J. A. Hartigan and M. A. Wong. Algorithm as 136 : A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1) :100–108, 1979.
- [92] A. Hinneburg, D. A. Keim, et al. An efficient approach to clustering in large multimedia databases with noise. In *Proc. KDD*, pages 58–65, 1998.
- [93] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. Squeezenet : Alexnet-level accuracy with 50x fewer parameters and < 0.5 mb model size. *arXiv preprint arXiv :1602.07360*, 1(10), 2016.
- [94] A. Ibrahim, S. Rahnamayan, M. V. Martin, and K. Deb. 3d-radvis : Visualization of pareto front in many-objective optimization. In *Congress on Evolutionary Computation (CEC)*, pages 736–745. IEEE, 2016.
- [95] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima. Performance comparison of nsga-ii and nsga-iii on various many-objective test problems. In *IEEE Congress on Evolutionary Computation (CEC)*, pages 3045–3052. IEEE, 2016.
- [96] S. Jamali and M. Hatami. Coverage aware scheduling in wireless sensor networks : An optimal placement approach. *Wireless Personal Communications*, 85(3) :1689–1699, 2015.
- [97] H. Jawad, R. Nordin, S. Gharghan, A. Jawad, and M. Ismail. Energy-efficient wireless sensor networks for precision agriculture : A review. *Sensors*, 17(8) :1781, 2017.
- [98] M. Jhuria, A. Kumar, and R. Borse. Image processing for smart farming : Detection of disease and fruit grading. In *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*, pages 521–526, Dec 2013.
- [99] Z. Jian and Z. Wei. Support vector machine for recognition of cucumber leaf diseases. In *2010 2nd International Conference on Advanced Computer Control*, pages 264–266, March 2010.
- [100] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang. Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access*, 7 :59069–59080, 2019.
- [101] S. Jiang, Y.-S. Ong, J. Zhang, and L. Feng. Consistencies and contradictions of performance metrics in multiobjective optimization. *IEEE Trans. Cybernetics*, 44(12) :2391–2404, 2014.
- [102] T. Kalaivani, A. Allirani, and P. Priya. A survey on zigbee based wireless sensor networks in agriculture. In *3rd International Conference on Trendz in Information Sciences & Computing (TISC2011)*, pages 85–89. IEEE, 2011.
- [103] E. Kalantari, M. Z. Shakir, H. Yanikomeroğlu, and A. Yongacoglu. Backhaul-aware robust 3d drone placement in 5g+ wireless networks. In *IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 109–114. IEEE, 2017.
- [104] E. Kalantari, H. Yanikomeroğlu, and A. Yongacoglu. On the number and 3d placement of drone base stations in wireless cellular networks. In *84th Vehicular Technology Conference (VTC-Fall)*, pages 1–6. IEEE, 2016.
- [105] A. Kamilaris and F. X. Prenafeta-Boldu. Deep learning in agriculture : A survey. *Computers and Electronics in Agriculture*, 147 :70–90, 2018.
- [106] M. Kantardzic. *Data Mining : Concepts, Models, Methods, and Algorithms*. Wiley-IEEE Press, second edition, 2011.

- [107] A. K. Kar, S. K. Patel, and R. Yadav. A comparative study & performance evaluation of different clustering techniques in data mining. In *ACEIT Conference Proceeding*, pages 139–142, 2016.
- [108] D. Karaboga and B. Basturk. A powerful and efficient algorithm for numerical function optimization : artificial bee colony (abc) algorithm. *Journal of global optimization*, 39(3) :459–471, 2007.
- [109] S. Karthik, A. Paul, and N. Karthikeyan. *Deep Learning Innovations and Their Convergence With Big Data (Advances in Data Mining and Database Management)*. IGI Global, first edition, 2017.
- [110] G. Karypis, E.-H. S. Han, and V. Kumar. Chameleon : Hierarchical clustering using dynamic modeling. *Computer*, (8) :68–75, 1999.
- [111] P. Katsigiannis, L. Misopolinos, V. Liakopoulos, T. K. Alexandridis, and G. Zalidis. An autonomous multi-sensor uav system for reduced-input precision agriculture applications. In *2016 24th Mediterranean Conference on Control and Automation (MED)*, pages 60–64. IEEE, 2016.
- [112] R. Kaur and S. Singla. Classification of plant leaf diseases using gradient and texture feature. In *2016 International Conference on Advances in Information Communication Technology & Computing*, page 96. ACM, 2016.
- [113] Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi. Basic study of automated diagnosis of viral plant diseases using convolutional neural networks. In *2015 International Symposium on Visual Computing*, pages 638–645. Springer, 2015.
- [114] I. Khoufi, P. Minet, A. Laouiti, and S. Mahfoudh. Survey of deployment algorithms in wireless sensor networks : coverage and connectivity issues and challenges. *Int. J. Autonomous and Adaptive Communications Systems*, 2017.
- [115] Y.-h. Kim, C.-M. Kim, D.-S. Yang, Y.-j. Oh, and Y.-H. Han. Regular sensor deployment patterns for p-coverage and q-connectivity in wireless sensor networks. In *The International Conference on Information Network 2012*, pages 290–295. IEEE, 2012.
- [116] Y. H. Kim, S. J. Yoo, Y. H. Gu, J. H. Lim, D. Han, and S. W. Baik. Crop pests prediction method using regression and machine learning technology : Survey. *IERI Procedia*, 6 :52–56, 2014.
- [117] A. King et al. The future of agriculture. *Nature*, 544(7651) :S21–S23, 2017.
- [118] A. Kishor, P. K. Singh, and J. Prakash. Nsabc : Non-dominated sorting based multi-objective artificial bee colony algorithm and its application in data clustering. *Neurocomputing*, 216 :514–533, 2016.
- [119] G. Kokkonis, S. Kontogiannis, and D. Tomtsis. Fitra : A neuro-fuzzy computational algorithm approach based on an embedded water planting system. In *2nd International Conference on Internet of Things, Data and Cloud Computing*, pages 1–8, 2017.
- [120] K. R. Krishna. *Push button agriculture : Robotics, drones, satellite-guided soil and crop management*. CRC Press, 2017.
- [121] K. R. Krishna. *Agricultural drones : a peaceful pursuit*. Apple Academic Press, 2018.

- [122] A. Krizhevsky. One weird trick for parallelizing convolutional neural networks. *arXiv preprint arXiv :1404.5997*, 2014.
- [123] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [124] S. Kumar, S. Mishra, P. Khanna, et al. Precision sugarcane monitoring using svm classifier. *Procedia Computer Science*, 122 :881–887, 2017.
- [125] R. Lal and B. A. Stewart. *Soil-specific farming : precision agriculture*. CRC Press, 2015.
- [126] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4) :541–551, 1989.
- [127] J. Y. Lee. *Multi-objective optimisation using the Bees Algorithm*. Cardiff University, 2010.
- [128] P. J. R. Leonard Kaufman. *Finding groups in data : an introduction to cluster analysis (Ninth Edition)*. Wiley-Interscience, 1990.
- [129] M. Li, S. Yang, and X. Liu. Diversity comparison of pareto front approximations in many-objective optimization. *IEEE Transactions on Cybernetics*, 44(12) :2568–2584, 2014.
- [130] R.-a. Li, X. Sha, and K. Lin. Smart greenhouse : A real-time mobile intelligent monitoring system based on wsn. In *2014 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1152–1156. IEEE, 2014.
- [131] X. Li, L. Ci, M. Yang, C. Tian, and X. Li. Deploying three-dimensional mobile sensor networks based on virtual forces algorithm. In *China Conference on Wireless Sensor Networks*, pages 204–216. Springer, 2012.
- [132] Y. Li, M. Qian, P. Liu, Q. Cai, X. Li, J. Guo, H. Yan, F. Yu, K. Yuan, J. Yu, et al. The recognition of rice images by uav based on capsule network. *Cluster Computing*, pages 1–10, 2018.
- [133] Z. Li and L. Lei. Sensor node deployment in wireless sensor networks based on improved particle swarm optimization. In *2009 International Conference on Applied Superconductivity and Electromagnetic Devices*, pages 215–217. IEEE, 2009.
- [134] W.-H. Liao, Y. Kao, and R.-T. Wu. Ant colony optimization based sensor deployment protocol for wireless sensor networks. *Expert Systems with Applications*, 38(6) :6599–6605, 2011.
- [135] B. R. Limited. Towards smart farming : Agriculture embracing the iot vision. Technical report, Reasearch report, UK, London, USA, Boston, 2015.
- [136] M. Liu, A. K. Skidmore, T. Wang, X. Liu, L. Wu, and L. Tian. An approach for heavy metal pollution detected from spatio-temporal stability of stress in rice using satellite images. *International Journal of Applied Earth Observation and Geoinformation*, 80 :230–239, 2019.
- [137] Y. Liu, L. Suo, D. Sun, and A. Wang. A virtual square grid-based coverage algorithm of redundant node for wireless sensor network. *Journal of network and computer applications*, 36(2) :811–817, 2013.

- [138] P. Lottes, M. Hoeflerlin, S. Sander, M. Müter, P. Schulze, and L. C. Stachniss. An effective classification system for separating sugar beets and weeds for precision farming applications. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5157–5163. IEEE, 2016.
- [139] H. Lu, X. Fu, C. Liu, L.-g. Li, Y.-x. He, and N.-w. Li. Cultivated land information extraction in uav imagery based on deep convolutional neural network and transfer learning. *Journal of Mountain Science*, 14(4) :731–741, 2017.
- [140] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang. Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, 267 :378–384, 2017.
- [141] M. Maksimović and V. Milošević. Evaluating the optimal sensor placement for smoke detection. *Yugoslav Journal of Operations Research*, 26(1) :33–50, 2016.
- [142] M. Mancuso and F. Bustaffa. A wireless sensors network for monitoring environmental variables in a tomato greenhouse. In *2006 IEEE International Workshop on Factory Communication Systems*, pages 107–110, 2006.
- [143] O. Matei, T. Rusu, A. Petrovan, and G. Mihaș. A data mining system for real time soil moisture prediction. *Procedia Engineering*, 181 :837–844, 2017.
- [144] I. Maza and A. Ollero. Multiple uav cooperative searching operation using polygon area decomposition and efficient coverage algorithms. In *Distributed Autonomous Robotic Systems 6*, pages 221–230. Springer, 2007.
- [145] T. McKinnon. Agricultural drones : What farmers need to know. *Boulder, Colorado*, 2016.
- [146] A. Meola. Internet of things, big data and smart farming are the future of agriculture. *business insider*, Dec, 2016.
- [147] B. Milovic and V. Radojevic. Application of data mining in agriculture. *Bulgarian Journal of Agricultural Science*, 21(1) :26–34, 2015.
- [148] S. Mishra, D. Mishra, and G. H. Santra. Applications of machine learning techniques in agricultural crop production : a review paper. *Indian Journal of Science and Technology*, 9(38) :1–14, 2016.
- [149] U. R. Mogili and B. Deepak. Review on application of drone systems in precision agriculture. *Procedia Computer Science*, 133 :502–509, 2018.
- [150] O. Moh’d Alia and A. Al-Ajourri. Maximizing wireless sensor network coverage with minimum cost using harmony search algorithm. *IEEE Sensors Journal*, 17(3) :882–896, 2016.
- [151] D. Moshou, C. Bravo, R. Oberti, J. West, H. Ramon, S. Vougioukas, and D. Bochtis. Intelligent multi-sensor system for the detection and treatment of fungal diseases in arable crops. *Biosystems Engineering*, 108(4) :311–321, 2011.
- [152] K. Mougou, S. Mahfoudh, P. Minet, and A. Laouiti. Redeployment of randomly deployed wireless mobile sensor nodes. In *2012 IEEE Vehicular Technology Conference (VTC Fall)*, pages 1–5. IEEE, 2012.
- [153] M. Murad, K. M. Yahya, and G. M. Hassan. Web based poultry farm monitoring system using wireless sensor network. In *7th International Conference on Frontiers of Information Technology*, page 7. ACM, 2009.

- [154] A. M. Mustapha Reda Senouci. *Deploying Wireless Sensor Networks : Theory and Practice*. ISTE Press - Elsevier, first edition, 2016.
- [155] K. Nagar. Data mining clustering methods : A review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 5 :575–579, April 2015.
- [156] M. Nagaraju and P. Chawla. Iot implementation and management for smart farming. *International Journal of Innovative Technology and Exploring Engineering*, 8, 2019.
- [157] N. Nasri, S. Mnasri, and T. Val. 3d node deployment strategies prediction in wireless sensors network. *International Journal of Electronics*, pages 1–31, 2019.
- [158] J. Natividade, J. Prado, and L. Marques. Low-cost multi-spectral vegetation classification using an unmanned aerial vehicle. In *International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 336–342. IEEE, 2017.
- [159] D. Ngom. *Optimisation de la durée de vie dans les réseaux de capteurs sans fil sous contraintes de couverture et de connectivité réseau*. PhD thesis, Mulhouse, 2016.
- [160] A. Niimi, M. Wada, K. Ito, M. Toda, K. Hatanaka, and O. Konishi. Broiler-house environment monitoring system using sensor network and mail delivery system. *Artificial Life and Robotics*, 13(1) :264–270, 2008.
- [161] L. R. Oded Maimon. *Data Mining and Knowledge Discovery Handbook*. Texts and Monographs in Physics. Springer US, second edition, 2010.
- [162] C. Ozturk, D. Karaboga, and B. Gorkemli. Probabilistic dynamic deployment of wireless sensor networks by artificial bee colony algorithm. *sensors*, 11(6) :6056–6065, 2011.
- [163] C. Öztürk, D. Karaboğa, and B. GÖRKEMLİ. Artificial bee colony algorithm for dynamic deployment of wireless sensor networks. *Turkish Journal of Electrical Engineering & Computer Sciences*, 20(2) :255–262, 2012.
- [164] P. Padalalu, S. Mahajan, K. Dabir, S. Mitkar, and D. Javale. Smart water dripping system for agriculture/farming. In *2017 2nd International Conference for Convergence in Technology (I2CT)*, pages 659–662, April 2017.
- [165] P. B. Padol and A. A. Yadav. Svm classifier based grape leaf disease detection. In *2016 Conference on Advances in Signal Processing (CASP)*, pages 175–179. IEEE, 2016.
- [166] J. Pang, Z.-y. Bai, J.-c. Lai, and S.-k. Li. Automatic segmentation of crop leaf spot disease images by integrating local threshold and seeded region growing. In *2011 international conference on image analysis and signal processing*, pages 590–594. IEEE, 2011.
- [167] S. P. Patil and R. S Zambre. Classification of cotton leaf spot disease using support vector machine. *International Journal of Engineering Research and Applications*, 4 :92–97, May 2014.
- [168] S. S. Patil and S. A. Thorat. Early detection of grapes diseases using machine learning and iot. In *2016 Second International Conference on Cognitive Computing and Information Processing (CCIP)*, pages 1–5, Aug 2016.
- [169] P. Pawar, V. Turkar, and P. Patil. Cucumber disease detection using artificial neural network. In *2016 International Conference on Inventive Computation Technologies (ICICT)*, pages 1–5. IEEE, 2016.

- [170] S. M. Pedersen, K. M. Lind, et al. *Precision Agriculture : Technology and Economic Perspectives*. Springer, 2017.
- [171] R. G. Perea, E. C. Poyato, P. Montesinos, and J. R. Díaz. Prediction of irrigation event occurrence at farm level using optimal decision trees. *Computers and Electronics in Agriculture*, 157 :173–180, 2019.
- [172] H. X. Pham, H. M. La, D. Feil-Seifer, and L. Van Nguyen. Cooperative and distributed reinforcement learning of drones for field coverage. *arXiv preprint arXiv :1803.07250*, 2018.
- [173] X. Pham and M. Stack. How data analytics is transforming agriculture. *Business horizons*, 61(1) :125–133, 2018.
- [174] R. Polikar, C. Zhang, and Y. Ma. *Ensemble Machine Learning : Methods and Applications*. Springer-Verlag New York, first edition, 2012.
- [175] I. Pölönen, H. Saari, J. Kaivosoja, E. Honkavaara, and L. Pesonen. Hyperspectral imaging based biomass and nitrogen content estimations from light-weight uav. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XV*, volume 8887. International Society for Optics and Photonics, 2013.
- [176] D. Popescu, F. Stoican, G. Stamatescu, O. Chenaru, and L. Ichim. A survey of collaborative uav–wsn systems for efficient monitoring. *Sensors*, 19(21) :4690, 2019.
- [177] H. P. Prajapati, B. Yatharth, and M. Singh. Imaging of crop canopies for the remote diagnosis and quantification of field responses. In *International Conference on Inventive Systems and Control (ICISC)*, pages 1–5. IEEE, 2017.
- [178] R. Priyadarshi, B. Gupta, and A. Anurag. Deployment techniques in wireless sensor networks : a survey, classification, challenges, and future research issues. *The Journal of Supercomputing*, pages 1–41, 2020.
- [179] L. D. P. Pugliese, F. Guerriero, D. Zorbas, and T. Razafindralambo. Modelling the mobile target covering problem using flying drones. *Optimization Letters*, 10(5) :1021–1052, 2016.
- [180] A. Pulver and R. Wei. Optimizing the spatial location of medical drones. *Applied Geography*, 90 :9–16, 2018.
- [181] V. Puri, A. Nayyar, and L. Raja. Agriculture drones : A modern breakthrough in precision agriculture. *Journal of Statistics and Management Systems*, 20(4) :507–518, 2017.
- [182] J. R. Quinlan. Induction of decision trees. *Machine learning*, 1(1) :81–106, 1986.
- [183] J. R. Quinlan. *C4. 5 : programs for machine learning*. Elsevier, 2014.
- [184] S. Rajeswari, K. Suthendran, and K. Rajakumar. A smart agricultural model by integrating iot, mobile and cloud-based big data analytics. In *2017 International Conference on Intelligent Computing and Control (I2C2)*, pages 1–5. IEEE, 2017.
- [185] A. Rani, A. Chaudhary, N. Sinha, M. Mohanty, and R. Chaudhary. Drone : the green technology for future agriculture. *Harit Dhara*, 2(1) :3–6, 2019.
- [186] G. Ravichandran and R. Koteeshwari. Agricultural crop predictor and advisor using ann for smartphones. In *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*, pages 1–6. IEEE, 2016.
- [187] A.-u. Rehman. *Smart Agriculture : An Approach towards Better Agriculture Management*. OMICS Group eBook, 2015.

- [188] T. U. Rehman, M. S. Mahmud, Y. K. Chang, J. Jin, and J. Shin. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Computers and electronics in agriculture*, 156 :585–605, 2019.
- [189] D. Reina, T. Camp, A. Munjal, and S. Toral. Evolutionary deployment and local search-based movements of 0th, responders in disaster scenarios. *Future Generation Computer Systems*, 88 :61–78, 2018.
- [190] P. Revathi and M. Hemalatha. Classification of cotton leaf spot diseases using image processing edge detection techniques. In *2012 International Conference on Emerging Trends in Science, Engineering and Technology (INCOSSET)*, pages 169–173, Dec 2012.
- [191] M. Rout and R. Roy. Dynamic deployment of randomly deployed mobile sensor nodes in the presence of obstacles. *Ad Hoc Networks*, 46 :12–22, 2016.
- [192] L. Ruan, J. Wang, J. Chen, Y. Xu, Y. Yang, H. Jiang, Y. Zhang, and Y. Xu. Energy-efficient multi-uav coverage deployment in uav networks : A game-theoretic framework. *China Communications*, 15(10) :194–209, 2018.
- [193] R. Rupnik, M. Kukar, P. Vračar, D. Košir, D. Pevec, and Z. Bosnić. Agrodss : A decision support system for agriculture and farming. *Computers and Electronics in Agriculture*, 2018.
- [194] G. Ruß and A. Brenning. Data mining in precision agriculture : management of spatial information. In *2010 International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 350–359. Springer, 2010.
- [195] T. Sakamoto, A. A. Gitelson, A. L. Nguy-Robertson, T. J. Arkebauer, B. D. Wardlow, A. E. Suyker, S. B. Verma, and M. Shibayama. An alternative method using digital cameras for continuous monitoring of crop status. *Agricultural and Forest Meteorology*, 154 :113–126, 2012.
- [196] B. J. Samajpati and S. D. Degadwala. Hybrid approach for apple fruit diseases detection and classification using random forest classifier. In *2016 International Conference on Communication and Signal Processing (ICCSP)*,, pages 1015–1019. IEEE, 2016.
- [197] S. S. Sannakki, V. S. Rajpurohit, V. B. Nargund, and P. Kulkarni. Diagnosis and classification of grape leaf diseases using neural networks. In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, pages 1–5, July 2013.
- [198] A. V. Savkin and H. Huang. Asymptotically optimal deployment of drones for surveillance and monitoring. *Sensors*, 19(9) :2068, 2019.
- [199] N. Schor, A. Bechar, T. Ignat, A. Dombrovsky, Y. Elad, and S. Berman. Robotic disease detection in greenhouses : combined detection of powdery mildew and tomato spotted wilt virus. *IEEE Robotics and Automation Letters*, 1(1) :354–360, 2016.
- [200] V. Sellam and E. Poovammal. Prediction of crop yield using regression analysis. *Indian Journal of Science and Technology*, 9(38), pages 1–5, 2016.
- [201] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat : Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv :1312.6229*, 2013.

- [202] V. Sharma, K. Srinivasan, H.-C. Chao, K.-L. Hua, and W.-H. Cheng. Intelligent deployment of uavs in 5g heterogeneous communication environment for improved coverage. *Journal of Network and Computer Applications*, 85 :94–105, 2017.
- [203] Shruti and D. R. Patil. An approach for agricultural field monitoring and control using iot. *International Research Journal of Engineering and Technology (IRJET)*, 3 :86–89, 2016.
- [204] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv :1409.1556*, 2014.
- [205] V. Singh, A. Misra, et al. Detection of unhealthy region of plant leaves using image processing and genetic algorithm. In *2015 International Conference on Advances in Computer Engineering and Applications (ICACEA)*, pages 1028–1032. IEEE, 2015.
- [206] N. Srinivas and K. Deb. Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary computation*, 2(3) :221–248, 1994.
- [207] I. Strumberger, N. Bacanin, S. Tomic, M. Beko, and M. Tuba. Static drone placement by elephant herding optimization algorithm. In *25th Telecommunication Forum (TELFOR)*, pages 1–4. IEEE, 2017.
- [208] H. Sundmaeker, C. Verdouw, S. Wolfert, and L. Pérez Freire. Internet of food and farm 2020. *Digitising the Industry-Internet of Things connecting physical, digital and virtual worlds*, 2, 2016.
- [209] K. Suresh and S. Krishna Priya. A study on pre-harvest forecast of sugarcane yield using climatic variables. *Stat. Appl. 7&8 (1&2)(New Series)*, pages 1–8, 2009.
- [210] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *2015 IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.
- [211] H. Tan, Y. Wang, X. Hao, Q.-S. Hua, and F. C. Lau. Arbitrary obstacles constrained full coverage in wireless sensor networks. In *International Conference on Wireless Algorithms, Systems, and Applications*, pages 1–10. Springer, 2010.
- [212] R. Tang, X. Qian, and X. Yu. On virtual-force algorithms for coverage-optimal node deployment in mobile sensor networks via the two-dimensional yukawa crystal. *International Journal of Distributed Sensor Networks*, 15(9), 2019.
- [213] F. Tian, B. Wu, H. Zeng, X. Zhang, and J. Xu. Efficient identification of corn cultivation area with multitemporal synthetic aperture radar and optical images in the google earth engine cloud platform. *Remote Sensing*, 11(6) :629, 2019.
- [214] N. C. Tri, T. Van Hoai, H. N. Duong, N. T. Trong, V. Van Vinh, and V. Snasel. A novel framework based on deep learning and unmanned aerial vehicles to assess the quality of rice fields. In *International Conference on Advances in Information and Communication Technology*, pages 84–93. Springer, 2016.
- [215] A. Tripathy, J. Adinarayana, S. Merchant, U. Desai, S. Ninomiya, M. Hirafuji, and T. Kiura. Data mining and wireless sensor network for groundnut pest/disease precision protection. In *2013 National Conference on Parallel Computing Technologies (PARCOMPTECH)*, pages 1–8. IEEE, 2013.

- [216] A. Tripathy, J. Adinarayana, D. Sudharsan, S. Merchant, U. Desai, K. Vijayalakshmi, D. R. Reddy, G. Sreenivas, S. Ninomiya, M. Hirafuji, et al. Data mining and wireless sensor network for agriculture pest/disease predictions. In *2011 World Congress on Information and Communication Technologies (WICT)*, pages 1229–1234. IEEE, 2011.
- [217] A. Tripathy, J. Adinarayana, K. Vijayalakshmi, S. Merchant, U. Desai, S. Ninomiya, M. Hirafuji, and T. Kiura. Knowledge discovery and leaf spot dynamics of groundnut crop through wireless sensor network and data mining techniques. *Computers and electronics in agriculture*, 107 :104–114, 2014.
- [218] D. C. Tsouros, S. Bibi, and P. G. Sarigiannidis. A review on uav-based applications for precision agriculture. *Information*, 10(11) :349, 2019.
- [219] E. Tuba, I. Tuba, D. Dolicanin-Djekic, A. Alihodzic, and M. Tuba. Efficient drone placement for wireless sensor networks coverage by bare bones fireworks algorithm. In *6th International Symposium on Digital Forensic and Security (ISDFS)*, pages 1–5. IEEE, 2018.
- [220] T. Ullmert and N. Fröhlich. A heuristic approach for the drone placement problem. In *Multikriterielle Optimierung und Entscheidungsunterstützung*, pages 45–59. Springer, 2019.
- [221] USGCRP. Impacts, risks, and adaptation in the united states : Fourth national climate assessment, volume ii. *US Global Change Research Program*, 2018.
- [222] D. Van Le, H. Oh, and S. Yoon. Virfid : A virtual force (vf)-based interest-driven moving phenomenon monitoring scheme using multiple mobile sensor nodes. *Ad Hoc Networks*, 27 :112–132, 2015.
- [223] J. L. Van Zanden. The first green revolution : the growth of production and productivity in european agriculture, 1870-1914. *Economic History Review*, pages 215–239, 1991.
- [224] S. A. M. Varman, A. R. Baskaran, S. Aravindh, and E. Prabhu. Deep learning and iot for smart agriculture using wsn. In *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pages 1–6. IEEE, 2017.
- [225] S. Veenadhari, B. Mishra, and C. Singh. Soybean productivity modelling using decision tree algorithms. *International Journal of Computer Applications*, 27(7) :11–15, 2011.
- [226] S. Veenadhari, B. Misra, and C. Singh. Machine learning approach for forecasting crop yield based on climatic parameters. In *2014 International Conference on Computer Communication and Informatics*, pages 1–5, Jan 2014.
- [227] P. Vinciya and A. Valarmathi. Agriculture analysis for next generation high tech farming in data mining. *International Journal of Advanced Research in Computer Science and Software Engineering (ijarcsse)*, 6(5) :481–488, 2016.
- [228] I. Wahab, O. Hall, and M. Jirström. Remote sensing of yields : Application of uav imagery-derived ndvi for estimating maize vigor and yields in complex farming systems in sub-saharan africa. *Drones*, 2(3) :28, 2018.
- [229] B. Wang. Coverage problems in sensor networks : A survey. *ACM Computing Surveys (CSUR)*, 43(4) :1–53, 2011.
- [230] G. Wang, G. Cao, and T. F. La Porta. Movement-assisted sensor deployment. *IEEE Transactions on Mobile Computing*, 5(6) :640–652, 2006.

- [231] H. Wang and Z. Ma. Prediction of wheat stripe rust based on support vector machine. In *7th International Conference on Natural Computation*, pages 378–382, July 2011.
- [232] S. Wang, G. Azzari, and D. B. Lobell. Crop type mapping without field-level labels : Random forest transfer and unsupervised clustering techniques. *Remote sensing of environment*, 222 :303–317, 2019.
- [233] W. Wang, J. Yang, R. Muntz, et al. Sting : A statistical information grid approach to spatial data mining. In *Proc. VLDB*, pages 186–195, 1997.
- [234] P. P. Warne and S. Ganorkar. Detection of diseases on cotton leaves using k-mean clustering method. *International Research Journal of Engineering and Technology (IRJET)*, 2 :04, 2015.
- [235] B. Whelan and J. Taylor. *Precision agriculture for grain production systems*. CSIRO Publishing, 2013.
- [236] K. Xiao, D. Xiao, and X. Luo. Smart water-saving irrigation system in precision agriculture based on wireless sensor network. *Transactions of the Chinese Society of Agricultural Engineering*, 26(11) :170–175, 2010.
- [237] L. Xiao and L. Guo. The realization of precision agriculture monitoring system based on wireless sensor network. In *2010 International Conference On Computer and Communication Technologies in Agriculture Engineering (CCTAE)*, pages 89–92. IEEE, 2010.
- [238] T. Xie, Z. Huang, Z. Chi, and T. Zhu. Minimizing amortized cost of the on-demand irrigation system in smart farms. In *3rd International Workshop on Cyber-Physical Systems for Smart Water Networks*, pages 43–46, 2017.
- [239] I.-T. R. Y.2238. Overview of smart farming based on networks. <http://handle.itu.int/11.1002/1000/12520>, 2015.
- [240] S. Yahata, T. Onishi, K. Yamaguchi, S. Ozawa, J. Kitazono, T. Ohkawa, T. Yoshida, N. Murakami, and H. Tsuji. A hybrid machine learning approach to automatic plant phenotyping for smart agriculture. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 1787–1793. IEEE, 2017.
- [241] N. Yang, D. Liu, Q. Feng, Q. Xiong, L. Zhang, T. Ren, Y. Zhao, D. Zhu, and J. Huang. Large-scale crop mapping based on machine learning and parallel computation with grids. *Remote Sensing*, 11(12) :1500, 2019.
- [242] E. Yanmaz, S. Yahyanejad, B. Rinner, H. Hellwagner, and C. Bettstetter. Drone networks : Communications, coordination, and sensing. *Ad Hoc Networks*, 68 :1–15, 2018.
- [243] Q. Yao, Z. Guan, Y. Zhou, J. Tang, Y. Hu, and B. Yang. Application of support vector machine for detecting rice diseases using shape and color texture features. In *2009 International Conference on Engineering Computation*, pages 79–83, May 2009.
- [244] J. Yu, S. M. Sharpe, A. W. Schumann, and N. S. Boyd. Deep learning for image-based weed detection in turfgrass. *European Journal of Agronomy*, 104 :78–84, 2019.
- [245] X. Yu, W. Huang, J. Lan, and X. Qian. A novel virtual force approach for node deployment in wireless sensor network. In *2012 IEEE 8th International Conference on Distributed Computing in Sensor Systems*, pages 359–363. IEEE, 2012.

- 
- [246] X. Yu, N. Liu, W. Huang, X. Qian, and T. Zhang. A node deployment algorithm based on van der waals force in wireless sensor networks. *International Journal of Distributed Sensor Networks*, 9(10) :505710, 2013.
- [247] X. Yu, P. Wu, W. Han, and Z. Zhang. Overview of wireless underground sensor networks for agriculture. *African journal of biotechnology*, 11(17) :3942–3948, 2012.
- [248] X. Yu, J. Zhang, J. Fan, and T. Zhang. A faster convergence artificial bee colony algorithm in sensor deployment for wireless sensor networks. *International Journal of Distributed Sensor Networks*, 9(10) :497264, 2013.
- [249] M. J. Zaki. Scalable algorithms for association mining. *IEEE transactions on knowledge and data engineering*, 12(3) :372–390, 2000.
- [250] Y. Zeng and R. Zhang. Energy-efficient uav communication with trajectory optimization. *IEEE Transactions on Wireless Communications*, 16(6) :3747–3760, 2017.
- [251] M. Zeni, E. Ondula, R. Mbitiru, A. Nyambura, L. Samuel, K. Fleming, and K. Weldemariam. Low-power low-cost wireless sensors for real-time plant stress detection. In *2015 Annual Symposium on Computing for Development*, pages 51–59. ACM, 2015.
- [252] C. Zhang and S. Zhang. *Association Rule Mining : Models and Algorithms*. Springer-Verlag Berlin Heidelberg, first edition, 2002.
- [253] J. Zhang, L. Feng, and F. Yao. To obtain the planting area of corn crop based on modis satellite data. In *2012 2nd International Conference on Remote Sensing, Environment and Transportation Engineering*, pages 1–4. IEEE, 2012.
- [254] M. Zhang, H. Lin, G. Wang, H. Sun, and J. Fu. Mapping paddy rice using a convolutional neural network (cnn) with landsat 8 datasets in the dongting lake area, china. *Remote Sensing*, 10 :1840, 2018.
- [255] T. Zhang, R. Ramakrishnan, and M. Livny. Birch : an efficient data clustering method for very large databases. In *Proc. ACM Sigmod Record*, pages 103–114. ACM, 1996.
- [256] E. Zitzler, M. Laumanns, and L. Thiele. Spea2 : Improving the strength pareto evolutionary algorithm. *TIK-report*, 103 :1–21, 2001.
- [257] D. Zorbas, L. D. P. Pugliese, T. Razafindralambo, and F. Guerriero. Optimal drone placement and cost-efficient target coverage. *Journal of Network and Computer Applications*, 75 :16–31, 2016.
- [258] D. Zorbas, T. Razafindralambo, F. Guerriero, et al. Energy efficient mobile target tracking using flying drones. *Procedia Computer Science*, 19 :80–87, 2013.
- [259] W. Zou, Y. Zhu, H. Chen, and B. Zhang. Solving multiobjective optimization problems using artificial bee colony algorithm. *Discrete dynamics in nature and society*, 2011, 2011.