

الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي و البحث العلمي
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
جامعة عمّار تليجي بالأغواط
UNIVERSITE AMAR TELIDJI LAGHOUAT



كلية العلوم
DÉPARTEMENT D'INFORMATIQUE

MÉMOIRE DE MASTER

Domaine : Mathématique et Informatique

Filière : Informatique

Option : Réseaux , Systèmes d'information et de décision

Présenté par : KOUIDRI Soundous et ZEBEIR Amina

Thème

**Gestion Intelligente du Trafic de Données Agricoles par
UAV via l'Apprentissage par Renforcement**

Mme. SAHLAOUI Atika	MC(B)	(Université de Laghouat)	President
M. BEN ARFA Abdelmajid	MC(B)	(Université de Laghouat)	Examineur
M. BOUZID Tarek	MC(B)	(Université de Laghouat)	Examineur
M. AMEUR Mohamed El Amine	MC(B)	(Université de Laghouat)	Encadreur

Année Universitaire 2024/2025

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dédicaces

Je dédie ce travail :

À mon père qui a toujours été là pour moi, qui m'a donné un bel exemple de travail acharné et de persévérance. J'espère qu'il le trouvera.

À ma mère : En toi je vois la mère parfaite, toujours prête se sacrifiant pour le bonheur de ses filles, merci pour toute la patience et le soutien.

À ma chère sœur Hiba, pour ton affection, ta compréhension et ton inquiétude.

À mon mari et soutien qui a toujours été à mes côtés.

À mon ami Soundous et compagnon dans ce travail.

À mon encadreur, Mr. Ameer Mohamed El Amine , pour sa patience et sa tolérance, et pour ses précieux conseils.

À tous ceux qui m'ont aidé à terminer ce travail, et à tous ceux qui J'aime et ceux qui m'aiment.

.....

Amina

Dédicaces

Je dédie ce modeste travail :

...

À l'âme de mon cher père qui a toujours été à mes côtés... qui m'a tracé le bon chemin...

là où je suis aujourd'hui.

Que Dieu lui fasse miséricorde et le reçoive dans son vaste paradis ...

Et ma chère mère, qui m'a aidé psychologiquement, merci à vous

A mes frères, ma sœur, ma famille.

A tous mes chers(es) amis(es), particulièrement les plus proches Soulaf et Fatima.

Et pour mon encadreur, Mr. Ameer Mohamed El Amine, je le remercie infiniment pour son soutien et ses conseils les plus précieux...

...

Soundous

Remerciement

Tout d'abord, je remercie mon Dieu, le tout puissant de nous avoir donné la santé et la volonté d'entamer et de terminer ce mémoire. Tout d'abord, ce travail ne serait pas aussi riche et n'aurait pas pu avoir le jour sans l'aide et l'encadrement de Mr Ameer Mohamed El Amine, on le remercie pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur et sa disponibilité durant notre préparation de ce mémoire. Nos remerciements s'adressent également à tous nos professeurs pour leurs générosités et la grande patience dont ils ont su faire preuve malgré leurs charges académiques et professionnelles. Tous les membres de nos Familles et nos ami(e)s.

Résumé

Ce mémoire traite de l'utilisation des techniques d'apprentissage par renforcement (Reinforcement Learning) pour la collecte efficace de données dans un environnement agricole à l'aide de drones (UAVs). L'étude s'intéresse particulièrement à l'optimisation de deux aspects critiques : la consommation énergétique des UAVs et la minimisation du délai de collecte, tout en assurant une couverture maximale des capteurs IoT déployés dans les champs. Le document présente une simulation complète mettant en œuvre différents algorithmes tels que PPO (Proximal Policy Optimization) et DDPG (Deep Deterministic Policy Gradient). Ces méthodes permettent aux UAVs d'apprendre des politiques optimales pour visiter les capteurs, en fonction de leur état énergétique, de leur localisation et de la priorité des données. Les résultats montrent que l'utilisation de l'apprentissage par renforcement améliore significativement l'efficacité énergétique et réduit les délais de traitement, comparé aux approches classiques comme les stratégies aléatoires ou gloutonnes.

Mots clés : Apprentissage par renforcement (RL), UAV, IoT, Agriculture intelligente, Optimisation énergétique, offloading.

Abstract

This thesis focuses on the use of Reinforcement Learning (RL) techniques for efficient data collection in an agricultural environment using Unmanned Aerial Vehicles (UAVs). The study addresses two key challenges : optimizing energy consumption of UAVs and minimizing data collection latency , while ensuring maximum coverage of IoT sensors deployed in fields. A full simulation is presented, implementing various RL algorithms such as PPO (Proximal Policy Optimization) and DDPG (Deep Deterministic Policy Gradient) . These methods allow UAVs to learn optimal policies for visiting sensors based on their energy levels, locations, and data priorities. Results show that the use of reinforcement learning significantly improves energy efficiency and reduces processing delays compared to classical approaches like random or greedy strategies.

Keywords : Reinforcement Learning, UAV, IoT, Smart agriculture, Energy optimization, offloading.

ملخص

تهدف هذه المذكرة إلى استخدام تقنيات التعلم المعزز لجمع البيانات بكفاءة في البيئات الزراعية باستخدام الطائرات المسيّرة بدون طيار (UAVs) تركّز الدراسة بشكل خاص على تحسين جانبيين حاسمين: استهلاك الطاقة للطائرات المسيّرة وتقليل زمن جمع البيانات، مع ضمان أقصى تغطية لأجهزة استشعار إنترنت الأشياء (IoT) المنتشرة في الحقول. يقدّم البحث محاكاة شاملة تنفّذ خوارزميات مختلفة مثل

DDPG (Deep Deterministic Policy Gradient) و PPO (Proximal Policy Optimization) تسمح هذه الطرق للطائرات المسيّرة بتعلّم سياسات مثلى لزيارة أجهزة الاستشعار بناءً على حالتها الطاقوية وموقعها وأولوية البيانات. تُظهر النتائج أنّ استخدام التعلم المعزز يحسّن بشكل كبير كفاءة الطاقة ويقلّل من زمن المعالجة مقارنةً بالطرق التقليدية مثل الاستراتيجيات العشوائية أو الجشعة. الكلمات المفتاحية: التعلم المعزز (RL)، الطائرات المسيّرة (UAV)، إنترنت الأشياء (IoT)، الزراعة الذكية، تحسين الطاقة، تفريغ البيانات (offloading).

TABLE DES MATIÈRES

Liste des Abréviations	xiii
Introduction générale	1
1 CONTEXTE DE L'ÉTUDE	1
1.1 Introduction	1
1.2 L'agriculture intelligente et l'Internet des Objets (IoT)	1
1.2.1 Internet des objets	2
1.2.2 Définition de l'agriculture intelligente	2
1.2.3 L'agriculture intelligente et de l'impact croissant des IoT	3
1.2.4 Application de l'IoT dans l'agriculture intelligente	4
1.3 Défis associés à la gestion des données IoT	5
1.3.1 Coût de la technologie	5
1.3.2 Manque de ressources financières	5
1.3.3 Niveau d'alphabétisation des agriculteurs	6
1.3.4 Manque d'intégration entre les systèmes	6
1.3.5 Infrastructures de télécommunications	6
1.3.6 Gestion des données	6
1.4 Consommation énergétique des capteurs et des IoT	7
1.4.1 Principales causes de consommation énergétique	7
1.5 Les drones et l'Internet des objets	7
1.5.1 Fonction des UAV dans la collecte des données IoT	8

1.5.2	Avantages de l'intégration des UAV dans l'IoT agricole	8
1.5.3	Défis liés à l'intégration des UAV dans l'IoT agricole	9
1.5.4	Des techniques de IoT dans l'agriculture	10
1.5.5	Lacunes identifiées dans les solutions actuelles	11
1.5.6	Recommandations clés pour combler les lacunes	12
1.6	Conclusion	12
2	DÉCHARGEMENT DES DONNÉES	13
2.1	Introduction	13
2.2	Déchargement Des Données (Data Offloading)	13
2.2.1	Processus typique de déchargement de données	14
2.3	Stratégies de déchargement des données	16
2.3.1	Déchargement local	16
2.3.2	Déchargement centralisé	16
2.3.3	Déchargement distribué	17
2.4	Critères d'efficacité pour le déchargement	17
2.5	Technologies utilisées dans l'agriculture intelligente	18
2.5.1	Système de positionnement global (GPS)	18
2.5.2	Les technologies de détection	19
2.5.3	Les technologies à taux variable (VRT)	19
2.5.4	Système d'information géographique (SIG)	19
2.5.5	Gestion des cultures	20
2.5.6	Capteurs pour le sol et les plantes	20
2.5.7	Contrôleurs de débit	20
2.5.8	Irrigation de précision	21
2.6	Défis énergétiques dans l'agriculture intelligente	21
2.7	Travaux connexes	21
2.8	Conclusion	23
3	APPRENTISSAGE PAR RENFORCEMENT (REINFORCEMENT LEARNING)	24
3.1	Introduction	24
3.2	Apprentissage automatique	25

3.2.1	Apprentissage profond (Deep Learning) :	25
3.3	Apprentissage par Renforcement	26
3.3.1	Concepts du RL	27
3.3.2	Les étapes du RL	28
3.4	Classification des algorithmes du RL	28
3.4.1	Model-Free Reinforcement Learning (RL sans modèle :)	28
3.4.2	Model-Based Reinforcement Learning (RL avec modèle) :	29
3.5	l'Apprentissage par Renforcement Profond (Deep Reinforcement Learning – DRL)	30
3.5.1	DRL pour résoudre des problèmes de déchargement dans IoT	31
3.5.2	Avantages de l'utilisation du RL pour l'agriculture intelligente	31
3.6	Algorithme courants	32
3.6.1	DDPG (Deep Deterministic Policy Gradient)	32
3.6.2	PPO (on-policy policy)	35
3.6.3	SAC (Soft Actor-Critic)	36
3.7	Conclusion	38
4	SIMULATION ET RÉSULTATS	40
4.1	Introduction	40
4.2	Configuration de l'Environnement de Simulation	41
4.2.1	Paramètres de l'Environnement	41
4.2.2	Outils et plateformes	44
4.3	Algorithmes Comparés	44
4.4	Métriques d'Évaluation	49
4.4.1	Latence Moyenne (ms)	49
4.4.2	Consommation Énergétique (Joules)	49
4.4.3	Taux de Succès (%)	50
4.4.4	Débit Moyen (Throughput, en KB/s)	50
4.4.5	Efficacité Énergétique (Mo/J)	51
4.4.6	Récompense Moyenne	51
4.5	Résultats des Simulations	52
4.5.1	Consommation Énergétique (Joules)	52

4.5.2	Latence Moyenne (ms)	53
4.5.3	Récompense Moyenne	55
4.5.4	Taux de Succès (%)	56
4.5.5	Débit Moyen (Throughput, en KB/s)	58
4.6	Tableau de Comparaison des Algorithmes	59
4.7	Discussion	60
4.8	Conclusion	61
	Conclusion générale	63
	Bibliographie	64

TABLE DES FIGURES

1.1	Utilisation de drones pour l'agriculture.[6]	5
2.1	Système de déchargement de données par drone.[10]	14
2.2	Technologies utilisées dans l'agriculture intelligente.[36]	18
3.1	Les réseaux neuronaux. [37]	26
3.3	Hierarchie des algorithmes de Apprentissage par Renforcement (RL).[35]	29
4.1	Presentation de champ agricole .	43
4.2	Energy Consumption comparison.	52
4.3	Latency comparison.	53
4.4	Rewards comparison.	55
4.5	Success Rate comparison.	56
4.6	Throughput comparison.	58

LISTE DES TABLEAUX

4.1 Comparaison des performances moyennes des algorithmes 59

LISTE DES ABRÉVIATIONS

- RL** : Reinforcement Learning (Apprentissage par Renforcement)
- DRL** : Deep Reinforcement Learning (Apprentissage par Renforcement Profond)
Deep
- IoT** : Internet of Things (Internet des Objets)
- IdO** : Internet des Objets
- UAV** : Unmanned Aerial Vehicle
- SAC** : Soft Actor-Critic
- DDPG** : Deep Deterministic Policy Gradient
- PPO** : Proximal Policy Optimization
- GAE** : Generalized Advantage Estimation
- Q-network** : Réseau Q pour l'estimation des valeurs d'action
- Policy Network** : Réseau définissant la politique de l'agent
- Target Network** : Réseau cible utilisé pour stabiliser l'apprentissage
- Replay Buffer** : Tampon de mémoire pour stocker les transitions passées
- MSE** : Mean Squared Error (Erreur Quadratique Moyenne)
- API** : Application Programming Interface
- SIG** : Système d'Information Géographique
- VRT** : Variable Rate Technology (Technologie à Taux Variable)
- NDVI** : Normalized Difference Vegetation Index (Indice de Végétation)

Edge Computing : Traitement local des données à proximité de la source

D2D : Device-to-Device (Communication directe entre appareils)

Fog Computing : Inter médiaire entre cloud et périphérique

INTRODUCTION GÉNÉRALE

L'agriculture constitue un pilier fondamental de l'économie mondiale, garantissant la sécurité alimentaire et soutenant les efforts de développement durable. Toutefois, face à l'augmentation rapide de la population mondiale et aux défis croissants liés au changement climatique, les méthodes agricoles traditionnelles atteignent leurs limites. C'est dans ce contexte que l'agriculture intelligente s'impose comme une solution innovante, en intégrant les technologies de l'information et de la communication pour accroître la productivité, optimiser l'utilisation des ressources naturelles et réduire l'impact environnemental.

L'Internet des Objets (IoT) joue un rôle central dans cette transition vers une agriculture connectée. Il permet aux agriculteurs de collecter des données en temps réel à l'aide de capteurs intelligents mesurant l'humidité du sol, la température ambiante, ou encore la santé des cultures. Ces données facilitent une gestion plus précise et efficace des exploitations agricoles. Les dispositifs connectés mobilisés incluent notamment des stations météorologiques, des capteurs de terrain et des systèmes d'irrigation intelligents.

Cependant, l'intégration massive de l'IoT dans le secteur agricole soulève plusieurs défis, en particulier en matière de gestion et de traitement des grandes quantités de données générées. Cela nécessite une infrastructure robuste pour assurer la collecte, la transmission et l'analyse efficace des données. De plus, la gestion énergétique représente un enjeu majeur, notamment dans les zones rurales où les ressources sont limitées. Les capteurs alimentés par batterie sont particulièrement vulnérables aux contraintes d'énergie.

Pour répondre à ces problématiques, la recherche se tourne vers l'utilisation de drones (UAVs) capables de collecter périodiquement les données des capteurs au sol, réduisant

ainsi les coûts de communication et la consommation énergétique des dispositifs IoT. Dans ce cadre, l'objectif est d'optimiser le processus de déchargement des données en milieu agricole à l'aide de techniques d'apprentissage par renforcement. Cela inclut l'identification des trajectoires optimales pour les drones, assurant une couverture maximale des capteurs tout en minimisant la consommation énergétique et le délai de transmission. L'efficacité de cette approche est évaluée en comparaison avec des méthodes classiques.

Cette méthode repose sur la modélisation d'un environnement agricole virtuel, dans lequel sont appliqués des algorithmes d'apprentissage par renforcement pour mesurer la performance du système, en particulier en termes de consommation d'énergie et de latence de transmission.

Afin de structurer cette étude, le mémoire est organisé en quatre chapitres principaux :

— **Chapitre 1 : Contexte de l'étude**

Ce chapitre présente les fondements de l'agriculture intelligente et souligne l'importance croissante de l'IoT dans ce domaine. Il aborde également les défis liés à la gestion énergétique des dispositifs connectés, et introduit l'apport potentiel des drones dans cette problématique.

— **Chapitre 2 : Déchargement des données**

Ce chapitre s'intéresse aux différentes méthodes de transfert de données dans les réseaux IoT agricoles. Il met en lumière les enjeux énergétiques associés et le rôle stratégique que peuvent jouer les UAVs. Une revue des travaux de recherche existants y est également proposée.

— **Chapitre 3 : Apprentissage par renforcement**

Il explore les concepts fondamentaux de l'apprentissage par renforcement et décrit comment ces techniques peuvent être appliquées à l'optimisation du déchargement des données. Une attention particulière est portée à l'approche Deep Reinforcement Learning, adaptée à la complexité des environnements agricoles dynamiques.

— **Chapitre 4 : Simulation et résultats**

Ce chapitre détaille le protocole de simulation, les outils utilisés, ainsi que les scénarios expérimentaux. Il expose et analyse les résultats obtenus, en mettant en avant l'efficacité énergétique et les performances du système proposé.

Enfin, ce mémoire se conclut par une synthèse des contributions apportées, ainsi que par des perspectives de recherche futures.

CHAPITRE 1

CONTEXTE DE L'ÉTUDE

1.1 Introduction

L'agriculture est essentielle pour les nations en tant que principale source alimentaire, mais elle doit relever le défi de produire efficacement tout en améliorant la durabilité. Cela nécessite une transition vers des méthodes modernes, avec l'agriculture intelligente comme solution clé. Cette approche intègre diverses technologies, y compris l'Internet des objets, l'intelligence artificielle et les drones, pour optimiser les pratiques agricoles. Le chapitre examine les outils d'agriculture intelligente, en mettant l'accent sur les drones et leurs applications.

1.2 L'agriculture intelligente et l'Internet des Objets (IoT)

L'agriculture intelligente propose un contrôle précis des cultures, une collecte de données et des techniques automatisées, offrant ainsi de nombreux avantages. Un rapport de Beecham intitulé "Towards Smart Farming : Agriculture Embracing the IoT Vision" [1] de 70 % d'ici 2050 pour une population mondiale de 9,6 milliards, face à des défis comme le changement climatique et les ressources limitées. Alors, quelle est la solution pour surmonter ces préoccupations ?

1.2.1 Internet des objets

L'IoT peut être compris comme un réseau d'appareils intelligents interconnectés, capables de communiquer entre eux et de générer des données pertinentes sur l'environnement dans lequel ils opèrent. Ainsi, pratiquement tout appareil capable d'établir une connexion à Internet peut être considéré comme un [objet] dans le contexte de l'IoT, comme les appareils électroménagers, l'électronique, les meubles, les machines agricoles ou industrielles, et même les personnes [2].

Bien que le concept d'IoT ne soit pas nouveau, son adoption s'est accrue ces dernières années, principalement grâce au développement des technologies qui le soutiennent, parmi lesquelles l'amélioration du matériel (avec la réduction de la taille et de la consommation d'énergie qui en résulte), l'amélioration de la connectivité à Internet et entre les appareils via la connexion sans fil, le cloud computing, l'intelligence artificielle et le big data. Tous ces composants technologiques contribuent à la construction d'un réseau d'appareils capables de partager des données et des informations, et d'agir activement en fonction des entrées du réseau.[3]

l'architecture des systèmes IoT est similaire à l'architecture des autres systèmes informatiques mais elle doit prendre en compte les particularités de ce paradigme, telles que les capacités de calcul limitées des appareils, l'identification, la détection et le contrôle des objets distants.[4]

1.2.2 Définition de l'agriculture intelligente

L'agriculture intelligente peut être définie comme l'application de technologies complémentaires aux techniques de production agricole afin de minimiser les déchets et d'accroître la productivité. À cet effet, les fermes intelligentes utilisent des ressources technologiques qui interviennent à différentes étapes du processus de production, telles que la surveillance des plantations, la gestion des sols, l'irrigation, la lutte antiparasitaire, le suivi des livraisons, etc. . Ces ressources comprennent, entre autres, la température, la luminosité, l'humidité, la pression, la concentration chimique du sol, les drones, les caméras vidéo, les systèmes de gestion de l'information agricole, les systèmes de positionnement global (GPS) et les réseaux de communication . L'intégration des ressources technologiques au processus

de production agricole est un enjeu crucial offrant des opportunités aux fournisseurs de technologies, aux fournisseurs d'équipements et de machines agricoles, aux producteurs et aux autres acteurs du secteur. De plus, les fermes intelligentes devraient permettre d'optimiser la production alimentaire en améliorant l'application des nutriments au sol, en réduisant la quantité de pesticides et la consommation d'eau pour l'irrigation.

1.2.3 L'agriculture intelligente et de l'impact croissant des IoT

L'adoption de technologies basées sur les capteurs et l'IoT améliore le rendement des cultures par rapport aux méthodes agricoles conventionnelles

a) Culture sous serre et culture protégée

La culture en environnement contrôlé, comme la culture sous serre, est une méthode d'agriculture intelligente de plus en plus populaire. Les cultures en intérieur sont moins affectées par l'environnement et cultivées à tout moment grâce à des capteurs et des dispositifs de communication. Le succès de cette production dépend de divers facteurs, notamment des structures de hangars, des systèmes d'aération et de la précision des paramètres de surveillance. La surveillance précise des paramètres environnementaux est essentielle, nécessitant plusieurs points de mesure. Les serres intelligentes automatisent les travaux agricoles, protègent les plantes et permettent un contact direct entre l'agriculteur et le consommateur. L'IoT permet une surveillance et un contrôle efficaces de divers paramètres, augmentant ainsi la productivité.

b) Hydroponie

Une sous-division de l'hydroculture, consiste à cultiver des plantes hors sol pour améliorer la culture en serre. Les systèmes hydroponiques permettent un dosage équilibré des nutriments dissous dans l'eau. Les systèmes et capteurs détectent et analysent un large éventail de paramètres à intervalles prédéterminés. La mesure et la surveillance précises de la teneur en nutriments sont essentielles à la croissance des plantes. Un prototype basé sur des capteurs sans fil fournit une solution pour la culture hors-sol en temps réel. Un système hydroponique intelligent et automatisé, intégré à l'IoT, surveille la culture des

plantes via Internet en analysant des paramètres tels que le pH, l'eau, la température et l'humidité. La technique du flux profond assure la circulation continue de la solution nutritive, avec des données acquises par des capteurs intégrés au Raspberry Pi.

c) Agriculture verticale

L'agriculture verticale (AV) permet de maintenir les plantes dans un environnement contrôlé avec précision, réduisant considérablement la consommation de ressources et augmentant la production. L'AV nécessite une surface de sol réduite et est efficace pour augmenter les rendements et réduire la consommation d'eau. [5]

1.2.4 Application de l'IoT dans l'agriculture intelligente

Les systèmes intelligents sont principalement associés à l'IoT, car il représente l'épine dorsale de toutes les applications intelligentes ; les applications les plus importantes de l'agriculture intelligente IoT :

IoT basé sur les drones

Les drones sont utilisés commercialement dans l'agriculture depuis 1980, mais leur utilisation a été limitée en raison des progrès des technologies de communication et de l'IdO (Internet des Objets). Les drones peuvent remplir diverses fonctions, telles que l'irrigation, la surveillance de la croissance des cultures, la plantation, la récolte, l'analyse des sols et la lutte contre les ravageurs. Grâce à l'utilisation de plusieurs caméras, de caméras 3D, de caméras thermiques, de caméras multispectrales et de caméras optiques, les drones peuvent être utilisés pour surveiller la croissance des cultures, les maladies, les indicateurs phytosanitaires, la densité, la prospection de pesticides, la fertilisation, etc. Ils permettent également de calculer directement des indices de végétation à partir d'images multispectrales, comme l'indice de normalisation de la différence (NDVI). Les drones ont un impact significatif sur le développement de l'agriculture verte, depuis la collecte et le traitement des récoltes jusqu'à la moisson et la récolte dans les zones reculées. Comme le montre la Figure 1.1, Utilisation de drones pour l'agriculture.

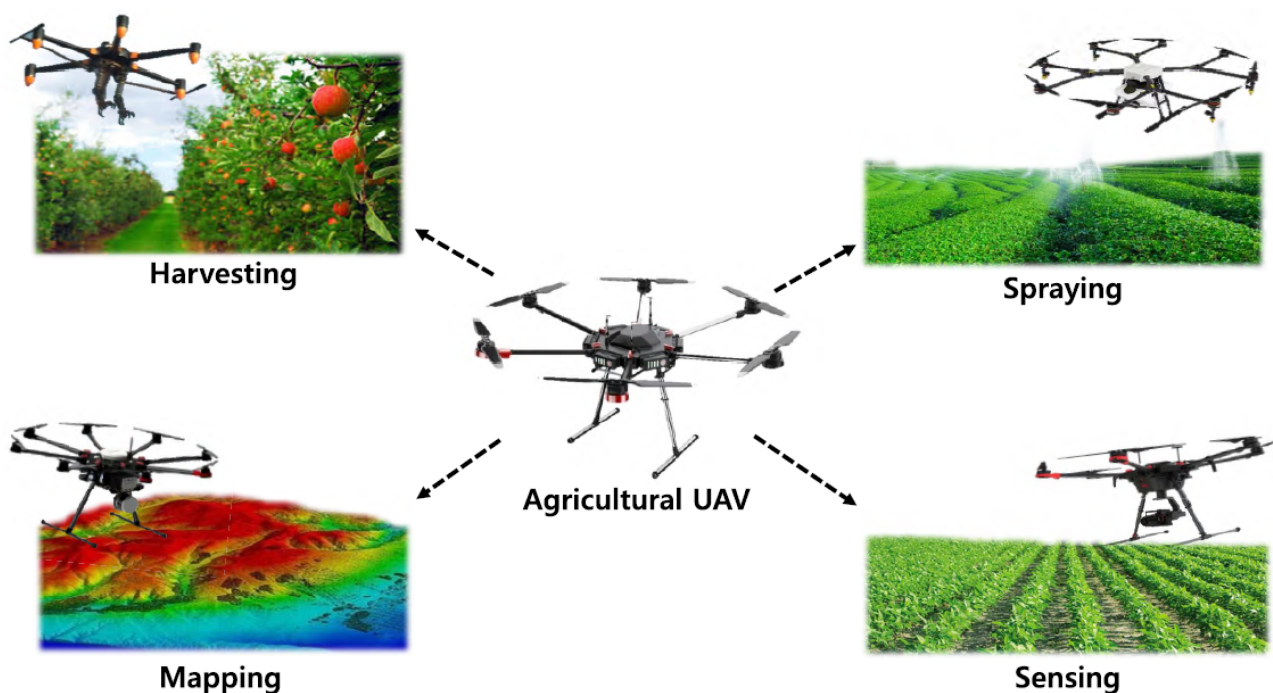


FIGURE 1.1 – Utilisation de drones pour l'agriculture. [6]

1.3 Défis associés à la gestion des données IoT

L'adoption des technologies agricoles a apporté précision, efficacité et réduction des contraintes de temps. Cependant, leur adoption pose encore des défis.

1.3.1 Coût de la technologie

Les technologies existantes sont très efficaces et précises, mais elles pourraient remplacer le travail humain à l'avenir en raison de la pauvreté dans de nombreux pays où le travail manuel constitue la principale source de revenus.

1.3.2 Manque de ressources financières

L'accès au financement de ces technologies nécessite des ressources importantes, que les agriculteurs peinent à obtenir. Cela constitue également un défi pour eux, car ils peuvent ne pas bénéficier de prêts adéquats en raison de catastrophes imprévisibles telles que la sécheresse, les inondations, les ravages et les maladies affectant les cultures.

1.3.3 Niveau d'alphabétisation des agriculteurs

Le niveau de formation des agriculteurs est crucial pour la mise en œuvre de ces technologies, car ils doivent acquérir des connaissances et des compétences sur les nouvelles technologies. L'agriculture traditionnelle est préférée à l'agriculture intelligente, mais les agriculteurs trouvent souvent son utilisation complexe et ne parviennent pas à reconnaître les icônes des applications mobiles.

1.3.4 Manque d'intégration entre les systèmes

L'intégration entre les systèmes constitue un autre défi, car ces technologies nécessitent des outils plus avancés, la gestion des téléphones et une aide à la décision. La communication entre les universités et les groupes interdisciplinaires est essentielle pour améliorer l'efficacité des utilisateurs. Le développement d'un système d'information peut contribuer à générer davantage d'informations et de connaissances.[7]

1.3.5 Infrastructures de télécommunications

Les infrastructures de télécommunications sont également essentielles à la réussite de la mise en œuvre de ces technologies. Les activités rurales sont plus efficaces sur les terres agricoles que sur les terres contaminées. Cependant, l'absence de connexions internet fiables rend la transmission de données difficile, notamment via les téléphones portables et les tablettes. L'agriculture intelligente requiert une connexion internet en temps réel pour l'utilisation des informations.

1.3.6 Gestion des données

La gestion des données constitue un autre défi pour les agriculteurs, qui peinent à organiser et à manipuler les données obtenues par les capteurs. L'IoT joue un rôle crucial dans la collecte, la surveillance et le contrôle de l'approvisionnement en eau. Les capteurs et la connectivité cloud permettent aux agriculteurs de surveiller à distance l'état des sols et des plantes, de remédier aux carences en nutriments et aux maladies, et de lutter contre ces dernières. Le développement des technologies IoT a joué un rôle important dans le secteur agricole moderne, notamment grâce aux infrastructures de communication.

1.4 Consommation énergétique des capteurs et des IoT

Dans de nombreux environnements, il est compliqué de remplacer ou de recharger régulièrement les batteries des dispositifs IoT qui y sont souvent installés. Cependant, leur fonctionnement continu, la collecte et le transfert de données ainsi que leur connectivité sans fil requièrent de l'énergie.

1.4.1 Principales causes de consommation énergétique

Diffusion de données par voie sans fil :

Le Wi-Fi, le Bluetooth consomment beaucoup d'énergie, surtout lorsqu'ils sont utilisés pour des communications fréquentes ou sur de longues distances.

Capteurs actifs en permanence :

Certains capteurs consomment de l'énergie même lorsqu'ils n'envoient pas de données.

Microcontrôleurs et processeurs peu performants :

Les puces employées pourraient avoir de longs cycles d'activité, sans un mode veille efficace.

Algorithmes mal ajustés :

L'exécution de calculs localement ou la transmission de données non traitées plutôt que de données prétraitées peut nécessiter davantage d'énergie.

Taux d'échantillonnage élevé :

L'extraction trop fréquente de données accroît inutilement le coût énergétique.

1.5 Les drones et l'Internet des objets

Les drones sont de plus en plus utilisés par les agriculteurs pour surveiller la croissance des plantations, appliquer des nutriments, de l'eau et des pesticides sur des terrains complexes et à différents stades de culture. Ils offrent une meilleure vaporisation, une surface

et une élution supérieures à celles des appareils conventionnels. Les avancées technologiques permettent d'intégrer les drones à divers contrôleurs et caméras 3D, offrant ainsi des outils complets de gestion des terres. La technologie des drones dans l'agriculture intelligente fournit des informations sur la fertilisation, l'irrigation, la croissance des plantes, la gestion des herbicides, les maladies des cultures et le phénotypage afin d'améliorer les pratiques culturales. Une nouvelle méthode de modélisation 3D est utilisée pour surveiller les paramètres de croissance et déterminer la hauteur des plantes dans des conditions difficiles grâce aux drones.[8]

1.5.1 Fonction des UAV dans la collecte des données IoT

Les drones sont essentiels pour collecter et transmettre les données des pilotes, mais ils nécessitent une planification précise en raison des limitations de temps et d'énergie. Les algorithmes d'intelligence artificielle (IA), tels que l'intelligence de groupe et l'apprentissage par renforcement, peuvent gérer des environnements dynamiques et fournir des solutions presque optimales en temps réel. La littérature sur la collecte de données par drone se concentre sur trois indicateurs : la vitesse des données, l'efficacité et la consommation d'énergie. Une vitesse de vol élevée peut réduire le temps de vol et la zone d'intérêt, mais une consommation d'énergie accrue est nécessaire. La position de vol stationnaire du drone joue un rôle crucial dans la détermination du temps de transmission et de la consommation d'énergie. L'optimisation de la consommation d'énergie, de la localisation des vols stationnaires et de la zone d'intérêt est nécessaire pour améliorer l'acquisition de données. Le vieillissement des drones impacte significativement la trajectoire de vol optimale. L'algorithme Learning Q a été utilisé pour planifier le vol du drone afin d'assurer la livraison des données en temps opportun, réduisant ainsi l'expiration des données.

1.5.2 Avantages de l'intégration des UAV dans l'IoT agricole

a) Diminution des mouvements physiques :

Les drones prennent la place des visites répétées sur le terrain pour la collecte d'informations.

Réduction des déplacements → économie de carburant, d'énergie et de temps.

b) Surveillance et diminution des transmissions :

Grâce à des caméras multi/hyperspectrales, les UAV sont capables d'identifier les zones problématiques.

Cela autorise l'activation exclusive des capteurs ou équipements requis à proximité, ce qui réduit la consommation d'énergie totale du système IoT.

c) Récupération rapide et centralisée des données :

Les drones ont la capacité de survoler rapidement de vastes zones et d'accumuler des informations provenant de divers capteurs éloignés.

Moins de capteurs fixes entraîne une diminution de la consommation d'énergie distribuée.

d) Complémentarité avec les capteurs terrestres :

Les drones ont la capacité de compenser les capteurs qui sont hors service ou momentanément sans alimentation, augmentant ainsi la robustesse énergétique du réseau IoT.

1.5.3 Défis liés à l'intégration des UAV dans l'IoT agricole**a) Limitation de l'autonomie des UAV :**

Les drones sont alimentés par batterie et ont une durée de vol limitée.

Ceci pourrait nécessiter des vols à répétition ou des points de recharge, ce qui entraînerait une augmentation indirecte de la consommation énergétique s'il n'est pas optimisé.

b) Infrastructure pour la recharge et les communications :

L'installation de stations de recharge solaires ou de microgrids agricoles sera nécessaire pour fournir de l'énergie aux UAV, ce qui représente un investissement énergétique initial.

La communication avec les drones (Wi-Fi, LTE, 5G) peut consommer beaucoup d'énergie si elle n'est pas bien gérée.

c) Coût en énergie des traitements de données intégrés :

Les drones peuvent intégrer des modules d'intelligence artificielle pour effectuer un traitement local (edge computing), mais cette opération peut nécessiter plus d'énergie que la simple transmission de données.

d) Conditions climatiques et robustesse :

Les drones sont affectés par le vent, la pluie et les températures extrêmes, ce qui restreint leur utilisation constante et peut conduire à des reprises de vol énergétiquement inefficaces.[9]

1.5.4 Des techniques de IoT dans l'agriculture**Réseaux de capteurs sans fil :**

Servant à recueillir des informations sur le sol, le taux d'humidité, la température, les éléments nutritifs, etc.

Opèrent en temps réel pour fournir des informations sur la condition des cultures et du sol.

Cloud computing & big data :

Les informations recueillies sont envoyées sur le cloud pour être traitées et contribuer à la prise de décisions.

Facilite l'utilisation des systèmes de soutien à la décision intelligents pour optimiser la gestion agricole.

Usage de réseaux de communication :

La connectivité est fournie par les technologies mobiles en fonction des exigences de portée et de consommation.

Capteurs de haute précision :

Surveillance et mesure exactes de l'eau, des fertilisants et des pesticides basées sur les informations en temps réel.

Appareils IoT intégrés dans les drones agricoles :

L'emploi de drones et robots interconnectés pour la cartographie, l'épandage, la récolte ou le suivi des cultures.[8][9]

1.5.5 Lacunes identifiées dans les solutions actuelles**Forte consommation d'énergie :**

Une alimentation permanente est requise pour de nombreux appareils IoT, tels que les capteurs, les drones et les passerelles.

Dans certaines configurations, l'absence de systèmes d'auto-alimentation (comme les panneaux solaires ou la récupération d'énergie) rend leur mise en œuvre peu durable.

Surproduction et surcharge en données :

Les systèmes actuels génèrent d'énormes quantités de données (big data), généralement sans filtrage préliminaire.

Cela conduit à une hausse de la consommation d'énergie pour le traitement et l'envoi des données, notamment en l'absence de filtrage local (edge computing).

Restrictions des réseaux dans les régions rurales :

Les infrastructures 4G/5G ne sont pas systématiquement accessibles, surtout dans les pays en voie de développement ou les régions isolées.

Cela compromet la fiabilité du déchargement des informations et contraint les appareils à maintenir ou réémettre plus souvent, ce qui amplifie la consommation d'énergie.

Manque de standards interopérables :

L'absence de standardisation des protocoles et architectures IoT agricoles entraîne des problèmes d'interopérabilité entre différents fournisseurs et plateformes logicielles.

Cela restreint les options pour la centralisation ou l'amélioration du traitement des informations.

Coût et complexité des solutions avancées

L'emploi de technologies de pointe (robots, drones intelligents) demeure onéreux et gourmand en énergie.

Ils ne prouvent pas toujours leur efficacité énergétique à l'échelle de la mise en œuvre réelle.

1.5.6 Recommandations clés pour combler les lacunes

Edge computing : Incorporer un prétraitement local des données afin d'éviter une saturation du réseau.

Capteurs de réveil : Des capteurs au repos qui ne se mettent en marche que lors de modifications détectées.

Exploitation d'énergie : Mise en œuvre de solutions solaires ou cinétiques sur les capteurs ou les drones.

Priorisation des données : Ne décharger que les données critiques ou anormales.

Normes ouvertes : Promouvoir l'interopérabilité entre les équipements, les logiciels et les plateformes IoT.[8][9]

1.6 Conclusion

Dans ce chapitre, nous explorons l'intégration de l'Internet des objets (IoT) à l'agriculture intelligente, qui a permis des avancées significatives dans la surveillance, la gestion et l'optimisation des ressources agricoles. Cependant, l'utilisation énergivore de capteurs, de drones, de robots et d'appareils de communication reste un défi majeur, notamment dans les zones rurales où l'approvisionnement énergétique est limité. De plus, la gestion de grandes quantités de données collectées sans tri préalable entraîne une congestion du réseau et un gaspillage d'énergie. Malgré des solutions telles que les capteurs basse consommation, le traitement décentralisé et l'intégration de drones, des lacunes subsistent en termes d'indépendance énergétique, d'efficacité de transmission et de standardisation. Pour assurer la durabilité de l'agriculture numérique, il est essentiel de concevoir des systèmes IoT économes en énergie et compatibles avec les besoins des exploitations agricoles de toutes tailles.

2.1 Introduction

Le déchargement des données est une stratégie cruciale pour optimiser les performances et l'utilisation des ressources des systèmes IoT. Il s'agit de déléguer les opérations de traitement des données à des ressources limitées ou à des nœuds ou serveurs plus performants, de réduire la charge de traitement des appareils IoT, d'économiser de l'énergie et d'améliorer les performances globales.

Il existe différentes méthodes d'épuisement des données : totale, partielle, collaborative et dynamique. La mise en œuvre de stratégies efficaces de déchargement des données optimise les performances du système, prolonge l'autonomie des appareils et garantit une utilisation optimale des ressources. L'agriculture intelligente utilise des technologies comme l'Internet des objets, les drones et l'intelligence artificielle pour accroître la productivité agricole. La collecte et le transfert de données sont essentiels pour optimiser le traitement et réduire la latence.

2.2 Déchargement Des Données (Data Offloading)

Déchargement des données correspond à la procédure de transfert des informations produites ou recueillies par un appareil (tels qu'un capteur IoT, un drone) vers un autre système de traitement ou d'entreposage, généralement dans l'objectif d'alléger la charge

informatique ou de communication sur l'appareil initial.

Dans le cadre de l'agriculture intelligente, cela suppose généralement la transmission des informations recueillies sur le terrain (par exemple : humidité, température, images, géolocalisation) vers un serveur local, une passerelle ou le cloud pour analyse, stockage ou prise de décision. Comme le montre la Figure 2.1, Système de déchargement de données par drone.

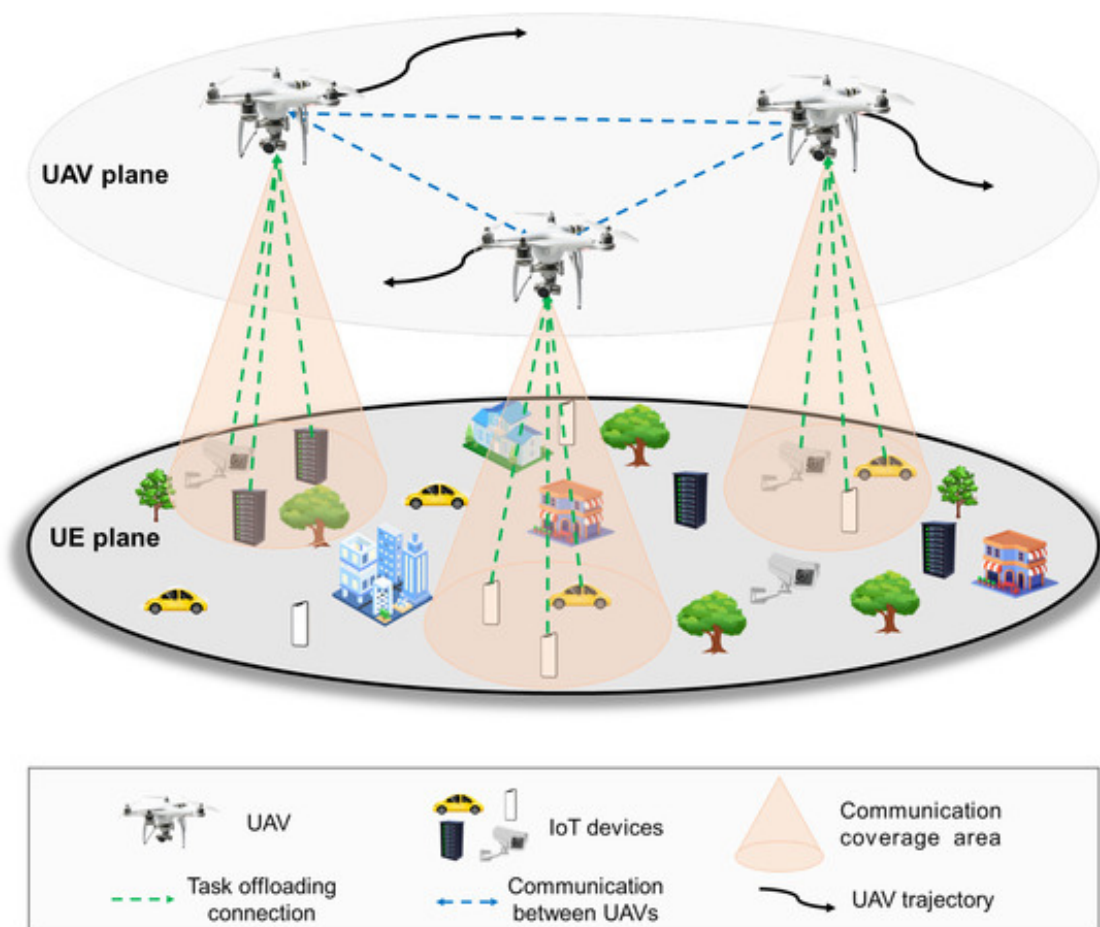


FIGURE 2.1 – Système de déchargement de données par drone.[10]

2.2.1 Processus typique de déchargement de données

Détection du besoin de déchargement :

Le système détecte qu'une tâche consomme excessivement des ressources (batterie, puissance de calcul, etc.) et détermine s'il convient de la décharger.

Sélection de la cible de déchargement :

Le système détermine la destination pour l'envoi des données ou de la tâche :

Serveur cloud distant.

Serveur en périphérie (Edge/Fog).

Appareil voisin (D2D).

Drone (UAV)

Évaluation des conditions réseau :

Le système évalue la disponibilité, la rapidité, la stabilité et l'efficacité énergétique du réseau (par exemple : Wi-Fi, 4G, 5G).

Si les circonstances sont défavorables, il peut patienter pour une opportunité plus propice.

Transfert des données :

Les données ou la tâche sont transmises à la destination sélectionnée par le biais du réseau.

Traitement ou stockage à distance :

Les données sont soit analysées, soit conservées sur une plateforme en nuage, un serveur périphérique ou une station locale. Le système décide à l'avance quand décharger les données, souvent en fonction des attentes (par exemple, la congestion du réseau, la disponibilité de l'énergie).

Retour des résultats (si nécessaire) :

Si la tâche génère un résultat (par exemple : analysis, prédiction), celui-ci est renvoyé au dispositif d'origine.

Mise à jour de l'état :

Le système actualise l'état de la tâche (succès, échec, nouvelle tentative, etc.).

2.3 Stratégies de déchargement des données

2.3.1 Déchargement local

Le traitement ou l'archivage des données se fait directement sur l'appareil de l'Internet des objets ou à proximité par exemple (microcontrôleur, microserveur embarqué).

Avantages :

Économie d'énergie en limitant le transfert de fil.

Plus grande dépendance à l'égard de la connectivité du réseau.

Réactivité rapide (faible latence).

Limitations :

Capacité de calcul et de stockage limitée.

Moins est plus adapté aux analyses complexes ou aux grands ensembles de données.[11]

2.3.2 Déchargement centralisé

Pour être stockées et analysées, les données provenant de plusieurs dispositifs de capture sont envoyées à un seul endroit, généralement un serveur en nuage.

Avantages :

Inclure la capacité de calculer et de stocker des quantités illimitées.

L'accès aux données à l'échelle mondiale, le traitement de données volumineuses et l'analyse d'impact avancée.

Une maintenance et une surveillance facilitées.

Limitations :

Consommation d'énergie élevée associée aux transmissions.

Dépendance à l'égard des connexions réseau (4G, LTE, etc.).

2.3.3 Déchargement distribué

Un certain nombre de nœuds intermédiaires (tels que les passerelles, les serveurs de périphérie et les collecteurs de drones) se partagent le traitement et la transmission des données.

Avantages :

Équilibre entre latence, autonomie et performance.

Plus de résilience en cas de défaillance du réseau.

Diminution du trafic réseau et des frais liés à l'informatique en nuage.

Limitations :

La coordination complexe entre les nœuds.

À chaque niveau, des dispositifs plus sophistiqués sont nécessaires.[11]

2.4 Critères d'efficacité pour le déchargement

Le temps de latence :

Les temps de réponse doivent être aussi courts que possible, en particulier pour les applications sensibles aux informations en temps réel (telles que la réalité augmentée et Véhicule aérien sans pilote UAV).

Consommation d'énergie :

En évitant les traitements locaux intensifs, la décharge devrait permettre de réduire la consommation d'énergie des appareils mobiles.

Fiabilité :

Pour garantir l'intégrité des traitements, la connexion entre l'appareil et le serveur MEC doit être stable et le taux de perte de données faible. Le serveur MEC (Mobile Edge Computing) est un nœud de calcul local proche des objets IoT, qui réduit la latence et la consommation d'énergie.

2.5 Technologies utilisées dans l'agriculture intelligente



FIGURE 2.2 – Technologies utilisées dans l'agriculture intelligente.[36]

Comme le montre la Figure 2.2, Les technologies utilisées dans l'agriculture intelligente sont suivant :

2.5.1 Système de positionnement global (GPS)

Les données relatives à la latitude, à la longitude et à l'altitude sont enregistrées avec précision par le GPS. [12]Le système de positionnement global utilise des satellites pour envoyer des signaux aux récepteurs GPS, ce qui leur permet de déterminer leur position en temps réel et de suggérer continuellement des positions lorsqu'ils sont en mouvement. Ces informations de localisation précises aident les agriculteurs à connaître l'emplacement spécifique des champs, comme les mauvaises herbes, le type de sol, la prévalence des insectes et d'autres obstacles, ce qui leur permet d'administrer les intrants nécessaires

(semences, engrais, herbicides, pesticides et eau) à un champ spécifique.[13]

2.5.2 Les technologies de détection

Notamment la photoélectricité, l'électromagnétisme, la conductivité et les ultrasons, sont utilisées pour évaluer la texture du sol, les niveaux de nutriments, la végétation, l'humidité, la vapeur, l'air et la température. Les données de télédétection permettent de différencier les types de cultures, de classer les parasites et les mauvaises herbes, de localiser le stress et de surveiller la sécheresse. La santé des plantes est affectée par des facteurs tels que l'humidité du sol, la disponibilité des nutriments, l'exposition à la lumière, l'humidité, les précipitations et la couleur des feuilles. Les capteurs détectent les changements et les transmettent aux microcontrôleurs pour qu'ils prennent les mesures nécessaires. Les capteurs sans fil sont essentiels pour collecter des données sur l'état des cultures et fournir d'autres informations. Ils peuvent être intégrés à des outils et machines agricoles avancés.[14]

2.5.3 Les technologies à taux variable (VRT)

L'échantillonnage du sol en grille est une méthode utilisée pour collecter des échantillons de sol à partir d'une grille systématique afin d'établir une carte pour chaque paramètre. Ces données sont chargées dans un applicateur à taux variable, qui contrôle les changements dans l'apport d'engrais. Les nouvelles technologies telles que l'échantillonnage des sols en grille peuvent améliorer la gestion de la fertilité des sols et évaluer la distribution des éléments nutritifs. La variabilité du phosphore et du potassium en fonction du champ nécessite une application uniforme d'engrais et de fumier pour l'agriculture de précision.[15]

2.5.4 Système d'information géographique (SIG)

Le SIG est un outil matériel et logiciel qui collecte, stocke, récupère, analyse les attributs et localise les données pour les cartes et l'analyse statistique. Il fournit des informations sur les types de sol, l'état des nutriments, la topographie, l'irrigation, les applications chimiques et la production agricole.

2.5.5 Gestion des cultures

Les images satellite offrent des informations en temps réel sur l'état des sols et la performance des cultures, permettant aux agriculteurs de contrôler des facteurs tels que les semences, les engrais et les pesticides afin d'améliorer le rendement et l'efficacité. Les indices de végétation, en particulier l'indice de végétation par différence normalisée (NDVI), sont utilisés pour évaluer la santé de la végétation et la production agricole.[16] Les méthodes de surveillance des cultures comparent l'état actuel des cultures aux saisons précédentes ou normales, ce qui permet d'estimer rapidement le rendement des cultures. La gestion automatisée des champs offre des fonctionnalités plus complexes, notamment l'acquisition de données, le traitement, la surveillance, la prise de décision et la gestion des opérations agricoles, y compris la production de cultures, les profits et les pertes, les prévisions météorologiques, la cartographie des champs et le suivi des éléments nutritifs du sol.[17]

2.5.6 Capteurs pour le sol et les plantes

Les capteurs de sol et les vêtements pour plantes surveillent en temps réel les signaux physiques et chimiques du sol, tels que l'humidité, le pH, la température et les polluants, afin d'optimiser les conditions de croissance des cultures et d'augmenter les rendements. Les éléments nutritifs les plus importants pour la production végétale sont les matières organiques du sol, l'azote, le phosphore et le potassium. Les capteurs basés sur la réflectance NIR (Near-Infrared Reflectance) mesurent les variations de l'azote du sol en surface et sous la surface, tandis que la technologie de spectrophotométrie NIR prédit l'azote et le phosphore du sol. Les capteurs de conductivité électrique apparente du sol recueillent des informations sur la surface du champ.[18]

2.5.7 Contrôleurs de débit

En suivant la vitesse des camions sur le terrain et en ajustant le débit de matériau en temps réel pour atteindre le taux souhaité, les régulateurs de taux sont conçus pour réguler le taux de livraison des intrants. Les régulateurs de débit sont fréquemment utilisés comme systèmes autonomes .

2.5.8 Irrigation de précision

Afin de surveiller le sol et les conditions climatiques et d'évaluer l'irrigation, les récents progrès des systèmes d'irrigation ont permis d'introduire des machines d'irrigation, des commandes de mouvement, des contrôleurs basés sur le GPS, des technologies de capteurs et des communications sans fil.

Les systèmes d'irrigation sont capables de gérer des facteurs tels que le débit et la pression, afin d'accroître l'efficacité de l'utilisation de l'eau par les cultures. Ces technologies sont très prometteuses, mais elles doivent encore être développées avant d'être commercialisées.[19]

2.6 Défis énergétiques dans l'agriculture intelligente

L'agriculture intelligente s'appuie principalement sur des réseaux de capteurs IoT (Internet des objets) pour surveiller l'environnement, les cultures, le sol et même l'irrigation. Cependant, ces dispositifs fonctionnent souvent dans des zones rurales éloignées, avec un accès limité à l'énergie électrique. Ainsi, la consommation d'énergie des capteurs, des antennes, des passelles et des unités informatiques embarquées devient une contrainte critique.

Un système énergétique inefficace peut entraîner une fréquence importante de remplacement des batteries, une durée de vie plus courte du réseau et des interruptions de service. Ces limitations augmentent les coûts de maintenance, compromettent la longévité des systèmes agricoles IoT et diminuent la fiabilité des décisions basées sur les données collectées. Il est donc essentiel d'utiliser des techniques d'économie d'énergie telles que l'éveil adaptatif, l'agrégation de données ou même le passage à l'informatique de périphérie pour accroître l'autonomie des appareils et garantir la continuité du service.[20]

2.7 Travaux connexes

L'agriculture intelligente et le déchargement des données représentent un sujet de recherche très actif. De nombreuses études ont été menées sur différents aspects de cette thématique. Dans la référence [21], les auteurs présentent un aperçu de l'agriculture intel-

ligente et du transfert de données, en exploitant les drones (UAV) associés aux technologies de l'informatique mobile de pointe (MEC) et de l'Internet des objets (IoT), dans le but d'optimiser la gestion des ressources et d'augmenter la productivité agricole. Cette étude propose une stratégie d'apprentissage par renforcement profond (Deep Reinforcement Learning, DRL) multi-agents qui gère de manière adaptative le transfert de tâches, réduisant ainsi les délais et la consommation d'énergie tout en garantissant la sécurité des transmissions de données. Cette approche favorise des solutions durables et économes en énergie pour les futurs réseaux agricoles intelligents.

Dans la référence [22], les chercheurs proposent une vue d'ensemble du transfert de données dans l'agriculture intelligente, impliquant l'échange d'informations géographiquement dispersées entre agriculteurs et serveurs périphériques (Edge Servers - ES), afin d'améliorer les contrôles intelligents en temps réel tels que la fertilisation ou l'irrigation. L'étude introduit une méthode de placement orientée service (Service Offloading Placement - SOP) permettant de répartir les charges entre les ES et d'optimiser les délais de transmission. Cette stratégie prend en compte les préférences des agriculteurs dans la mise en œuvre des services, contribuant ainsi à l'automatisation et à l'amélioration du rendement des cultures.

La référence [23] explore des solutions de déchargement de données mobiles basées sur la confiance. L'étude propose un système fondé sur un modèle de confiance pour améliorer l'efficacité du déstaging des données dans un environnement agricole intelligent, notamment en cas de perturbations de connectivité. L'algorithme employé permet d'économiser de l'énergie, de réduire les délais et d'assurer une meilleure prise de décision en garantissant que seuls des contributeurs fiables participent au processus.

L'article de la référence [24] aborde l'utilisation des drones pour la diffusion de données agricoles intelligentes, en se concentrant sur des tâches telles que la catégorisation d'images et la surveillance de l'environnement. Il propose une approche d'apprentissage profond basée sur le Deep Q-Learning afin d'optimiser la consommation d'énergie et de réduire la latence, ce qui améliore l'efficacité opérationnelle des réseaux agricoles.

La référence [25] traite d'une stratégie de déstaging des calculs dans l'agriculture intelligente en s'appuyant sur l'informatique en périphérie multi-accès (MEC). Cette approche permet d'accroître l'efficacité du traitement des données malgré les limitations énergé-

tiques des capteurs, tout en favorisant une surveillance durable grâce à une planification optimisée des ressources et à l'exploitation de l'énergie solaire.

La référence [26] présente une stratégie de migration des données impliquant la gestion de sources multiples, notamment les composants IoT. Le transfert efficace de données y est considéré comme essentiel pour le traitement et le stockage optimisés, facilitant des applications agricoles telles que la gestion de l'eau ou la maintenance des équipements.

Dans la référence [27], l'accent est mis sur l'importance de systèmes efficaces de gestion de données, permettant un traitement en temps réel ou en mode batch, et favorisant une meilleure utilisation des données générées par les dispositifs IoT dans les exploitations agricoles.

Enfin, la référence [28] examine le déchargement des données à la périphérie, à proximité de la source, plutôt que via des serveurs cloud centralisés. Cette approche réduit la latence, améliore la prise de décision en temps réel et optimise la gestion des ressources, éléments essentiels à l'efficacité des opérations agricoles.

2.8 Conclusion

la décharge de données est une solution stratégique essentielle pour répondre aux demandes informatiques et énergétiques des systèmes IoT dans l'agriculture intelligente. En utilisant une variété d'approches, qu'elles soient distribuées, centralisées ou locales, cette stratégie permet d'optimiser la gestion des ressources tout en réduisant la latence et la consommation d'énergie. Cette dynamique est renforcée par l'incorporation de technologies de pointe comme les drones et les caméras intelligentes, qui facilitent la collecte et l'analyse des données agricoles. Cependant, il existe encore des problèmes, tels que ceux liés à la couverture du réseau, l'indépendance énergétique et la coordination entre les nombreux nœuds du système. Ces résultats soulignent l'importance des stratégies d'optimisation complémentaires, telles que celles basées sur l'intelligence artificielle, pour assurer une agriculture numériquement durable et efficace.

CHAPITRE 3

APPRENTISSAGE PAR RENFORCEMENT (REINFORCEMENT LEARNING)

3.1 Introduction

À l'ère du digital, les machines ne se limitent plus à suivre des directives immuables. Elles ont maintenant la capacité d'apprendre de manière autonome à partir des données. C'est à ce moment-là que le Machine Learning (ML) prend son importance : une technologie innovante qui donne la capacité aux systèmes informatiques de progresser au fil de l'expérience, sans nécessiter d'être spécifiquement programmés à chaque phase. Aujourd'hui, le Machine Learning occupe une place centrale dans de nombreuses applications.

Le Deep Reinforcement Learning (DRL) est l'un des sous-domaines les plus prometteurs de l'apprentissage automatique. Cette méthode s'inspire de la façon dont les êtres humains acquièrent des connaissances : en faisant des essais et en apprenant directement par l'expérience. Dans le cadre de l'apprentissage par renforcement profond, un agent (qu'il s'agisse d'un robot ou d'un logiciel) se forme à la prise de décisions appropriées en interagissant avec son environnement.

3.2 Apprentissage automatique

C'est un domaine de l'informatique qui donne aux ordinateurs la capacité d'apprendre à partir de données, de prendre des décisions et de faire des prédictions sans être explicitement programmés. Il comprend des techniques telles que l'apprentissage supervisé, non supervisé et par renforcement. [30]

3.2.1 Apprentissage profond (Deep Learning) :

Il s'agit de la branche la plus avancée et complexe de l'apprentissage automatique, qui utilise des réseaux de neurones à plusieurs couches pour identifier les caractéristiques complexes dans les données.

L'apprentissage profond est un sous-domaine de l'apprentissage supervisé dans lequel l'intervention humaine dans l'extraction des caractéristiques est réduite et remplacée par les couches profondes des réseaux neuronaux.[30] Les techniques de DL incluent notamment :

a) Réseaux de neurones artificiels (ANNs) :

Inspirés du réseau neuronal humain. Un réseau neuronal de base comprend des neurones artificiels interconnectés en couches : une couche d'entrée, une ou plusieurs couches cachées, et une couche de sortie. Comme le montre la Figure 3.1 :

- **La couche d'entrée** : reçoit les informations provenant du monde extérieur. Les nœuds d'entrée traitent les données, les analysent ou les classent, puis les transmettent à la couche suivante.
- **La couche cachée** : reçoit ses entrées de la couche d'entrée ou d'autres couches cachées. Chaque couche cachée analyse la sortie de la couche précédente, la retravaille, puis la transmet à la couche suivante. Il est à noter que l'information circule dans une seule direction, comme dans les réseaux Feedforward (à propagation avant).
- **La couche de sortie** : fournit le résultat final de tout le traitement des données effectué par le réseau de neurones artificiel. [31]

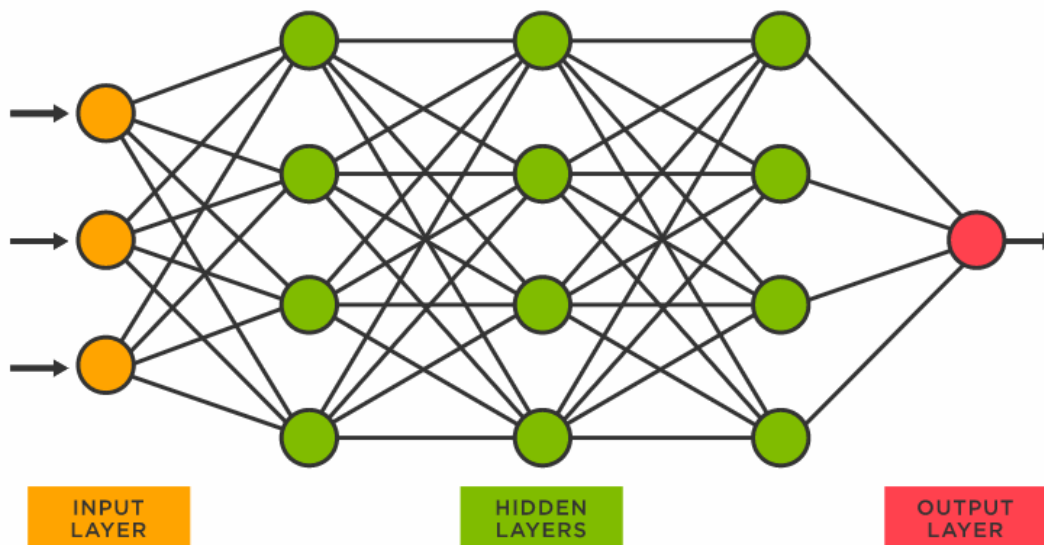


FIGURE 3.1 – Les réseaux neuronaux. [37]

b) les réseaux de neurones profonds (DNN) :

Les modèles d'apprentissage automatique basés sur les réseaux neuronaux profonds s'inspirent de la manière dont le cerveau humain fonctionne. Ils sont constitués de multiples strates de neurones synthétiques, agencées pour acquérir de manière autonome des représentations sophistiquées à partir de données brutes.

3.3 Apprentissage par Renforcement

L'apprentissage par renforcement (Reinforcement Learning RL) est une forme de machine learning où un agent apprend à interagir avec un environnement en effectuant des actions qui maximisent un signal de récompense cumulatif. L'objectif de l'apprentissage par renforcement est de former un agent à prendre des décisions qui contribuent à atteindre un objectif précis en choisissant les actions les plus appropriées. L'agent acquiert ses connaissances en recevant des retours de l'environnement, sous forme de récompenses positives ou négatives. [32] Ce genre de formation nous est très utile lorsqu'il s'agit de discuter de la conduite autonome, car nous savons que les véhicules autonomes évoluent dans des environnements incertains, où ils doivent prendre des décisions et effectuer des actions en fonction de leur perception du monde, comme déterminer le moment idéal pour freiner ou accélérer en se basant sur leur vitesse actuelle et la distance qui les sépare

d'autres véhicules. L'apprentissage par renforcement permet à l'agent d'apprendre de ses propres expériences via l'essai et l'erreur. Cela est particulièrement utile en termes de mise à l'échelle dans le monde réel et des milliers de scénarios impossibles à maîtriser uniquement grâce aux données d'experts.

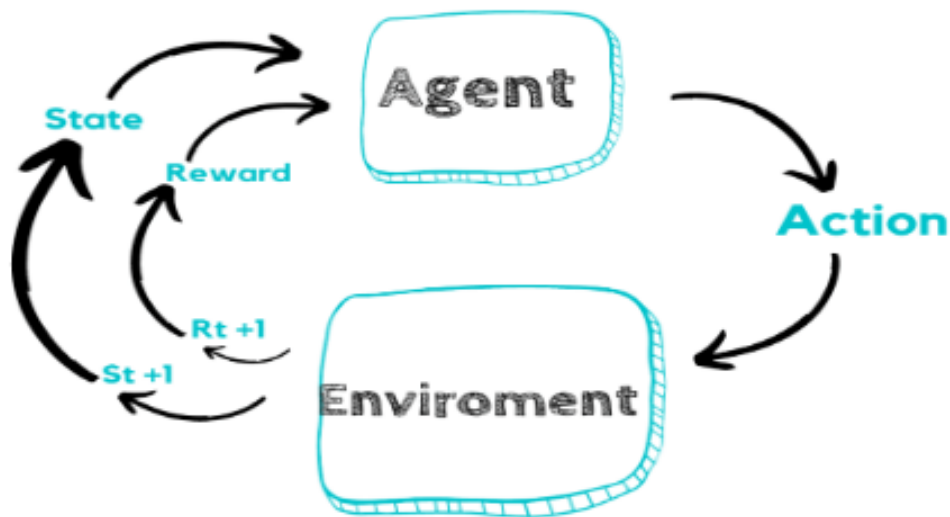


FIGURE 3.2 – Apprentissage par Renforcement (RL)

3.3.1 Concepts du RL

Comme le montre la Figure 3.3 :

- **Agent** : Le système qui prend des décisions.
- **Environnement** : Le monde dans lequel l'agent évolue.
- **État (state, (S))** : La situation actuelle de l'environnement.
- **Action (A)** : Ce que l'agent choisit de faire.
- **Récompense (Reward, (R))** : Le retour (positif ou négatif) que l'agent reçoit.
- **Politique (Policy, ())** : La stratégie de l'agent pour choisir ses actions.
- **Fonction de valeur (V)** : Estimation de la récompense future.
- **Exploration vs Exploitation** : Un dilemme des agents RL : explorer pour découvrir de nouvelles actions ou exploiter les actions connues pour maximiser les récompenses.

3.3.2 Les étapes du RL

- **Observation et Exploration** : L'agent examine l'état présent de son environnement et, en se basant sur cette analyse, détermine la marche à suivre. Par la suite, il analyse les conséquences de ses actes.
- **Récompense et Exploitation** : Suite à l'exécution d'une action, l'agent obtient une rétroaction de son environnement, qu'elle se manifeste sous forme de récompense ou de sanction. La récompense est bénéfique si l'action était appropriée et efficace, ou défavorable (sanction) si l'action était inappropriée. L'état de l'environnement évolue en conséquence.

L'étape d'exploitation implique de tirer parti de ces retours pour orienter les actions prochaines de l'agent.
- **Apprentissage** : L'agent utilise les informations obtenues pour perfectionner ses actions futures, c'est le processus d'apprentissage. Le processus observation-action-récompense se continue jusqu'à l'achèvement de l'apprentissage.
- **Appréciation des Actions** : En se basant sur la gratification obtenue, l'agent est capable d'évaluer si une action doit être reproduite dans des circonstances analogues (si elle est positive) ou évitée (si elle est négative). Ce processus d'interaction constante permet à l'agent de progresser au fil du temps.

3.4 Classification des algorithmes du RL

Les algorithmes RL peuvent être classés en deux grandes catégories :

Comme le montre la Figure 3.3, les algorithmes sans modèle et les algorithmes basés sur un modèle.

3.4.1 Model-Free Reinforcement Learning (RL sans modèle :)

Le RL sans modèle est une méthode où l'agent acquiert ses connaissances directement à partir des échanges avec l'environnement, sans élaborer de modèle prévisionnel de ce dernier. L'agent perfectionne sa prise de décision par le biais de l'essai-erreur, exploitant les récompenses observées pour peaufiner sa stratégie. [33] Les algorithmes sans modèle

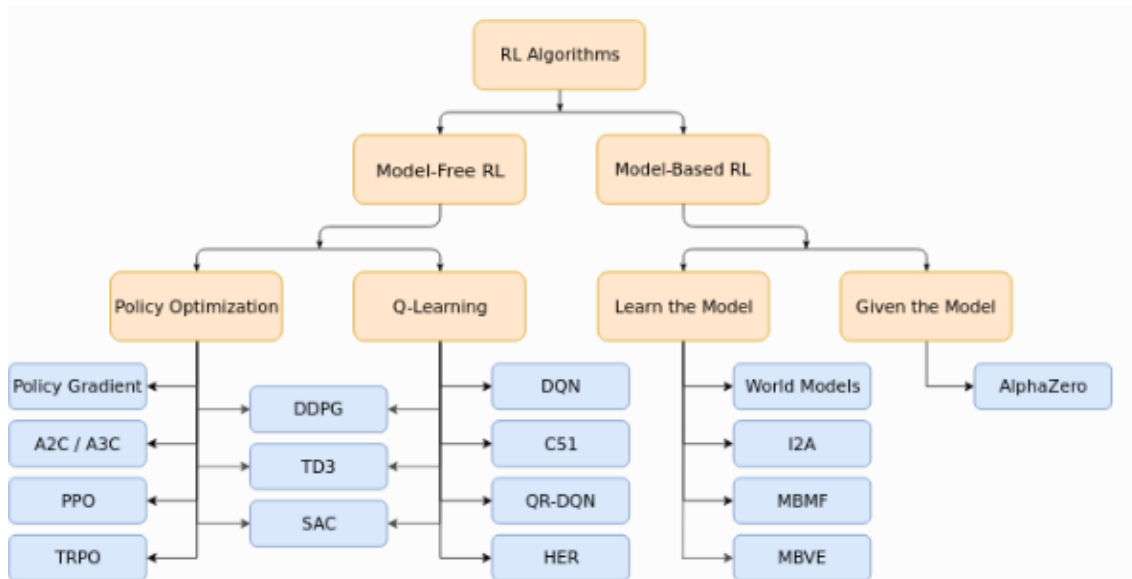


FIGURE 3.3 – Hiérarchie des algorithmes de Apprentissage par Renforcement (RL).[35]

(MFRL) ne génèrent pas de représentation explicite de l’environnement. Ils opèrent généralement par le biais de la méthode d’essais et d’erreurs de l’agent, en interagissant directement avec l’environnement à travers des actions dans le but de déterminer la politique optimale.

Il est recommandé d’opter pour l’apprentissage par renforcement sans modèle lorsque l’environnement est étendu, complexe et ardu à décrire. On recommande aussi cette option quand l’environnement est incertain et en constante évolution.[34]

3.4.2 Model-Based Reinforcement Learning (RL avec modèle) :

Le RL basé sur un modèle signifie que l’agent élabore une représentation interne des dynamiques environnementales et s’en sert pour anticiper des états futurs, prévoir des récompenses et perfectionner ses actions de façon efficiente. [34]

L’Apprentissage par Renforcement Basé sur un Modèle (MBRL) est une approche où l’agent s’efforce de saisir et de modéliser la dynamique de son environnement. À l’opposé des méthodes MFRL, où l’agent acquiert ses connaissances uniquement par des interactions directes avec son environnement. Le MBRL implique la modélisation de l’environnement afin de reproduire ses dynamiques (permettant de tester et d’évaluer diverses stratégies plus rapidement et en toute sécurité, sans interaction directe avec le monde réel). Cette simulation permet à l’agent de prévoir les résultats et de planifier ses actions

en conséquence, optimisant ainsi la prise de décision et minimisant le besoin d'interactions concrètes.[35]

La force majeure des méthodes MBRL se trouve dans leur efficacité. Elles requièrent moins d'interactions avec l'environnement pour mettre en place une politique efficace. Il y a deux catégories de techniques pour le MBRL :

a) Apprendre le modèle :

Ces techniques essaient d'élaborer un modèle de l'environnement en comprenant le fonctionnement des transitions et des récompenses. Parmi les algorithmes fréquemment utilisés basés sur la technique d'apprentissage du modèle, on trouve : l'Estimation de Valeur Basée sur le Modèle (MBVE), les Agents Augmentés par l'Imagination (I2A), etc.

b) Application du modèle :

Cette méthode présuppose que le modèle de l'environnement est déjà connu ou mis à disposition. Cet agent exploite ce modèle pour organiser et perfectionner ses actions. Par exemple, l'algorithme Alpha Zero.

3.5 l'Apprentissage par Renforcement Profond (Deep Reinforcement Learning – DRL)

L'apprentissage par renforcement profond (DRL) est une branche de l'intelligence artificielle appartenant à l'Apprentissage automatique (Machine Learning) qui combine de l'apprentissage par renforcement (RL) classique et de l'apprentissage profond (Deep Learning), d'autre terme le DRL est combinaison du RL avec les réseaux de neurones profonds (DNN). Il s'agit d'un domaine où un agent apprend à prendre des décisions complexes en utilisant des réseaux de neurones profonds pour approximer des fonctions (comme la politique ou la fonction de valeur), ce qui permet de traiter des environnements très complexes avec des observations de haute dimension.

3.5.1 DRL pour résoudre des problèmes de déchargement dans IoT

Le déchargement des données dans les systèmes IoT consiste à déplacer certaines tâches de traitement des objets connectés vers des serveurs de calcul plus puissants (edge, fog ou cloud) afin d'économiser de l'énergie, de réduire la latence ou d'améliorer la fiabilité. Cependant, le choix du temps, de la quantité et de la destination de déchargement est un problème d'optimisation séquentielle influencé par :

- l'état de la batterie.
- la congestion du réseau.
- les priorités des tâches.
- la dynamique environnementale.

3.5.2 Avantages de l'utilisation du RL pour l'agriculture intelligente

Des capteurs IoT recueillent des informations environnementales (comme l'humidité du sol, la température, les précipitations, etc.) qui nécessitent un traitement ou une transmission afin de faciliter la prise de décisions (comme l'irrigation, la fertilisation, la détection de maladies, etc.). L'emploi de l'apprentissage par renforcement (RL) autorise une adaptation dynamique des tactiques de déchargement de données selon plusieurs facteurs clés :

Adaptation aux conditions fluctuantes :

Si le réseau est instable (zones rurales), l'agent RL peut retarder ou compresser les données avant de les transmettre.

En cas de fluctuations climatiques, les priorités de déchargement peuvent changer (par exemple, une fréquence de mesure accrue avant une pluie attendue).

Optimisation de l'énergie :

Les capteurs alimentés par l'énergie solaire ou par batterie peuvent ajuster leur fréquence d'envoi de données ou le choix de la destination selon leur niveau d'énergie.

Le RL a la capacité de réduire les transmissions non nécessaires tout en garantissant la transmission des données cruciales.

Diminution du délai de réponse et amélioration de la prise de décision :

À l'aide du edge computing, l'agent est capable de traiter localement les données essentielles pour une prise de décision rapide (par exemple, déclencher un système d'irrigation).

Le RL offre la possibilité de déterminer quand procéder à une décharge locale ou globale, en se basant sur les délais jugés acceptables.

3.6 Algorithme courants

3.6.1 DDPG (Deep Deterministic Policy Gradient)

DDPG est un algorithme d'apprentissage par renforcement profond, hors politique et sans modèle, basé sur la critique des acteurs, conçu pour gérer les espaces d'action continus. Il combine les concepts du gradient de politique déterministe (DPG) et du Q-Learning profond (DQN), et est particulièrement utile lorsque l'environnement requiert des actions précises et fluides (par exemple, la robotique, le contrôle de drones, la conduite autonome).

Algorithm 1: Deep Deterministic Policy Gradient (DDPG) Algorithm

```

1 Initialize ;;
2   Actor network  $\mu(s | \theta^\mu)$  with random weights  $\theta^\mu$ ;
3   Critic network  $Q(s, a | \theta^Q)$  with random weights  $\theta^Q$ ;
4   Target networks ;;
5      $\mu' \leftarrow \mu$  with weights  $\theta^{\mu'}$ ;
6      $Q' \leftarrow Q$  with weights  $\theta^{Q'}$ ;
7   Replay buffer  $\mathcal{B} \leftarrow \emptyset$ ;
8   Exploration noise process  $\mathcal{N}$  (e.g., Ornstein-Uhlenbeck or Gaussian);
9 for  $episode = 1$  to  $M$  do
10   Initialize a random process  $\mathcal{N}$  for action exploration;
11   Receive initial state  $s_0$ ;
12   for  $t = 1$  to  $T$  (episode length) do
13     Select action :  $a_t = \mu(s_t | \theta^\mu) + \mathcal{N}_t$ ;
14     Execute action  $a_t$ , observe reward  $r_t$  and new state  $s_{t+1}$ ;
15     Store transition  $(s_t, a_t, r_t, s_{t+1})$  in replay buffer  $\mathcal{B}$ ;
16     Sample random minibatch of  $N$  transitions  $(s_i, a_i, r_i, s_{i+1})$  from  $\mathcal{B}$ ;
17     Compute target Q-value ;;
18        $y_i = r_i + \gamma \cdot Q'(s_{i+1}, \mu'(s_{i+1} | \theta^{\mu'}) | \theta^{Q'})$ ;
19     Update critic by minimizing loss ;;
20        $L = \frac{1}{N} \sum_i [Q(s_i, a_i | \theta^Q) - y_i]^2$ ;
21     Update actor using sampled policy gradient ;;
22        $\nabla_{\theta^\mu} J \approx \frac{1}{N} \sum_i \nabla_a Q(s, a | \theta^Q) \Big|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s_i | \theta^\mu)$ ;
23     Soft update target networks ;;
24        $\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'}$ ;
25        $\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}$ ;
26   end
27 end

```

Signification des symboles de l'algorithme DDPG

- θ^μ : Paramètres (poids) du **réseau acteur principal** $\mu(s|\theta^\mu)$, qui prédit l'action optimale à partir d'un état donné.
- θ^Q : Paramètres du **réseau critique principal** $Q(s, a|\theta^Q)$, qui estime la valeur de chaque paire état-action.
- $\theta^{\mu'}$: Paramètres du **réseau acteur cible** $\mu'(s|\theta^{\mu'})$, mis à jour lentement à partir de θ^μ pour stabiliser l'apprentissage.
- $\theta^{Q'}$: Paramètres du **réseau critique cible** $Q'(s, a|\theta^{Q'})$, une copie lente du critique principal.
- τ : Taux de **mise à jour douce** (*soft update*) utilisé pour ajuster les poids des réseaux cibles : typiquement $\tau = 0.001$.
- γ : **Facteur d'actualisation** (discount factor), qui détermine l'importance des récompenses futures. Typiquement $\gamma \in [0.9, 0.99]$.
- \mathcal{N}_t : Processus de **bruit d'exploration** ajouté à l'action pour promouvoir l'exploration (bruit gaussien ou d'Ornstein-Uhlenbeck).
- a_t : Action choisie à l'instant t selon la politique perturbée.
- r_t : Récompense obtenue après exécution de l'action a_t .
- s_t, s_{t+1} : États avant et après exécution de l'action a_t .
- \mathcal{B} : **Buffer de jouabilité** utilisé pour stocker les transitions (s_t, a_t, r_t, s_{t+1}) .
- N : Taille du **minibatch** prélevé aléatoirement depuis le buffer pour mettre à jour les réseaux.
- L : **Fonction de perte** du critique, calculée comme l'erreur quadratique moyenne entre la valeur estimée et la cible.
- J : **Objectif de la politique** (acteur), visant à maximiser la valeur Q moyenne sur le minibatch.
- y_i : **Cible d'apprentissage** pour le critique, calculée à l'aide des réseaux cibles.

3.6.2 PPO (on-policy policy)

Il s'agit d'un algorithme d'optimisation des politiques qui améliore la politique en effectuant des mises à jour prudentes, évitant ainsi des changements drastiques de politique à l'aide d'une fonction objective tronquée.

Algorithm 2: Proximal Policy Optimization (PPO) with Clipping

```

1 Initialize ;;
2   Policy parameters  $\theta$ , Value function parameters  $\phi$ ;
3   Clip threshold  $\epsilon$ , Discount factor  $\gamma$ ;
4   GAE parameter  $\lambda$ , Target KL-divergence  $\delta$ ;
5 for  $iteration = 1, 2, \dots$  do
6   | Collect trajectories  $\{\tau_i\}$  using policy  $\pi_\theta$ ;
7   | Compute advantages  $\hat{A}_t$  using GAE( $\gamma, \lambda$ );
8   | Compute rewards-to-go  $\hat{R}_t$ ;
9   | for  $epoch = 1, 2, \dots, K$  do
10  | | Sample minibatches  $\mathcal{M}$  from trajectories;
11  | | for  $minibatch \{s_t, a_t\} \in \mathcal{M}$  do
12  | | | Compute ratio :  $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ ;
13  | | | Compute clipped objective ;;
14  | | |  $L_t^{CLIP}(\theta) = \min \left( r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A}_t \right)$ ;
15  | | | Update policy :  $\theta \leftarrow \theta + \alpha_\theta \nabla_\theta \frac{1}{|\mathcal{M}|} \sum L_t^{CLIP}(\theta)$ ;
16  | | | Update value function ;;
17  | | |  $\phi \leftarrow \phi - \alpha_\phi \nabla_\phi \frac{1}{|\mathcal{M}|} \sum \left( V_\phi(s_t) - \hat{R}_t \right)^2$ ;
18  | | end
19  | end
20 end

```

Signification des symboles de l'algorithme PPO

- θ : Paramètres du **réseau de politique** $\pi_\theta(a|s)$, utilisés pour générer les actions à partir des états.
- ϕ : Paramètres du **réseau de fonction de valeur** $V_\phi(s)$, utilisé pour estimer la ré-

compense attendue à partir d'un état.

- ϵ : Seuil de **clipping** utilisé pour limiter les mises à jour de la politique, typiquement $\epsilon \in [0.1, 0.3]$.
- γ : **Facteur de discount** utilisé pour pondérer les récompenses futures, généralement proche de 1 (ex. $\gamma = 0.99$).
- λ : Paramètre de **Generalized Advantage Estimation (GAE)**, permettant un compromis entre biais et variance dans l'estimation des avantages.
- δ : Seuil cible pour la **divergence KL** entre les anciennes et nouvelles politiques, servant à détecter des mises à jour trop agressives.
- τ_i : Trajectoire collectée en interagissant avec l'environnement, contenant une séquence d'états, d'actions et de récompenses.
- \hat{A}_t : **Avantage estimé** à l'instant t , calculé à l'aide de GAE pour guider la mise à jour de la politique.
- \hat{R}_t : **Récompense cumulée estimée** à l'instant t (aussi appelée reward-to-go), utilisée pour entraîner le réseau de valeur.
- $r_t(\theta)$: **Ratio de probabilité** entre la politique courante et l'ancienne .
- $L_t^{\text{CLIP}}(\theta)$: **Fonction objectif clipée**.
- \mathcal{M} : **Minibatch** d'exemples d'entraînement tiré des trajectoires collectées.
- α_θ : Taux d'apprentissage pour la mise à jour des paramètres de la politique.
- α_ϕ : Taux d'apprentissage pour la mise à jour des paramètres de la fonction de valeur.

3.6.3 SAC (Soft Actor-Critic)

est un algorithme d'apprentissage par renforcement hors politique, basé sur le modèle acteur-critique. Il vise à maximiser à la fois la récompense attendue et l'entropie de la politique, favorisant ainsi l'exploration et la stabilité.

Algorithm 3: Soft Actor-Critic (SAC) Algorithm

```

1 Initialize ;;
2   Policy parameters  $\theta$ ;
3   Q-function parameters  $\phi_1, \phi_2$ ;
4   Target Q-function parameters  $\bar{\phi}_1, \bar{\phi}_2 \leftarrow \phi_1, \phi_2$ ;
5   Temperature  $\alpha$ , Replay buffer  $\mathcal{B}$ , Target entropy  $\bar{\mathcal{H}}$ ;
6 for  $episode = 1, 2, \dots, M$  do
7   for  $time\ step\ t = 1, 2, \dots, T$  do
8     Sample action  $a_t \sim \pi_\theta(\cdot|s_t)$ ;
9     Execute  $a_t$ , observe  $r_t, s_{t+1}$ ;
10    Store  $(s_t, a_t, r_t, s_{t+1})$  in  $\mathcal{B}$ ;
11    for  $gradient\ step\ in\ range(G)$  do
12      Sample minibatch  $\{s_i, a_i, r_i, s_{i+1}\} \sim \mathcal{B}$ ;
13      Sample  $\tilde{a}_{i+1} \sim \pi_\theta(\cdot|s_{i+1})$ ;
14      Compute target Q-value ;;
15       $y_i = r_i + \gamma (\min_{j=1,2} Q_{\bar{\phi}_j}(s_{i+1}, \tilde{a}_{i+1}) - \alpha \log \pi_\theta(\tilde{a}_{i+1}|s_{i+1}))$ ;
16      Update Q-functions ;;
17       $\phi_j \leftarrow \phi_j - \eta_Q \nabla_{\phi_j} \frac{1}{N} \sum (Q_{\phi_j}(s_i, a_i) - y_i)^2$  for  $j = 1, 2$ ;
18      Update policy ;;
19       $\theta \leftarrow \theta - \eta_\pi \nabla_\theta \frac{1}{N} \sum (\alpha \log \pi_\theta(a_i|s_i) - \min_{j=1,2} Q_{\phi_j}(s_i, a_i))$ ;
20      Update temperature ;;
21       $\alpha \leftarrow \alpha - \eta \nabla_\alpha \frac{1}{N} \sum (-\alpha (\log \pi_\theta(a_i|s_i) + \bar{\mathcal{H}}))$ ;
22      Update target networks ;;
23       $\bar{\phi}_j \leftarrow \tau \phi_j + (1 - \tau) \bar{\phi}_j$  for  $j = 1, 2$ ;
24    end
25  end
26 end

```

Signification des symboles de l'algorithme SAC

- θ : Paramètres du **réseau de politique** $\pi_\theta(a|s)$, qui modélise la distribution stochastique des actions à partir d'un état.

- ϕ_1, ϕ_2 : Paramètres des deux **fonctions Q (critiques)** Q_{ϕ_1} et Q_{ϕ_2} , utilisées pour estimer la valeur des paires état-action.
- $\bar{\phi}_1, \bar{\phi}_2$: Paramètres des **réseaux Q cibles**, initialisés avec ϕ_1 et ϕ_2 et mis à jour par "soft update" pour stabiliser l'apprentissage.
- α : **Température de l'entropie**, un coefficient adaptatif qui contrôle l'équilibre entre exploration et exploitation via le terme d'entropie.
- $\bar{\mathcal{H}}$: **Entropie cible** que la politique doit atteindre, généralement fixée à $-\dim(\mathcal{A})$, où \mathcal{A} est l'espace d'action.
- \mathcal{B} : **Buffer de rejouabilité** (Replay Buffer), utilisé pour stocker les transitions (s, a, r, s') afin de permettre un apprentissage hors-ligne.
- a_t : Action échantillonnée depuis la politique stochastique à l'état s_t : $a_t \sim \pi_\theta(\cdot | s_t)$.
- r_t : Récompense obtenue après l'exécution de l'action a_t .
- y_i : **Cible pour les fonctions Q**, calculée avec l'estimateur à double Q et un terme d'entropie.
- \tilde{a}_{i+1} : Action échantillonnée selon la politique courante au prochain état s_{i+1} .
- η_Q : Taux d'apprentissage pour la mise à jour des fonctions Q.
- η_π : Taux d'apprentissage pour la mise à jour de la politique.
- η : Taux d'apprentissage pour la mise à jour de la température α .
- τ : Taux de **mise à jour douce** des réseaux Q cibles.

3.7 Conclusion

Ce chapitre met en évidence le potentiel de l'apprentissage par renforcement, pour améliorer la collecte et la gestion des données agricoles dans des environnements dynamiques et contraignants. Ces stratégies, qui sont basées sur des idées comme l'interaction agent-environnement, la fonction de récompense et les politiques d'apprentissage, permettent aux drones de prendre des décisions intelligentes et autonomes. Les algorithmes étudiés, DDPG, PPO et SAC, fournissent des cadres efficaces pour gérer la complexité des

tâches de collecte tout en tenant compte des contraintes énergétiques. Ainsi, l'apprentissage par renforcement est une approche prometteuse pour relever les défis de l'agriculture intelligente en fournissant des solutions flexibles, optimisées et économes en énergie.

4.1 Introduction

Dans le cadre de l'essor des technologies d'agriculture intelligente, l'intégration de drones (UAVs) et de capteurs IoT constitue une solution prometteuse pour améliorer l'efficacité de la collecte et de la transmission des données sur le terrain. Ce chapitre est consacré à la simulation de ce système complexe et à l'analyse approfondie des performances des différentes stratégies de coordination mises en œuvre.

L'objectif est de modéliser un environnement réaliste où plusieurs UAVs interagissent avec un réseau de capteurs déployés sur une zone agricole, tout en respectant des contraintes opérationnelles telles que la limitation énergétique, la couverture spatiale, et la fiabilité des communications. À travers des scénarios variés, la simulation permet d'évaluer l'impact des décisions prises par les UAVs sur les performances globales du système.

Ce chapitre décrit dans un premier temps l'environnement de simulation développé, incluant la définition des composants, les bibliothèques utilisées, ainsi que les actions possibles pour les UAVs et les capteurs. Ensuite, les métriques de performance sont présentées avec leurs formules de calcul. Enfin, les algorithmes testés — allant d'approches simples à des méthodes avancées d'apprentissage par renforcement profond — sont détaillés, suivis d'une analyse comparative des résultats obtenus.

4.2 Configuration de l'Environnement de Simulation

4.2.1 Paramètres de l'Environnement

L'environnement de simulation agricole est configuré avec les paramètres suivants :

Paramètres spatiaux

Zone agricole est Espace carré contenant des capteurs répartis régulièrement pour une couverture maximale.

- Zone de couverture : 100 m × 100 m (10 000 m²)
- Portée des capteurs IoT : 10 m
- Nombre de capteurs : calculé automatiquement pour assurer une couverture complète

Paramètres des UAV

Agent mobile disposant d'une batterie limitée, d'un buffer de données et d'une vitesse maximale. Il se déplace, collecte des données, et les transfère.

- Nombre d'UAV : 3
- Capacité de batterie : 500 unités d'énergie
- Vitesse maximale : 5 – 15 m/s (aléatoire par UAV)
- Capacité de buffer de données : 100 MB
- Portée de communication : 20 m

Les UAV effectuent les tâches suivantes à chaque pas de simulation :

- **Déplacement** : Se déplacer dans le champ agricole selon une direction $[dx, dy]$, normalisée.
- **Collecte de données** : S'approcher d'un capteur à portée (≤ 10 m) pour recevoir ses données.
- **Offloading (déchargement)** : Transmettre les données stockées dans leur mémoire tampon (buffer) vers la station de base, si à portée (≤ 100 m).

Mise à jour à chaque étape :

- Temps courant
- Niveaux de batterie

- Données collectées et transférées
- Métriques de performance

- Formule de déplacement :

$$\text{position}_{t+1} = \text{position}_t + \vec{d} \times v_{\max} \times dt$$

- Formule d'énergie consommée :

$$E_{\text{move}} = 0.1 \times \text{distance parcourue}$$

Rôle :

- Surveille la zone agricole.
- Se dirige vers les capteurs ayant des données.
- Transporte les données vers la station de base.
- Gère sa batterie pour éviter l'épuisement total.
- Fait des compromis entre énergie, distance et priorité.

Paramètres des capteurs IoT

Capteur statique générant périodiquement des données environnementales. Il a une portée fixe et une priorité de données.

- Fréquence de collecte : toutes les 10 unités de temps
- Taille des données : 1 – 10 MB (aléatoire)
- Priorité des données : valeur aléatoire entre 0.1 et 1.0
- Pénalité d'attente : $0.1 \times \text{temps d'attente}$ (plafonnée à 2.0)

Les capteurs sont fixes dans le champ et disposent de :

- Une Position fixe $[x, y]$
- Une Batterie limitée
- Un Rayon de couverture (ex : 10 m)
- Une mémoire tampon contenant des données périodiquement générées

Rôle :

- Surveiller des paramètres environnementaux agricoles (ex : humidité, température)
- Générer des données toutes les 10 étapes
- Attendre qu'un UAV vienne les interroger pour décharger les données

Paramètres de la station de base

Située au centre du champ, elle reçoit les données des UAVs et les transmet au cloud. -
 Position : centre du terrain (50 m, 50 m) - Portée de communication : 100 m - Connexion
 cloud : toujours disponible

Rôle :

- Valider les transferts de données (offloads),
- Mesurer les performances du système (latence, énergie, succès),
- Stocker et agréger les données collectées,
- Servir de point d’ancrage pour l’interaction UAV–infrastructure.

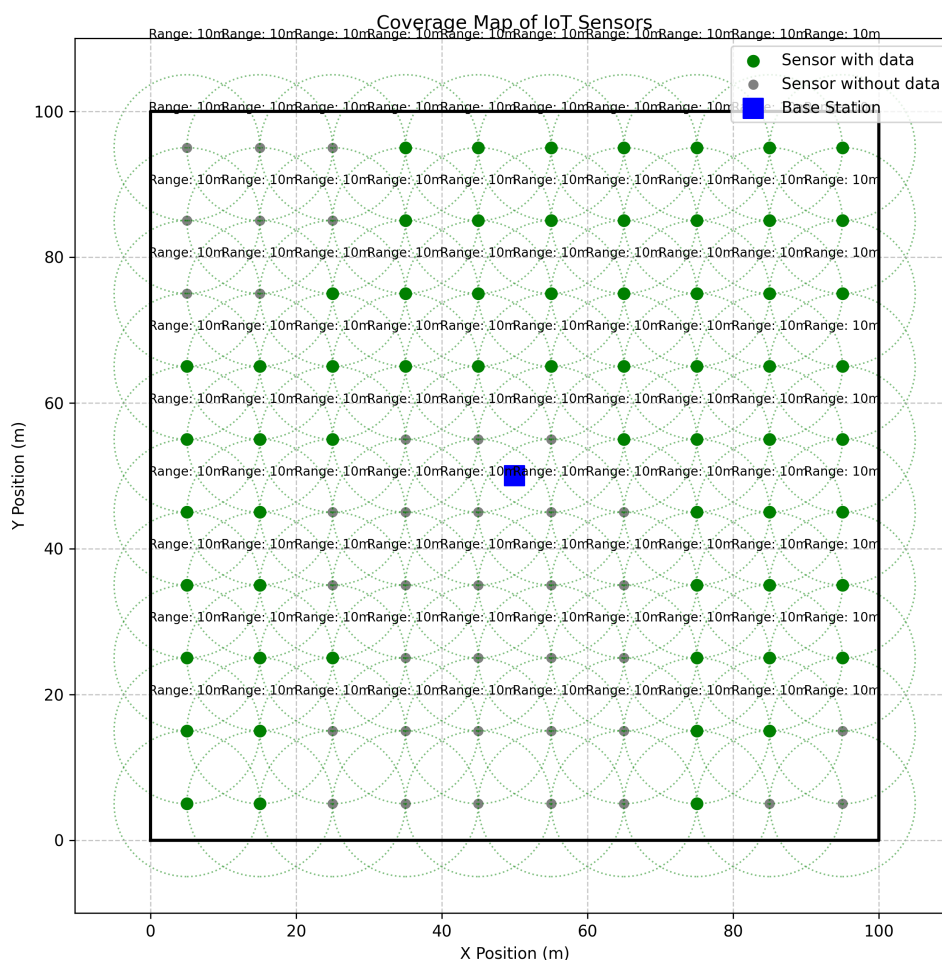


FIGURE 4.1 – Présentation de champ agricole .

4.2.2 Outils et plateformes

La simulation a été développée en Python (3.8.10) sur Ubuntu 20.04 LTS avec les caractéristiques suivantes :

— **Processeur** : Intel Core i5-6300U @ 2.40 GHz

— **Mémoire** : 8 Go de RAM

Les bibliothèques Python utilisées :

— pandas 2.1.1 : Analyse et traitement des résultats

— NumPy 1.26.0 : pour les vecteurs, distances, et calculs de couverture.

— Matplotlib 3.8.0 : visualisation de la couverture et des trajectoires.

— Torch 2.1.0 : Backend PyTorch pour les réseaux de neurones l'entraînement des algorithmes DRL (SAC, DDPG, PPO).

— gymnasium 0.29.1 : Structure d'environnement inspirée de Gym.

4.3 Algorithmes Comparés

Dans le but d'évaluer l'efficacité de la stratégie de déchargement de données assisté par drone basée sur l'apprentissage par renforcement profond (DRL), nous avons comparé plusieurs approches algorithmiques, classées en deux grandes catégories : heuristiques classiques et algorithmes d'apprentissage.

1. Algorithme Aléatoire (Random)

Cet algorithme sert de base de référence. Les drones se déplacent dans des directions aléatoires à chaque étape et collectent les données s'ils croisent un capteur dans leur rayon d'action. Il ne prend aucune décision basée sur l'état de l'environnement.

État (State) : - Position, état de la batterie et buffer de données de chaque UAV.

- Position de la station de base.

- Informations sur les capteurs à proximité.

Action : - Vecteur directionnel pour chaque UAV : $[dx, dy]$, normalisé.

- Dimension totale : Nombre d'UAVs \times 2.

2. Algorithme Glouton (Greedy)

Dans cette stratégie, chaque drone se dirige vers le capteur ayant la priorité de données la plus élevée et le plus proche. Une fois sa mémoire tampon pleine, il retourne vers la station pour le déchargement.

Hyperparamètres : Seuil de buffer pour retour base : 0.7

Seuil minimal du buffer : 0.5

Distance de collecte : 15.0

Portée station de base : 20.0

Fonction de score pondérée :

Poids distance : 0.4

Poids priorité : 0.3

Poids buffer : 0.3

État (State) : - Positions, état de la batterie et buffer de données des UAVs.

- Position de la station de base.

- Capteurs proches : position, priorité, statut des données.

Action : - Direction vers le capteur le plus pertinent (distance, priorité, remplissage du buffer), ou

- Vers la station de base si le buffer dépasse un seuil.

3. Collecte à Point Fixe (Static Sink)

Dans ce cas, les drones ne se déplacent pas vers les capteurs. Ce sont les capteurs qui envoient leurs données à une station fixe si celle-ci est dans le rayon de communication.

Hyperparamètres :

Taille de la grille : 6

Intervalle de mise à jour du chemin : 20 pas

Seuil de buffer pour retour base : 0.7

Seuil critique (urgence) : 0.9

Seuil minimal du buffer : 0.5

Distance seuil waypoint atteint : 8.0

Distance de collecte : 15.0

Portée station de base : 20.0

État (State) : - États des UAVs : position, buffer, batterie.

- Position de la station de base.
- Waypoints prédéfinis dans la grille.

Action : - Direction vers le waypoint courant (chemin prédéfini).

- Ou vers la station de base en cas de buffer plein.

Gestion des waypoints : - Chemins selon une spirale optimisée.

- Réassignation dynamique périodiquement.

4. DDPG (Deep Deterministic Policy Gradient)

L'algorithme DDPG repose sur un modèle d'apprentissage continu, où l'agent (le drone) apprend à prendre des décisions (direction de déplacement) en interagissant avec l'environnement.

Hyperparamètres : Couches cachées : 256 neurones chacune

Fonction d'activation : ReLU

Taux d'apprentissage (acteur) : 1e-4

Taux d'apprentissage (critique) : 1e-3

Facteur d'actualisation : 0.99

Taux de mise à jour des cibles : 0.005

Taille du batch : 128

Taille du buffer : 1e6

Bruit exploratoire initial : 0.05

Taux de décroissance du bruit : 0.99

Intervalle de mise à jour : 2

État (State) : - États complets des UAVs : position, batterie, buffer.

- Capteurs avec données disponibles.
- Station de base.

Action : - Vecteurs continus pour chaque UAV.

- Déterministes mais avec ajout de bruit Ornstein-Uhlenbeck pour exploration.

Apprentissage : - Basé sur un acteur (policy) et un critique (Q-value).

- Réseaux cibles et buffer de replay.
- Mise à jour par minimisation de la perte TD.

5. PPO (Proximal Policy Optimization)

PPO est un algorithme d'apprentissage par renforcement basé sur des politiques stochastiques. Il est réputé pour sa robustesse et sa stabilité, notamment dans les environnements bruités.

Hyperparamètres : Couches cachées : 256 neurones chacune

Taux d'apprentissage : $3e-4$ (souvent utilisé pour acteur et critique)

Facteur d'actualisation : 0.99

Ratio de clipping : 0.1

Nombre d'époques : 10

Taille du batch : 64

Taille maximale du buffer : 2048

État (State) : - États complets des UAVs : position, batterie, buffer.

- Capteurs avec données disponibles.
- Station de base.

Action : - Actions continues issues d'une distribution gaussienne paramétrée :

- Moyenne : sortie du réseau policy.
- Écart-type : appris ou fixe.

Métriques sauvegardées : - Log-probabilité des actions.

- Valeurs critiques (value network).
- Avantages calculés pendant l'entraînement (non implémentés dans le code simplifié).

Apprentissage : - PPO avec clipping de ratio :

$$L_{\text{CLIP}} = \mathbb{E} [\min(r_t(\theta)A_t, \text{clip}(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon)A_t)]$$

- Plusieurs époques d'entraînement par batch.

6. SAC (Soft Actor-Critic)

SAC est un algorithme DRL basé sur l'entropie maximale, favorisant une exploration plus efficace. Il est particulièrement utile dans des environnements complexes avec des actions continues.

Hyperparamètres :

Couches cachées : 256 neurones chacune

Fonction d'activation : ReLU

Taux d'apprentissage : 1e-3

Facteur d'actualisation : 0.99

Taux de mise à jour des cibles : 0.005

Température d'entropie : 0.1

Taille du batch : 182

Taille du buffer : 1e6

Intervalle de mise à jour : 2

État (State) : - États complets des UAVs : position, batterie, buffer.

- Capteurs avec données disponibles.
- Station de base.

Action : - Vecteurs de déplacement continus générés par un réseau stochastique.

- Ajout de bruit exploratoire.

Apprentissage : - Entropie maximisée pour favoriser l'exploration.

- Deux Q-networks et un réseau de politique.
- Buffer de replay + mise à jour douce des cibles (soft update).

4.4 Métriques d'Évaluation

Afin d'évaluer la performance du système proposé de déchargement de données IoT assisté par drones, plusieurs métriques quantitatives ont été utilisées. Ces métriques permettent d'évaluer la rapidité, l'efficacité énergétique et la fiabilité des transmissions au sein de l'environnement simulé.

4.4.1 Latence Moyenne (ms)

La latence mesure le délai total entre la collecte des données par le drone et leur déchargement à la station de base. Elle est définie par :

$$\text{Latence totale} = T_{\text{transmission}} + T_{\text{traitement}} + T_{\text{file}} \quad (4.1)$$

où :

$$T_{\text{transmission}} = \frac{D}{R + \varepsilon} \times 1000 \quad (4.2)$$

avec :

- D : taille des données (en Mo),
- R : débit de transmission (en Mo/s),
- ε : petit terme (ex. 10^{-6}) ajouté pour éviter la division par zéro.

4.4.2 Consommation Énergétique (Joules)

La consommation énergétique est composée de deux volets : l'énergie pour le déplacement du drone, et celle pour la transmission des données :

$$E_{\text{total}} = E_{\text{déplacement}} + E_{\text{transmission}} \quad (4.3)$$

où :

$$E_{\text{déplacement}} = 0.1 \times \|\Delta\vec{x}\| \quad (4.4)$$

$$E_{\text{transmission}} = \alpha \times d \times D \quad (4.5)$$

- $\|\Delta\vec{x}\|$: distance parcourue par le drone,
- d : distance entre le drone et le capteur ou la station,
- D : taille des données,
- α : coefficient empirique (ex. 0.01 pour collecte, 0.02 pour déchargement).

4.4.3 Taux de Succès (%)

Le taux de succès mesure la proportion d'opérations de déchargement effectuées avec succès :

$$\text{Taux de succès} = \frac{N_{\text{succès}}}{N_{\text{tentatives}}} \times 100 \quad (4.6)$$

où :

- $N_{\text{succès}}$: nombre de déchargements réussis,
- $N_{\text{tentatives}}$: nombre total de tentatives.

4.4.4 Débit Moyen (Throughput, en KB/s)

Le débit mesure la quantité de données transmises par unité de temps :

$$\text{Débit} = \frac{D_{\text{total}}}{T_{\text{latence total}} + \varepsilon} \quad (4.7)$$

où :

- D_{total} : données transmises (en KB),
- $T_{\text{latence total}}$: latence totale (en s),
- ε : terme pour éviter la division par zéro.

4.4.5 Efficacité Énergétique (Mo/J)

L'efficacité énergétique représente la quantité de données collectées pour chaque joule consommé :

$$\text{Efficacité énergétique} = \frac{D_{\text{collectées}}}{E_{\text{utilisée}} + \varepsilon} \quad (4.8)$$

4.4.6 Récompense Moyenne

La fonction de récompense \mathcal{R}_t est définie comme suit :

$$\mathcal{R}_t = \mathcal{R}_{\text{collecte}} + \mathcal{R}_{\text{offload}} - \mathcal{P}_{\text{batterie}} - \mathcal{P}_{\text{inactivité}} - \mathcal{P}_{\text{attente}} \quad (4.9)$$

où :

- $\mathcal{R}_{\text{collecte}}$: bonus basé sur la priorité, la proximité et la faible latence,
- $\mathcal{R}_{\text{offload}}$: forte récompense pour le transfert vers la station,
- $\mathcal{P}_{\text{batterie}}$: pénalité pour batterie faible ou vide,
- $\mathcal{P}_{\text{inactivité}}$: pénalité pour inactivité prolongée,
- $\mathcal{P}_{\text{attente}}$: pénalité pour capteurs restés trop longtemps sans service.

4.5 Résultats des Simulations

4.5.1 Consommation Énergétique (Joules)

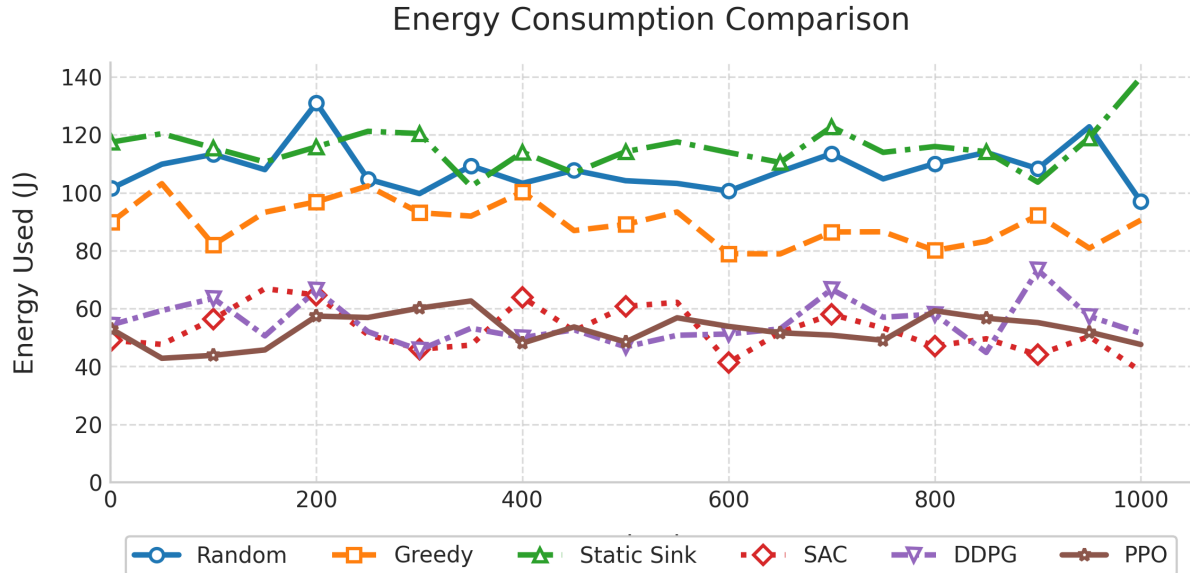


FIGURE 4.2 – Energy Consumption comparison.

Comme le montre la Figure 4.2, L’algorithme Aléatoire (Random) affiche une consommation relativement élevée, oscillant entre 100 et 130 J. Cette inefficacité énergétique s’explique par ses déplacements aléatoires non optimisés, qui génèrent des trajets souvent inutiles.

En revanche, l’algorithme Glouton (Greedy) présente une consommation plus modérée, généralement comprise entre 80 et 90 J. Grâce à sa stratégie d’optimisation locale, il parvient à réduire les déplacements superflus, ce qui améliore son efficacité énergétique.

L’algorithme Collecte à Point Fixe (Static Sink), quant à lui, consomme en moyenne entre 100 et 120 J. Bien qu’il soit plus efficace que le Random, il reste plus énergivore que le Greedy en raison de ses trajectoires fixes qui ne s’ajustent pas dynamiquement aux conditions environnementales.

Les algorithmes d’apprentissage par renforcement profond se révèlent nettement plus performants sur le plan énergétique. L’algorithme SAC (Soft Actor-Critic), par exemple, maintient sa consommation autour de 50 à 60 Joules, grâce à une exploration contrôlée et une capacité d’adaptation continue.

L’algorithme DDPG (Deep Deterministic Policy Gradient) démontre une efficacité encore meilleure, avec une consommation stable et faible, autour de 40 à 50 J. Sa politique déterministe contribue à minimiser les mouvements inutiles et à optimiser les trajectoires.

Enfin, l’algorithme PPO (Proximal Policy Optimization) se démarque par une performance énergétique remarquable, conservant une consommation très basse, également entre 40 et 50 J. Sa stabilité d’apprentissage et sa capacité à converger rapidement vers une stratégie optimale en font l’un des algorithmes les plus économes en énergie de cette étude.

En résumé, les algorithmes basés sur l’apprentissage par renforcement (SAC, DDPG, PPO) sont nettement plus économes en énergie que les algorithmes non apprenants (Random, Greedy, Static Sink). Parmi eux, PPO se révèle être le plus performant en termes d’efficacité énergétique.

4.5.2 Latence Moyenne (ms)

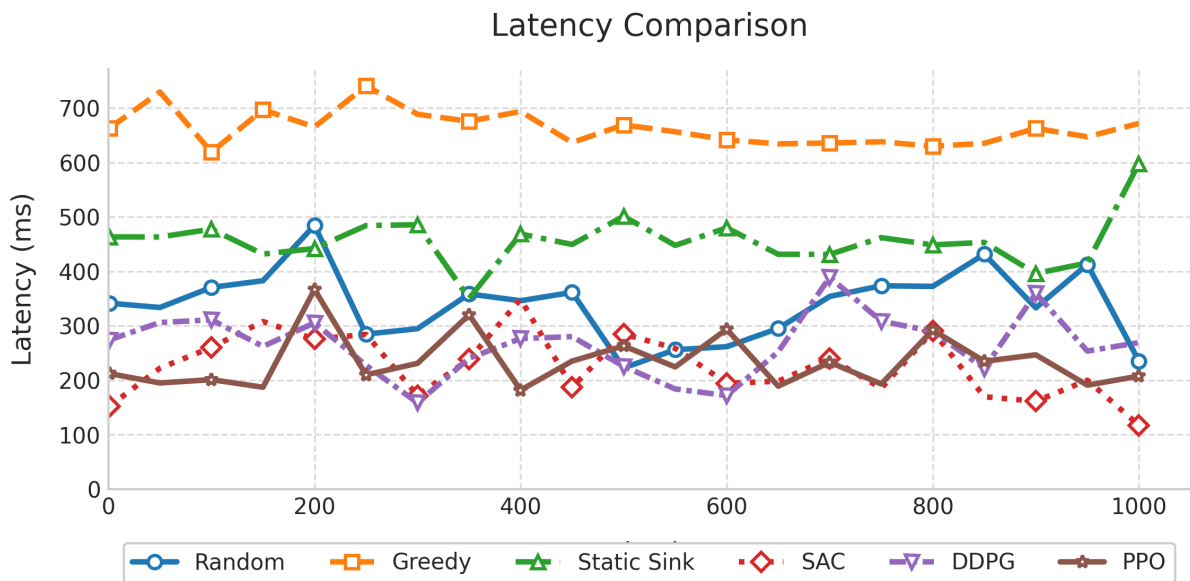


FIGURE 4.3 – Latency comparison.

Comme le montre la Figure 4.3, L’algorithme Aléatoire (Random) présente des latences relativement faibles, généralement comprises entre 200 et 350 ms. Toutefois, on observe une légère augmentation au fil du temps, probablement en raison de ses déplacements aléatoires qui peuvent allonger les trajets ou provoquer des collisions.

En comparaison, l'algorithme Glouton (Greedy) affiche des latences nettement plus élevées, stables entre 650 et 700 ms. Cette performance s'explique par son comportement myope, qui privilégie les décisions localement optimales au détriment de l'efficacité globale, ce qui engendre des retards dans la transmission des données.

L'algorithme Collecte à Point Fixe (Static Sink) se situe entre ces deux extrêmes, avec des latences oscillant entre 300 et 450 ms. Bien que plus élevées que celles du Random, ces valeurs restent modérées grâce aux trajectoires préétablies, qui assurent un certain degré de régularité.

Les algorithmes de type apprentissage par renforcement profond se distinguent à nouveau par leurs bonnes performances. L'algorithme SAC (Soft Actor-Critic) enregistre des latences variables allant de 200 à 300 ms. Ces variations sont probablement dues à son mécanisme d'exploration stochastique, mais globalement, il conserve des délais acceptables.

De leur côté, l'algorithme DDPG (Deep Deterministic Policy Gradient) et l'algorithme PPO (Proximal Policy Optimization) affichent tous deux des latences faibles et stables, oscillant entre 200 et 250 ms. Leurs approches respectives — déterministe pour DDPG et stable avec une bonne capacité de convergence pour PPO — leur permettent d'optimiser les trajets et de minimiser les retards dans la collecte et la transmission des données, garantissant ainsi une communication efficace dans l'environnement IoT agricole.

En résumé, les algorithmes DDPG et PPO offrent les meilleures performances en termes de latence, avec des valeurs stables et faibles. Le Greedy Algorithm, en revanche, souffre de latences élevées, ce qui pourrait affecter la qualité de service dans des scénarios sensibles au délai.

4.5.3 Récompense Moyenne

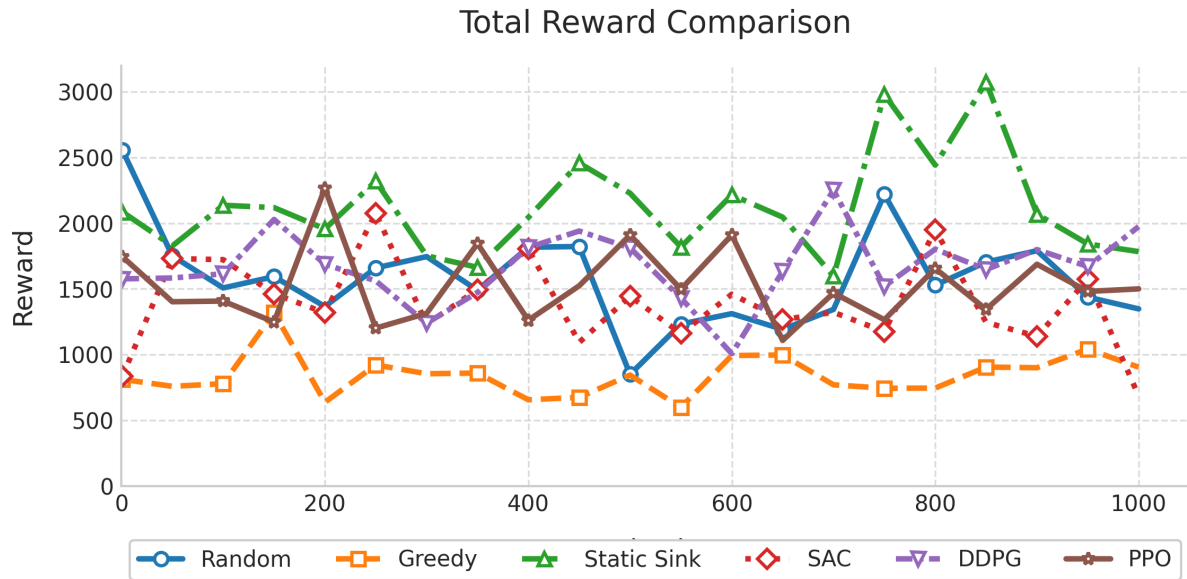


FIGURE 4.4 – Rewards comparison.

Comme le montre la Figure 4.4, L’algorithme Aléatoire (Random) montre une tendance à la baisse continue de ses récompenses, passant progressivement de 2500 à environ 1000. Cette diminution reflète son incapacité à interagir efficacement avec l’environnement, générant ainsi des pertes régulières au fil du temps.

L’algorithme Glouton (Greedy), quant à lui, maintient des récompenses plus stables, comprises entre 700 et 800. Bien qu’il surpasse le Random Algorithm, ses résultats restent limités en raison de son approche myope, qui privilégie les gains immédiats au détriment de l’optimisation globale.

L’algorithme Collecte à Point Fixe (Static Sink) présente une forte variabilité dans ses récompenses. Il atteint parfois des pics élevés allant de 2000 à 2500, mais ces gains sont suivis de chutes soudaines. Cette instabilité est probablement due à ses trajectoires rigides, peu adaptées aux changements dynamiques de l’environnement.

Les algorithmes basés sur l’apprentissage par renforcement profond montrent des résultats plus prometteurs. L’algorithme SAC (Soft Actor-Critic), par exemple, génère des récompenses fluctuantes situées entre 1500 et 2000. Malgré une certaine instabilité, ses performances s’améliorent progressivement grâce à son apprentissage continu et à sa stratégie d’exploration efficace.

L'algorithme DDPG (Deep Deterministic Policy Gradient) se distingue par des récompenses stables et relativement élevées, comprises entre 1500 et 2000. Sa capacité à apprendre des politiques déterministes efficaces lui permet d'assurer des performances constantes.

Enfin, l'algorithme PPO (Proximal Policy Optimization) enregistre les meilleures récompenses cumulées, oscillant entre 1800 et 2500. Grâce à sa convergence rapide vers des stratégies optimales et à son équilibre entre exploitation et exploration, il parvient à obtenir des résultats supérieurs aux autres algorithmes évalués dans ce contexte.

PPO se distingue clairement comme le meilleur algorithme en termes de récompense totale, suivi de près par DDPG. Les algorithmes non apprenants (Random, Greedy, Static Sink) affichent des performances inférieures, avec des récompenses instables ou décroissantes.

4.5.4 Taux de Succès (%)

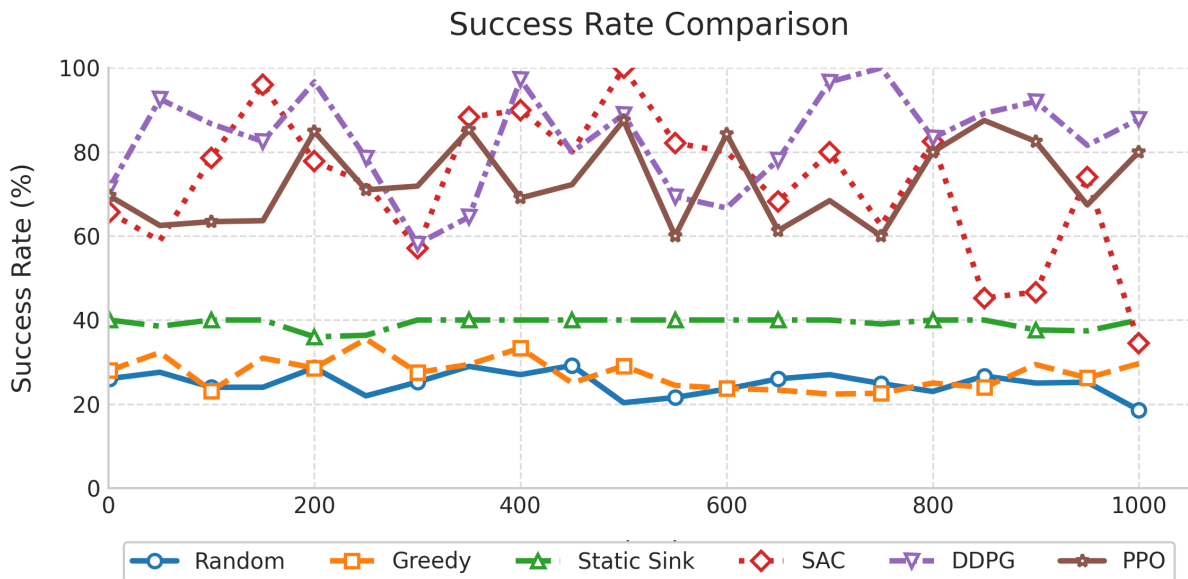


FIGURE 4.5 – Success Rate comparison.

Comme le montre la Figure 4.5, L'algorithme Aléatoire (Random) enregistre les performances les plus faibles, avec un taux de succès oscillant entre 20 et 30%. Ce résultat s'explique par l'absence totale de stratégie structurée, ce qui conduit à des actions inefficaces et à de nombreuses erreurs.

L'algorithme Glouton (Greedy), bien qu'un peu plus performant, reste limité avec un taux de succès moyen variant entre 30 et 40 %. Son approche orientée vers des décisions immédiates ne permet pas d'optimiser les résultats à long terme.

L'algorithme Collecte à Point Fixe (Static Sink) affiche des taux de succès similaires, entre 30 et 40 %, en raison de ses trajectoires figées qui ne tiennent pas compte des changements dynamiques dans l'environnement, entraînant ainsi des échecs fréquents.

En revanche, les algorithmes de type apprentissage profond offrent des résultats bien supérieurs. L'algorithme SAC (Soft Actor-Critic) présente des taux de succès plus élevés, variant entre 60 et 80 %, bien que sujets à des fluctuations. Ces performances en amélioration continue sont le reflet de sa capacité d'adaptation par apprentissage progressif.

L'algorithme DDPG (Deep Deterministic Policy Gradient) se distingue par des taux de succès à la fois élevés et stables, compris entre 80 et 90 %. Grâce à sa politique déterministe, il parvient à converger efficacement vers des stratégies optimales.

Enfin, l'algorithme PPO (Proximal Policy Optimization) atteint les meilleurs résultats en matière de taux de succès, avec des valeurs oscillant entre 80 et 95 %. Son équilibre entre exploration contrôlée et exploitation optimisée lui permet d'offrir des performances supérieures, stables et fiables dans divers contextes environnementaux.

PPO et DDPG dominent clairement en termes de taux de succès, avec des performances stables et élevées. Les algorithmes non apprenants (Random, Greedy, Static Sink) affichent des taux de succès inférieurs, souvent instables ou décroissants.

4.5.5 Débit Moyen (Throughput, en KB/s)

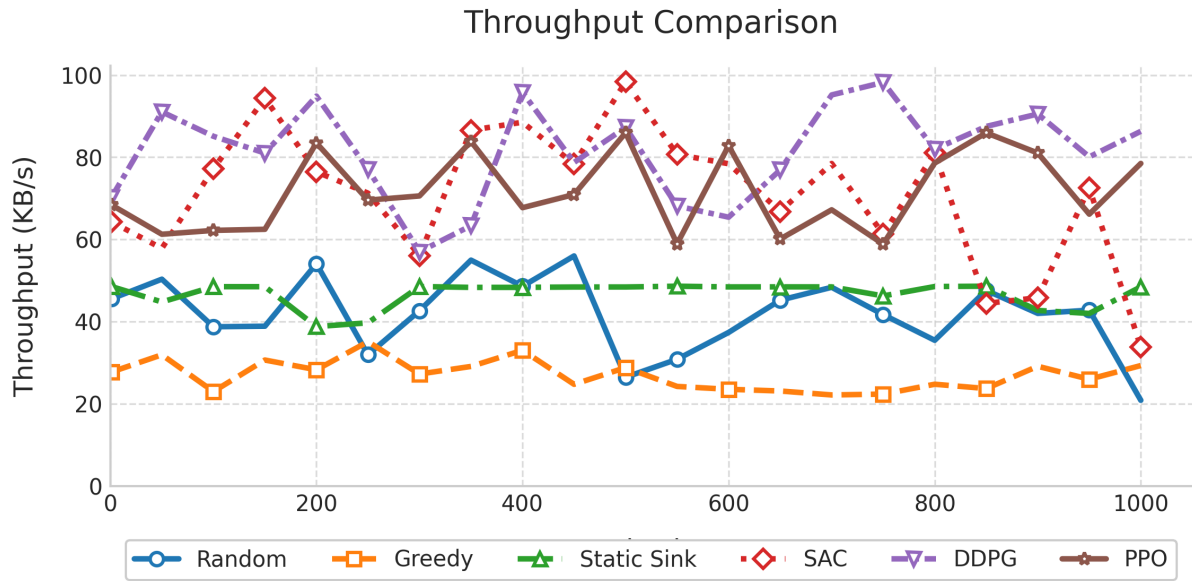


FIGURE 4.6 – Throughput comparison.

Comme le montre la Figure 4.6, L’algorithme Aléatoire (Random) affiche les plus faibles performances, avec des débits instables variant entre 20 et 40 KB/s. Cette faiblesse résulte directement de l’absence de stratégie structurée et de la nature aléatoire des déplacements, ce qui engendre fréquemment des pertes de données.

L’algorithme Glouton (Greedy), quant à lui, atteint des débits légèrement supérieurs, oscillant entre 20 et 30 KB/s. Toutefois, son approche myope limite son efficacité globale dans la transmission régulière des données.

En comparaison, l’algorithme Collecte à Point Fixe (Static Sink) présente des débits plus stables et relativement élevés, situés entre 40 et 50 KB/s. Bien que ses trajectoires prédéfinies assurent une certaine régularité, leur rigidité face aux variations de l’environnement empêche une adaptation optimale.

Les algorithmes d’apprentissage profond montrent, comme dans les autres métriques, de meilleures performances. L’algorithme SAC (Soft Actor-Critic) offre des débits comparables à ceux du Static Sink (40 à 50 KB/s), avec toutefois une plus grande variabilité due à son exploration stochastique. Néanmoins, son apprentissage progressif permet une amélioration continue.

Les meilleurs résultats sont obtenus par l’algorithme DDPG (Deep Deterministic Po-

licy Gradient) et l'algorithme PPO (Proximal Policy Optimization), tous deux affichant des débits élevés et constants, allant de 60 à 80 KB/s. Le caractère déterministe de DDPG et la capacité de PPO à équilibrer exploration et exploitation expliquent leur aptitude à atteindre une transmission de données efficace et régulière. Ces performances les positionnent comme les méthodes les plus adaptées pour maximiser le débit dans un environnement IoT agricole dynamique.

DDPG et PPO se distinguent par leurs débits élevés et stables, reflétant leur capacité à optimiser efficacement les communications. Les algorithmes non apprenants (Random, Greedy, Static Sink) affichent des débits inférieurs, souvent instables ou décroissants.

4.6 Tableau de Comparaison des Algorithmes

Algorithme	Latence (ms)	Énergie (J)	Succès (%)	Débit (KB/s)	Récomp.
Random	338.35	108.28	24.97	41.93	1584.198
Greedy	663.29	89.51	27.31	27.06	843.879
Static Sink	456.15	115.71	39.29	46.75	2116.925
SAC	226.09	52.45	72.40	71.08	1390.210
DDPG	265.04	55.14	82.95	81.51	1669.889
PPO	233.70	52.62	72.96	71.65	1525.887

TABLE 4.1 – Comparaison des performances moyennes des algorithmes

L'analyse comparative des algorithmes révèle des compromis significatifs entre différentes métriques de performance. SAC (Soft Actor-Critic) se distingue comme l'algorithme le plus efficace pour minimiser la latence (226,09 ms) et la consommation énergétique (52,45 J), établissant un équilibre remarquable entre réactivité et efficacité énergétique.

Cependant, DDPG (Deep Deterministic Policy Gradient) domine en termes de qualité de service avec le taux de succès le plus élevé (82,95 %) et le meilleur débit (81,51 KB/s), ce qui en fait la solution optimale pour les applications sensibles à la fiabilité des transmissions.

Fait intéressant, l'algorithme Collecte à Point Fixe (Static Sink) présente la récompense globale la plus élevée (2116,93), surpassant même les approches d'apprentissage par

renforcement. Cela suggère que sa simplicité opérationnelle peut compenser ses performances inférieures sur d'autres métriques.

Les algorithmes traditionnels (Aléatoire (Random) et Glouton (Greedy)) démontrent des limitations significatives, notamment en latence (663,29 ms pour Greedy) et en taux de succès (24,97 % pour Random), ce qui limite leur pertinence dans des environnements dynamiques et critiques.

L'algorithme PPO (Proximal Policy Optimization) montre des performances équilibrées, se positionnant comme un compromis viable entre SAC et DDPG.

Cette analyse souligne qu'aucun algorithme n'excelle simultanément sur toutes les métriques, et que le choix optimal dépend des priorités spécifiques de l'application :

- **Efficacité énergétique** : SAC
- **Fiabilité des communications** : DDPG
- **Maximisation de la récompense** : Static Sink

4.7 Discussion

L'analyse approfondie des différentes métriques (énergie consommée, latence, récompense, taux de succès et débit) permet d'établir une hiérarchie claire entre les algorithmes étudiés.

Les algorithmes classiques *Aléatoire (Random)*, *Glouton (Greedy)* et *Collecte à Point Fixe (Static Sink)* affichent des performances globalement inférieures. L'algorithme Random, en raison de son absence de stratégie, entraîne une consommation d'énergie élevée, une latence instable et des récompenses faibles. L'algorithme Greedy, bien qu'un peu plus structuré, reste limité par son approche localement optimisée, ce qui se traduit par des performances moyennes sur l'ensemble des métriques. Static Sink offre une meilleure stabilité énergétique et un débit modéré grâce à ses trajectoires prédéfinies, mais souffre d'un manque d'adaptabilité dans un environnement changeant, ce qui affecte son taux de succès.

En revanche, les algorithmes d'apprentissage par renforcement profond (DRL) SAC, DDPG et PPO surpassent nettement les méthodes traditionnelles sur toutes les métriques clés.

- **DDPG (Deep Deterministic Policy Gradient)** se distingue par un excellent compromis entre faible latence (200–250 ms), consommation énergétique réduite (40–50 J), débits élevés (60–80 KB/s) et un taux de succès stable (80–90%). Sa capacité à converger vers des politiques déterministes optimales garantit des performances constantes et fiables.
- **SAC (Soft Actor-Critic)** brille par son efficacité énergétique (50–60 J) et son apprentissage progressif, malgré des performances légèrement moins stables que DDPG en termes de récompense et de taux de succès (60–80%). Il constitue néanmoins un bon compromis pour des scénarios dynamiques grâce à son exploration stochastique.
- **PPO (Proximal Policy Optimization)** affiche les meilleurs taux de succès (jusqu'à 95%) et un débit comparable à DDPG. Sa stabilité d'apprentissage et sa capacité à atteindre rapidement des stratégies optimales en font l'un des meilleurs candidats pour des environnements complexes, bien qu'il reste légèrement en retrait de DDPG concernant la régularité des performances globales.

En résumé, ces résultats démontrent clairement que le recours à des techniques d'intelligence artificielle avancées permet d'améliorer significativement la coordination et l'efficacité des UAVs dans les réseaux de capteurs sans fil, en particulier dans des contextes agricoles intelligents. L'intégration de méthodes d'apprentissage profond permet non seulement de réduire la consommation énergétique et la latence, mais aussi d'accroître le taux de succès et le débit des données transmises, des éléments essentiels pour assurer une gestion intelligente et autonome des ressources agricoles.

4.8 Conclusion

Ce chapitre a présenté une analyse approfondie des performances de différentes stratégies de gestion du trafic de données agricoles par drones, en comparant des approches classiques (aléatoire, gloutonne) avec des algorithmes avancés d'apprentissage par renforcement (DDPG, PPO, SAC). Les résultats confirment la pertinence de notre approche RL pour résoudre les défis identifiés dans les chapitres précédents. La robustesse des algorithmes testés permet d'envisager leur application pratique dans des fermes intelligentes.

CONCLUSION GÉNÉRALE

Dans le cadre de ce mémoire, nous avons étudié les stratégies de déchargement de données dans un environnement intelligent, en mettant l'accent sur l'intégration de techniques d'apprentissage par renforcement profond (Deep Reinforcement Learning) afin d'optimiser les performances des systèmes IoT assistés par des véhicules aériens autonomes (UAVs). L'objectif principal était de minimiser les délais de traitement tout en prenant en compte les contraintes énergétiques et la dynamique de mobilité des drones.

Les simulations menées ont permis de comparer plusieurs approches, incluant des méthodes classiques (aléatoire, glouton) et des algorithmes avancés de RL tels que DDPG, PPO et SAC. Les résultats obtenus démontrent clairement la supériorité des algorithmes d'apprentissage par renforcement, notamment DDPG et PPO, qui surpassent largement les méthodes traditionnelles en termes de consommation énergétique, de latence et de taux de réussite des transmissions. Ces algorithmes permettent aux UAVs d'apprendre des politiques optimales pour la collecte et la transmission des données, tout en s'adaptant dynamiquement aux conditions changeantes de l'environnement.

Les principales contributions de ce travail peuvent être résumées comme suit :

Optimisation énergétique : Une réduction notable de la consommation d'énergie, grâce à la planification intelligente des trajectoires et à la gestion efficace des ressources embarquées.

Amélioration des performances : Des gains significatifs en matière de latence, de débit et de taux de succès, rendant le système particulièrement adapté aux exigences des applications agricoles en temps réel.

Capacité d'adaptation : Une meilleure résilience aux variations du terrain et aux besoins fluctuants en données, garantissant ainsi une couverture efficace et continue.

En conclusion, cette étude valide l'efficacité des techniques de RL pour une gestion intelligente et autonome des données agricoles via drones. Les perspectives futures incluent des expérimentations sur le terrain, l'intégration de sources de données hétérogènes (capteurs multispectraux, imagerie thermique, etc.), ainsi que l'amélioration de l'autonomie énergétique des UAVs. De telles avancées pourraient contribuer à transformer durablement les pratiques agricoles, en les rendant plus précises, efficaces et respectueuses de l'environnement.

- [1] B. R. Limited. Towards smart farming : Agriculture embracing the iot vision. Technical report, Reasearch report, UK, London, USA, Boston, 2015.
- [2] Madakam, S. ; Ramaswamy, R. ; Tripathi, S. Internet of Things (IoT) : A Literature Review. *J. Comput. Commun.* 2015, 3, 164–173. [Google Scholar] [CrossRef] [Green Version]
- [3] Zhao, J.C. ; Zhang, J.F. ; Feng, Y. ; Guo, J.X. The study and application of the IOT technology in agriculture. In *Proceedings of the IEEE 2010 3rd International Conference on Computer Science and Information Technology*, Chengdu, China, 9–11 July 2010 ; Volume 2, pp. 462–465.
- [4] Verdouw, C. ; Sundmaeker, H. ; Tekinerdogan, B. ; Conzon, D. ; Montanaro, T. Architecture framework of IoT-based food and farm systems : A multiple case study. *Comput. Electron. Agric.* 2019, 165, 104939.
- [5] Benke, K. ; Tomkins, B. Future food-production systems : Vertical farming and controlled-environment agriculture. *Sustain. Sci. Pract. Policy* 2017, 13, 13–26. [CrossRef].
- [6] Kim, J., Kim, S., Ju, C., Son, H.I., 2019. Unmanned aerial vehicles in agriculture : A review of perspective of platform, control, and applications. *IEEE Access* 7, 105100–105115.

-
- [7] Abdul Hakkim, V.M. ; Abhilash Joseph, E. ; Ajay Gokul, A.J. ; Mufeedha, K. Precision Farming : The Future of Indian Agriculture. *J.Appl. Biol. Biotechnol.* 2016, 4, 068–072. [CrossRef]
- [8] Dhanaraju, M., Chenniappan, P., Ramalingam, K., Pazhanivelan, S., and Kaliaperumal, R. (2022). Smart Farming : Internet of Things (IoT)-Based Sustainable Agriculture. *Agriculture*, 12(10), 1745.
- [9] "The UAV technology in smart agriculture provides information on fertilization, irrigation, use of pesticides, plant growth monitoring, weed management, crop disease management... UAVs equipped with multispectral cameras predict yield and detect stress with high spatial accuracy."
- [10] Mobility-Aware Task Offloading and Resource Allocation in UAV-Assisted Vehicular Edge Computing Networks. by Long Chen ,Jiaqi Du, and Xia Zhu
- [11] Dhanaraju et al. (2022). Smart Farming : Internet of Things (IoT)-Based Sustainable Agriculture. *Agriculture*, 12(10), 1745.
- [12] Lang, L. GPS + GIS + remote sensing : An overview. *Earth Obs. Mag.* 1992, 1, 23–26.
- [13] Batte, M.T. ; VanBuren, F.N. Precision farming—Factor influencing productivity. In *Proceedings of the Northern Ohio Crops Day Meeting, Wood County, OH, USA, 21 January 1999.*
- [14] Srisruthi, S. ; Swarna, N. ; Ros, G.M.S. ; Elizabeth, E. Sustainable agriculture using eco-friendly and energy efficient sensor technology. In *Proceedings of the 2016 IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT), Bangalore, India, 20–21 May 2016 ; IEEE : Bangalore, India, 2016 ; pp. 1442–1446.* [CrossRef]
- [15] Mallarino, A.P. ; Wittry, D.J. Use of DGPS, yield monitors, soil testing and variable rate technology to improve phosphorus and potassium management. In *Proceedings of the Integrated Crop Management Conference ; Iowa State University Extension and Outreach : Ames, IA, USA, 1997 ; pp. 267–275.*

-
- [16] Elijah, O. ; Rahman, T.A. ; Orikumhi, I. ; Leow, C.Y. ; Hindia, M.N. An Overview of Internet of Things (IoT) and Data Analytics in Agriculture : Benefits and Challenges. *IEEE Internet Things J.* 2018, 5, 3758–3773. [CrossRef]
- [17] Muthumanickam, D. ; Kannan, P. ; Kumaraperumal, R. ; Natarajan, S. ; Sivasamy, R. ; Poongodi, C. Drought assessment and monitoring through remote sensing and GIS in western tracts of Tamil Nadu, India. *Int. J. Remote Sens.* 2011, 32, 5157–5176.
- [18] Sudduth, K.A. ; Hummel, J.W. Soil Organic Matter, CEC, and Moisture Sensing with a Portable NIR Spectrophotometer. *Trans. ASAE* 1993, 36, 1571–1582.
- [19] Sushil, S. ; Radha Mohan, S. ; Manhas, S.S. ; Shiv Kumar, L. Potential of Variable Rate Application Technology in India. *AMA Agric.Mech. Asia Afr. Lat. Am.* 2014, 45, 74–89.
- [20] Hassan, A. ; Aitazaz, A.F. ; Farhat, A. ; Bishnu, A. ; Travis, E. Precision Irrigation Strategies for Sustainable Water Budgeting of Potato Crop in Prince Edward Island. *Sustainability* 2020, 12, 2419. [CrossRef]
- [21] Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, 143, 23–37.
- [22] Bao, T., Syed, A., Kennedy, W. S., & Erol-Kantarci, M. (2024). Sustainable Task Offloading in Secure UAV-assisted Smart Farm Networks : A Multi-Agent DRL with Action Mask Approach. <https://doi.org/10.48550/arxiv.2407.19657>
- [23] Zhang, J., Li, X., Zhang, X., Xue, Y., Srivastava, G., & Dou, W. (2021). Service offloading oriented edge server placement in smart farming. *Software - Practice and Experience*, 51(12), 2540–2557. <https://doi.org/10.1002/SPE.2847>
- [24] Sharma, P., Shukla, S., & Vasudeva, A. (2021). Trust-Based Opportunistic Network Offloaders for Smart Agriculture. *International Journal of Agricultural and Environmental Information Systems*, 12(1), 37–54. <https://doi.org/10.4018/IJAEIS.20210101.OA3>

-
- [25] Deep Reinforcement Learning for Task Offloading in UAV-Aided Smart Farm Networks. (2022). <https://doi.org/10.1109/fnwf55208.2022.00054>
- [26] Gao, J., Wu, R., Hao, J., Xu, C., & Guo, H. (2023). Energy-Efficient Resource Scheduling and Computation Offloading Strategy for Solar-Powered Agriculture WSN. *Journal of Sensors*, 2023, 1–17. <https://doi.org/10.1155/2023/7020104>
- [27] Ouafiq, E. M., Saadane, R., Chehri, A., & Wahbi, M. (2022). Data Lake Conception for Smart Farming : A Data Migration Strategy for Big Data Analytics (pp. 191–201). <https://doi.org/10.1007/978-981-19-3455-1-15>
- [28] El Mehdi, M., El Aissi, E., Benjelloun, S., Lakhrissi, Y., El, S., Ben Ali, H. (2023). A Scalable Smart Farming Big Data Platform for Real-Time and Batch Processing Based on Lambda Architecture. *Journal of System and Management Sciences*. <https://doi.org/10.33168/jsms.2023.0202>
- [29] Sathya, D., Thangamani, R., & Balaji, B. (2024). The Revolution of Edge Computing in Smart Farming (pp. 351–389). Springer Vienna. <https://doi.org/10.1007/978-3-031-51195-0-17>
- [30] Sarker, V. K., Alam, M. G. R., Kwak, K. S. (2020). Deep Reinforcement Learning-Based Offloading in Edge Computing for Internet of Things. *IEEE Access*, 8,148588–148601. [DOI : 10.1109/ACCESS.2020.3016094].
- [31] What is a neural network ? [online] available : <https://aws.amazon.com/fr/what-is/neural-network/>.
- [32] Reinforcement Learning : An Introduction, Richard S. Sutton and Andrew G. Barto,2014, The MIT Press Cambridge, Massachusetts.
- [33] GeeksforGeeks – Model-Free Reinforcement Learning - An Overview <https://www.geeksforgeeks.org/model-free-reinforcement-learning-an-overview/>
- [34] GeeksforGeeks – Model-Based Reinforcement Learning (MBRL) in AI <https://www.geeksforgeeks.org/model-based-reinforcement-learning-mbri-in-ai/>.

- [35] Sharma, A., Goyal, M., Singh, P. (2021). Reinforcement Learning Techniques for Smart Agriculture : A Review. International Journal of Information Technology, 13,249–256. [DOI : 10.1007/s41870-020-00563-4].
- [36] Smart Farming, The future of Agriculture. Lawrence Price and Co. Jun 8, 2021
- [37] Geekflare (2024), “Introduction aux réseaux neuronaux [7 Ressources didactiques]”, article de blog par Amrita Pathak [https ://geekflare.com/fr/neural-networks/](https://geekflare.com/fr/neural-networks/)