

Inria



ai
aivancity

SCHOOL FOR

TECHNOLOGY, BUSINESS & SOCIETY

Optimiser la sélection de tâches sous contrainte énergétique dans des systèmes IoT grâce à l'apprentissage par renforcement

Hugo Hadjur (aivancity Paris-Cachan, ENS de Lyon)

Doreid Ammar (aivancity Paris-Cachan)

Laurent Lefèvre (Inria, Univ Lyon, ENS de Lyon, UCBL, CNRS, LIP)

Mars 2023

Plan

1. Éléments de contexte – Problématique
2. L'apprentissage par renforcement
3. Modélisation de l'environnement
4. Résultats – Proximal Policy Optimization
5. Perspectives

Les activités de recherche passées et en cours

Conception et modélisation de systèmes connectés durables, autonomes, intelligents et à basse consommation énergétique appliqués à l'apiculture de précision.



Revue bibliographique des différents projets de recherche liés aux ruches connectées

Publication : Hugo Hadjur, Doreid Ammar, Laurent Lefèvre, **Toward an intelligent and efficient beehive: A survey of precision beekeeping systems and services**, Computers and Electronics in Agriculture, Volume 192, 2022, 106604, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2021.106604>



Design d'un système de ruche connectée innovant



Analyses énergétiques pour l'optimisation de ce système et des services qui lui sont intégrés. Publications :

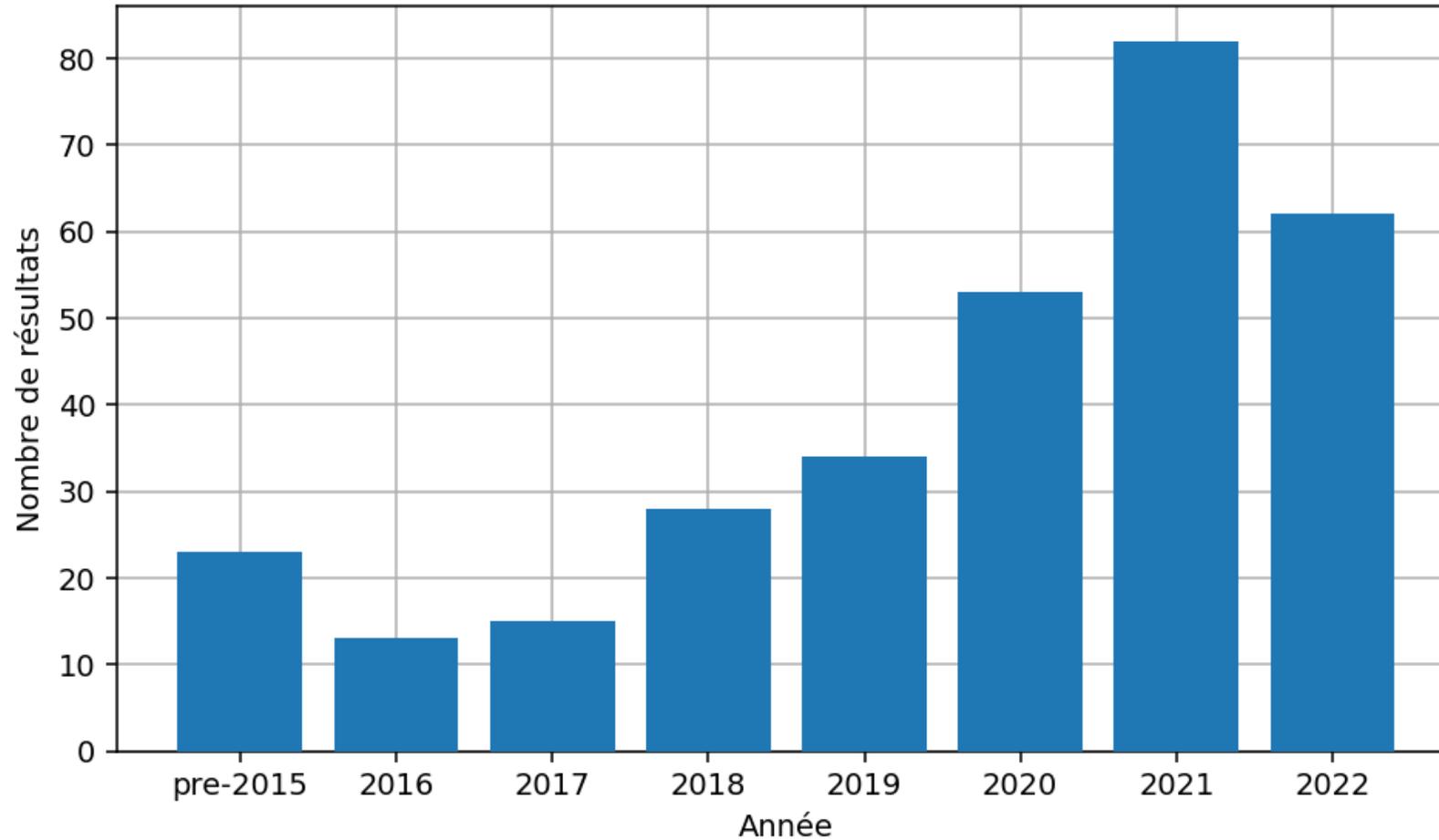
- Hugo Hadjur, Doreid Ammar, and Laurent Lefèvre. 2020. "Analysis of energy consumption in a precision beekeeping system." In *Proceedings of the 10th International Conference on the Internet of Things (IoT '20)*. Association for Computing Machinery, New York, NY, USA, Article 20, 1–8. <https://dl.acm.org/doi/10.1145/3410992.3411010>
- Hugo Hadjur, Doreid Ammar, Laurent Lefèvre, **Services Orchestration at the Edge and in the Cloud for Energy-Aware Precision Beekeeping Systems**, PAISE 2023 (IPDPS Workshop), May 2023



Optimiser la sélection de tâches sous contrainte énergétique dans des systèmes IoT grâce à l'apprentissage par renforcement

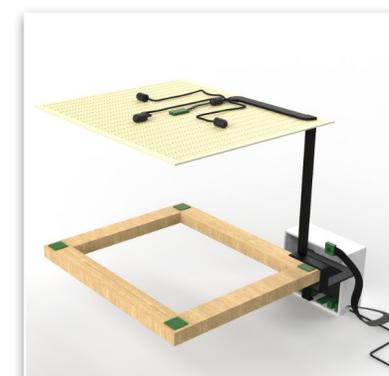
L'apiculture de précision : un thème en croissance jusqu'à 2021

Nombre de résultats Google Scholar pour le terme "Precision Beekeeping" au cours du temps



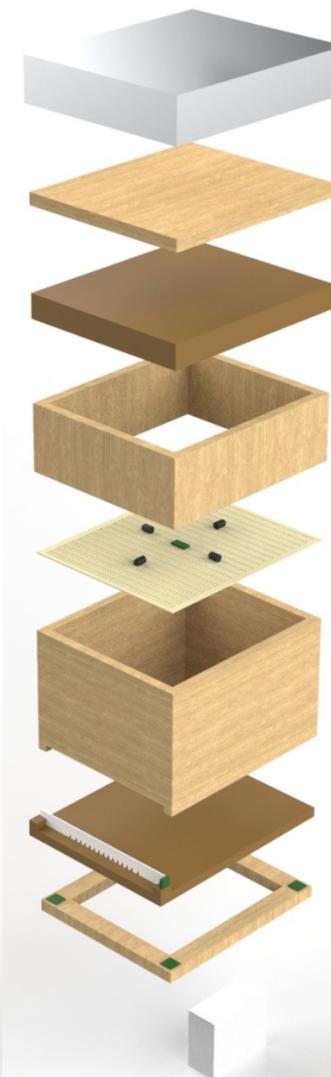
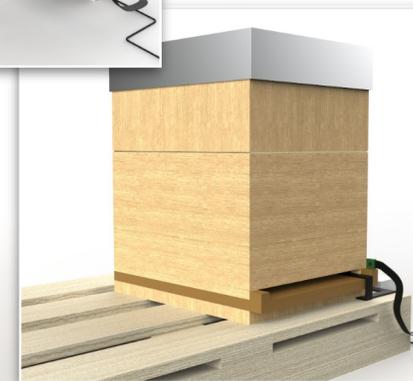
1. Éléments de contexte (1/2)

- Motivation générale : préserver les abeilles et épauler les apiculteurs.
- Phase de recherche expérimentale : déploiement de ruches urbaines connectées en 2022 (à Lyon et à Paris).



Entrées des câbles dans la ruche par la planche d'envol et la porte d'entrée

Vue d'ensemble des composants électroniques et capteurs



1. Éléments de contexte (2/2)

- Déploiement = Phase de collecte de **données apicoles** et de **consommation énergétique**.

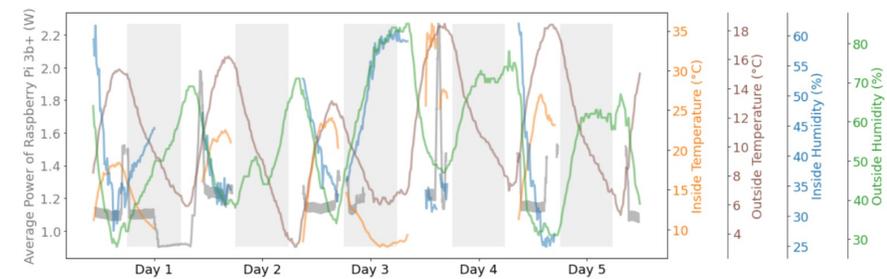
Le matériel informatique de base : un Raspberry Pi qui s'allume périodiquement pour collecter, traiter et transmettre des données.

- En complément de la collecte des données, on aimerait intégrer des méthodes capables de transformer ces données en diagnostic.

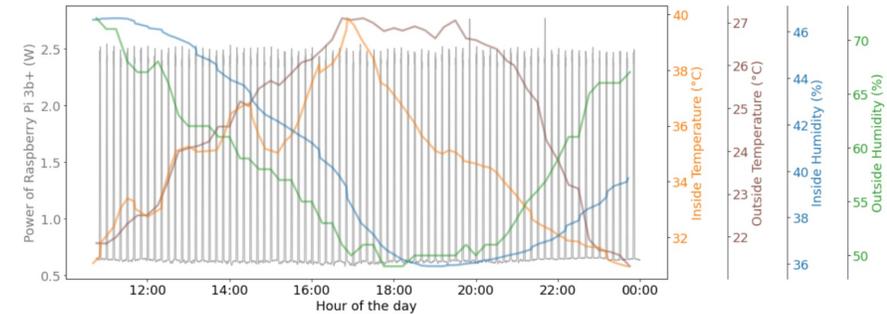
→ Dans la littérature, il y a au moins 8 grandes catégories de diagnostic.

→ Peut-on tout faire en continu ?

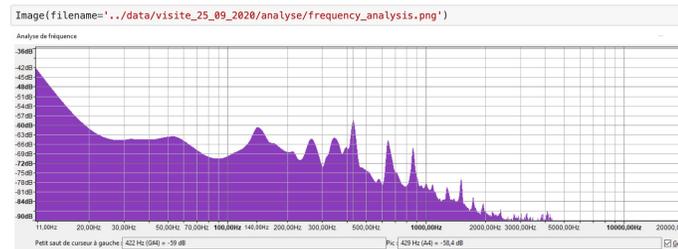
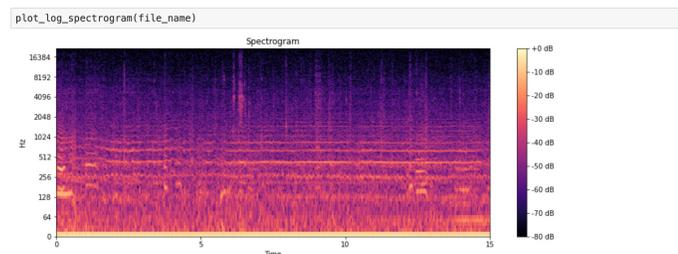
Doit-on faire des pauses et préserver l'énergie ?



(a)



(b)



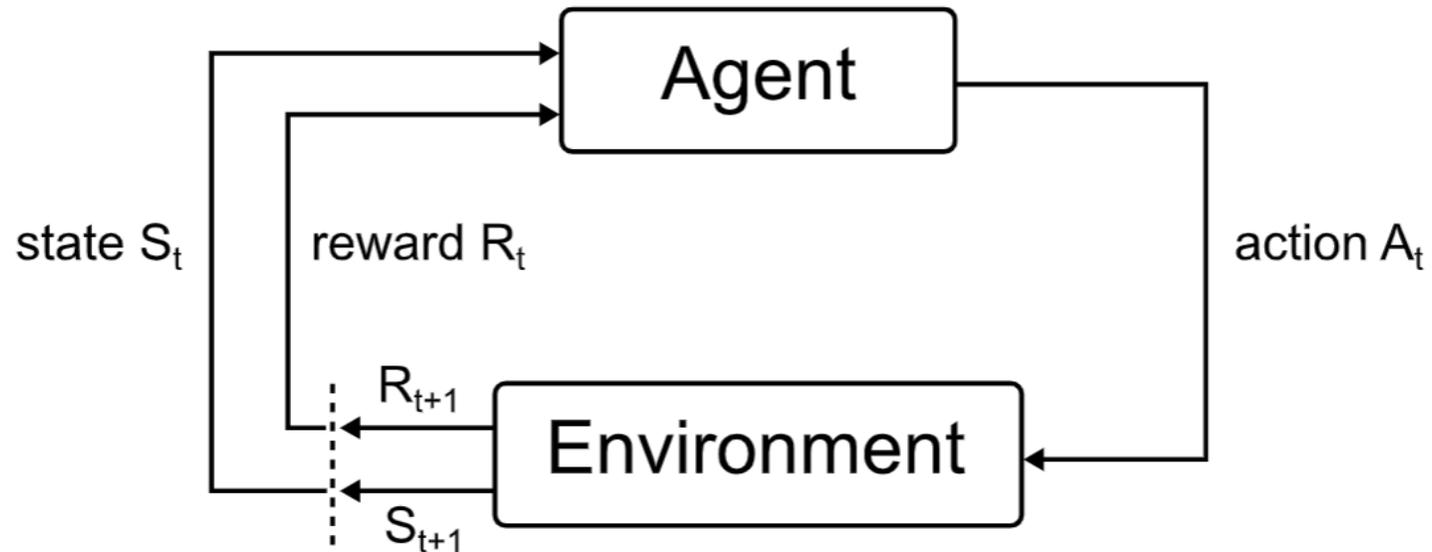
1. Problématique

- Comment optimiser la valeur produite par les actions de systèmes autonomes stationnaires produisant leur propre énergie ?
- On se concentre sur un sous-ensemble de systèmes de l'Internet des objets :
 1. **Immobile** et positionné à **l'extérieur**, donc soumis aux contraintes de son environnement.
 2. Basé sur un **ordinateur mono-carte** (type Raspberry Pi ou Arduino)
 3. Travaille sous un **budget énergétique restreint** et doit être capable de **produire de façon autonome sa propre énergie**
 4. Avoir accès à un **réseau** pour de la communication système-serveur ou système-système.



Sur le toit du campus d'aivancity, avril 2022

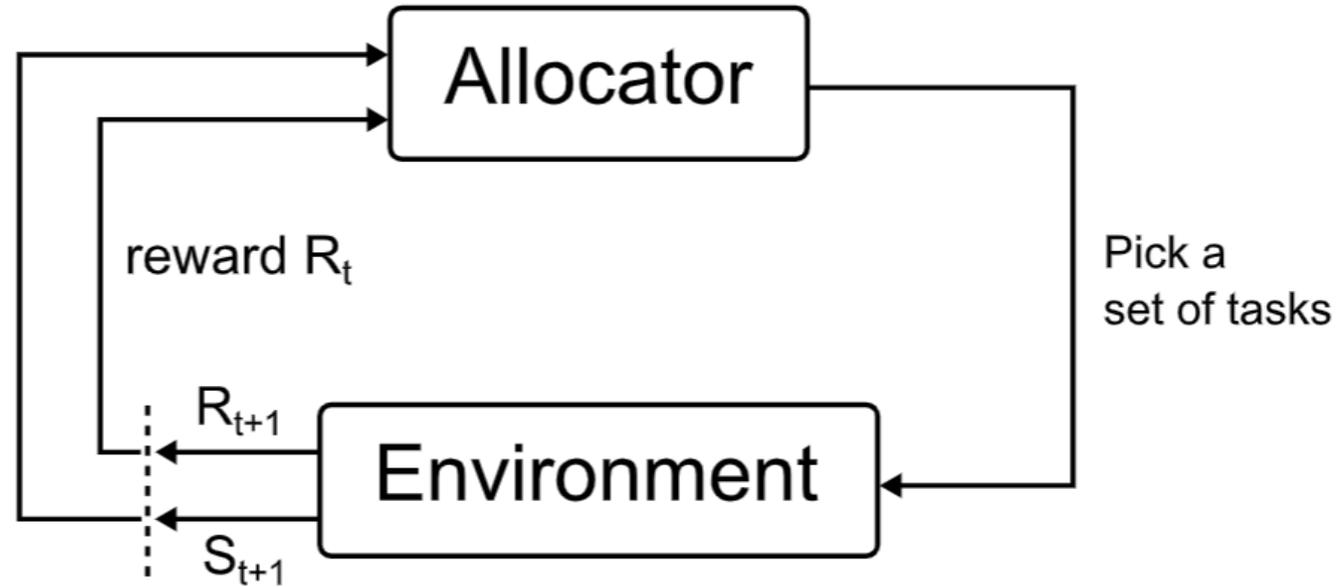
2. L'apprentissage par renforcement



2. L'apprentissage par renforcement appliqué à notre cas

S_t :

- weather
- success history
- battery % (optional)

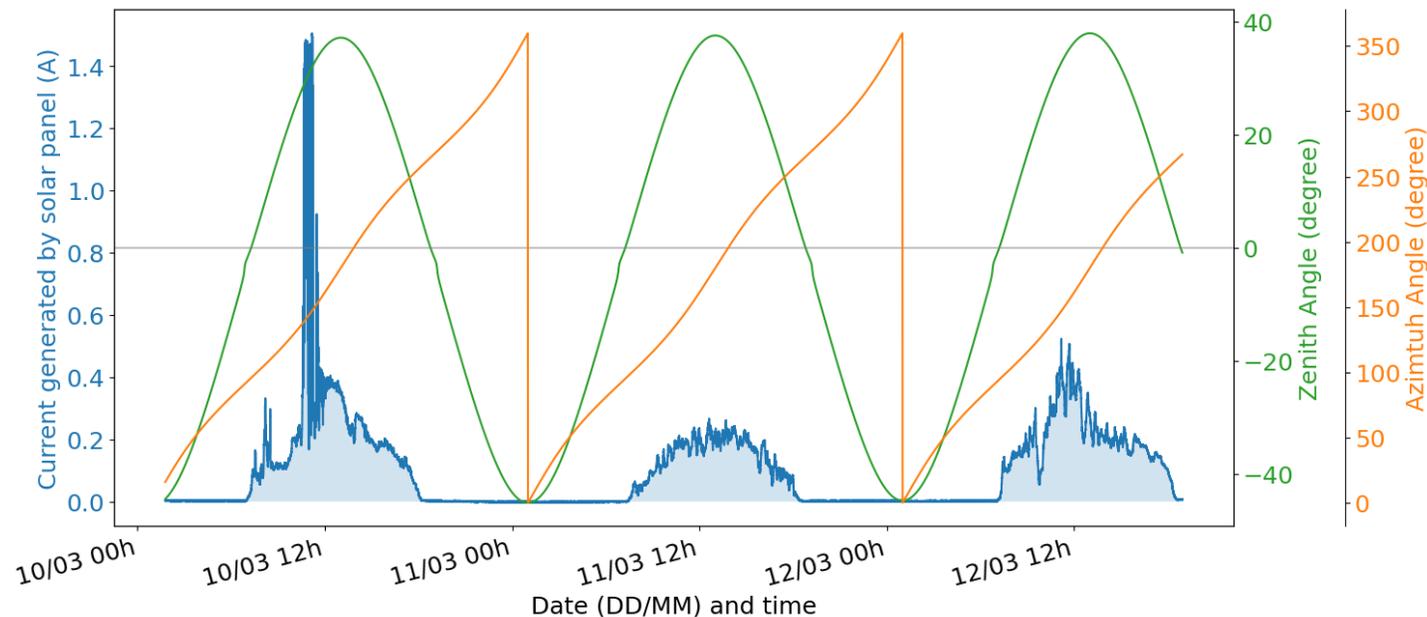


3. Modélisation de l'environnement – Les récompenses

- L'environnement aide à créer à la volée les données, lorsque l'agent apprend.
Des données de qualité = un modèle de qualité
- C'est aussi l'agent qui crée les données car il explore.
Exploration vs. Exploitation
- Premièrement, on doit définir les récompenses, car c'est grâce à celles-ci que l'agent peut ajuster sa politique (stratégie).
 -  Si on a assez de batterie pour effectuer les actions demandées par l'agent, alors on reçoit la somme des valeurs ajoutées des actions.
 -  Si on est dans un intervalle de pourcentages de batterie critique (de 0% à 20% par exemple), alors on reçoit une récompense négative, proportionnelle à la somme des potentielles valeurs ajoutées.
 -  Si le système ne peut même pas s'allumer pour réaliser les actions, on reçoit une pénalité maximale.

3. Modélisation de l'environnement – Les dynamiques

- L'environnement est dynamique : sa météo évolue, l'ensoleillement aussi. Ce sont les facteurs les plus importants lorsqu'on essaie d'estimer la production solaire.
- Période de collection de données de production solaire pour une estimation plus précise qu'actuellement (heuristique)
- Ensuite : intégration à l'environnement des modèles de Kraemer *et al.*, "Operationalizing Solar Energy Predictions for Sustainable, Autonomous IoT Device Management," 2020.



3. Modélisation de l'environnement – Les observations montrées à l'agent

- C'est un point tout aussi clé que les précédents : il définit l'information à disposition de l'agent afin de choisir l'action la plus adaptée.
- Réfléter la réalité :
 - Montrer l'historique des succès/échecs (en regardant les logs du système)
 - Montrer la météo et l'ensoleillement, nécessaire à l'estimation du budget énergétique.

4. Proximal Policy Optimization

PPO collecte des expériences de l'agent pendant un certain nombre d'itérations avant de mettre à jour la politique : moins de variance dans les mises à jour,

Maximisation par descente de gradient

Fonction avantage : détermine la valeur relative d'une action

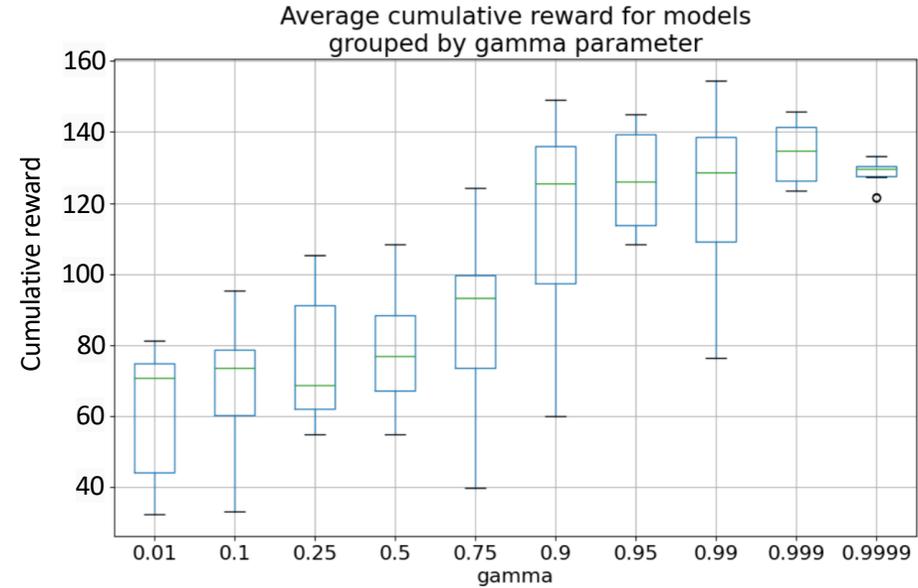
$$L^{CLIP}(\theta) = \hat{E}_t[\min(r_t(\theta) \hat{A}_t, \underbrace{\text{clip}(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon)}_{\text{Permet de fixer la nouvelle valeur dans un intervalle par trop loin de l'ancienne} \rightarrow \text{Régulariser l'apprentissage}}) \hat{A}_t)]$$

Fonction objectif

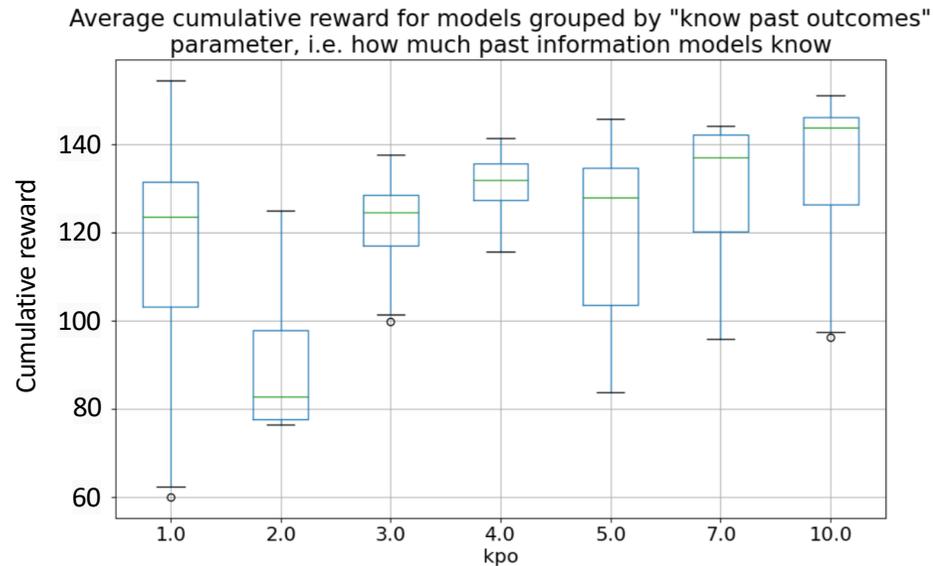
Rapport :
nouvelle politique / ancienne politique
(Politique : observation \rightarrow action suggérée)

4. Résultats – Analyse de sensibilité

- Optimiser les hyper-paramètres :

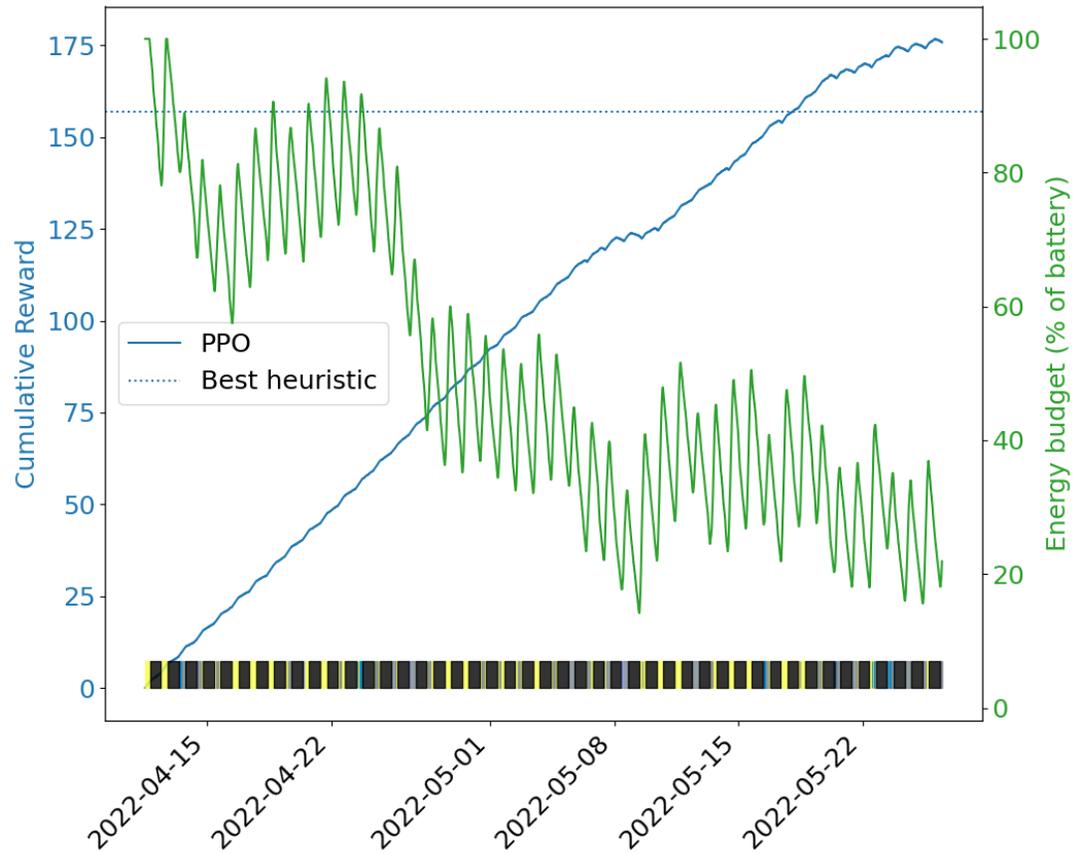


- Vérifier que l'on apprend mieux avec plus d'informations

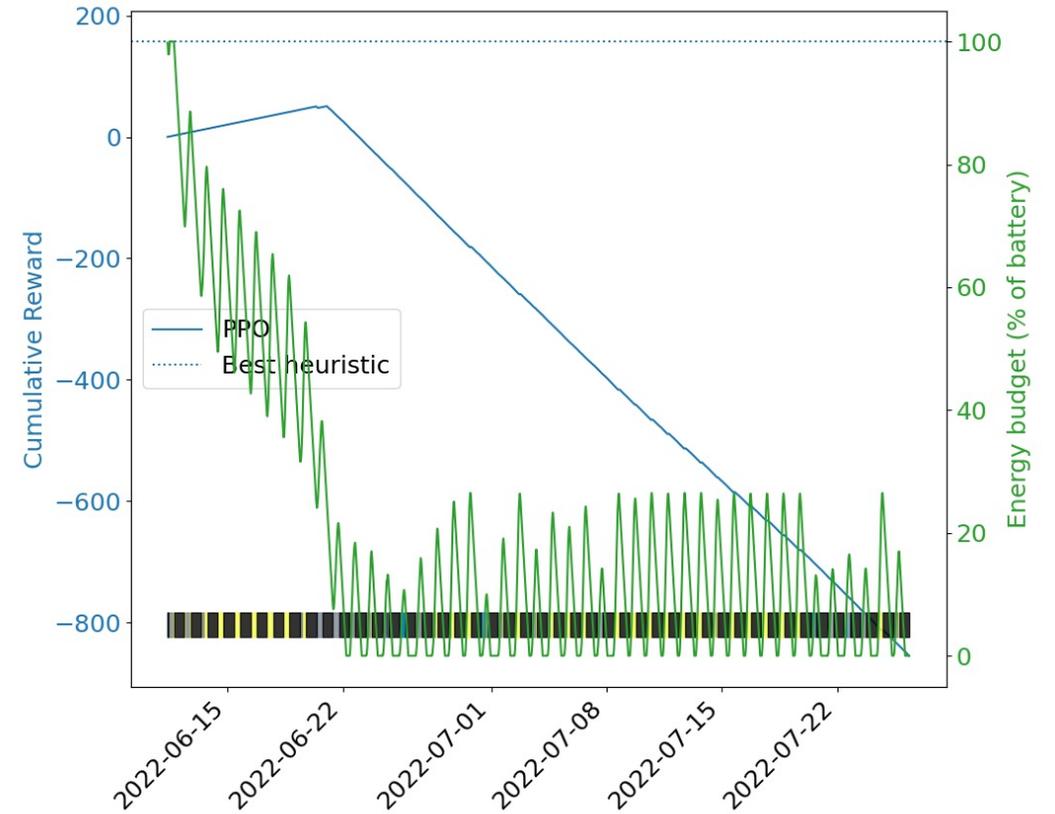


4. Résultats

Meilleur modèle (gamma = 0.999 ; kpo = 10)



Un mauvais modèle (paramètres pas optimaux)



5. Perspectives

- Améliorer les performances en explorant des d'apprentissages évolutifs
- Réaliser une étude sur l'explicabilité des stratégies apprises par les modèles.
- Intégrer le modèle de prédiction de production solaire dans les dynamiques de l'environnement
- Déployer les modèles en condition réelles, ajuster les dynamiques de l'environnement (exemple : coefficient de perte pour le budget énergétique).

Inria

UNIVERSITÉ
DE LYON



ENS DE LYON



aivancity

SCHOOL FOR

TECHNOLOGY, BUSINESS & SOCIETY

Merci, avez-vous des questions ?

Optimiser la sélection de tâches sous
contrainte énergétique dans des systèmes IoT
grâce à l'apprentissage par renforcement

hugo.hadjur@ens-lyon.fr

H. Hadjur, D. Ammar, L. Lefèvre

Mars 2023