

1. Informations générales

Intitulé	Irrigation intelligente : utilisation des données de capteurs connectés pour automatiser l'irrigation des cultures et réduire la consommation d'eau
Lieu de travail	Marseille – France
Type de contrat	Thèse CIFRE (CDD 3 ans)
Salaire mensuel	2000€ brut
Mots clés	IA, Data Science, Contrôle, Pronostic, Irrigation,
Contacts	<p>Telaqua : jf.pons@telaqua.com www.telaqua.com</p> <p>LIS : christophe.roman@lis-lab.fr https://www.lis-lab.fr/</p>

2. Problématique industrielle

2.1. Contexte général

Telaqua développe des systèmes intelligents de surveillance et de contrôle des systèmes d'irrigation présents dans les exploitations agricoles. Notamment ceux-ci permettent de surveiller le système d'irrigation à l'aide de relevés de pressions et de débit en temps réel, de piloter le système via la programmation de l'activation des électrovannes ou des pompes et de superviser l'efficacité de l'irrigation à l'aide de capteurs d'humidité du sol à différentes profondeur.

Aujourd'hui, les différentes solutions en place permettent à l'agriculteur ou à l'exploitant de connaître en temps réel les paramètres et caractéristiques de son système d'irrigation et de ces champs ce qui lui permet déjà d'adapter ses phases d'irrigation. L'objectif de Telaqua est d'aller plus loin en proposant un système d'irrigation entièrement autonome capable de contrôler les phases d'irrigation de manière à apporter à la plante la quantité d'eau suffisante ainsi que de détecter d'éventuelles anomalies. Les principaux objectifs étant de décharger l'agriculteur de ces tâches et de réduire la consommation d'eau nécessaire à l'irrigation.

2.2. Objectifs des travaux de recherche

2.2.1. Détection d'anomalies

Le premier objectif auquel ces travaux de recherche devront répondre est la détection d'anomalies sur les signaux issus des capteurs de Telaqua et utilisés pour réaliser l'irrigation automatique. Les principales anomalies à détecter sont l'absence d'eau dans les tensiomètres ainsi que les erreurs de mesures liées à un faible niveau de batterie des dispositifs, à de mauvais branchements des câbles ou à des capteurs défaillants.

2.2.2. L'irrigation automatique

Le second objectif auquel les travaux de recherche devront répondre est l'irrigation des champs de manière automatique c'est-à-dire irriguer le champ avant que le sol devienne trop sec au point de mettre en péril la récolte et avec la quantité d'eau juste nécessaire pour atteindre le taux d'humidité souhaité à la profondeur souhaitée. Ce second objectif pourra être subdivisé en plusieurs sous-objectifs comme la détermination de la propagation verticale de l'eau dans le sol et la mise en place

d'une planification optimale d'irrigation. Cette dernière devrait pouvoir être effective à partir exclusivement des données issues des capteurs du champ puis par des données complétées avec des informations complémentaires que sont par exemple les coordonnées GPS, des données météorologiques, le type de culture et les temps d'irrigation eux-mêmes. Dans la mesure du possible, le modèle d'irrigation automatique devrait s'adapter à chaque champ de manière automatique (personnalisation de la solution).

2.3. Proposition d'un plan de thèse

Afin de construire des modèles robustes et pouvant s'adapter à divers profils de sol et d'irrigation, une des premières étapes consistera à récolter suffisamment de données. Pour ce faire, la première étape sera de récolter plus de données grâce à nos futurs projets d'installations en France, en Espagne et en Afrique mais aussi au travers de la mise en place de protocoles de collecte de données spécifiques réalisées avec le support de centres de recherche en agronomie. Par ailleurs, les points que les travaux de recherche devraient couvrir et qui constituent le plan de thèse préliminaire sont les suivants :

1. Détection des anomalies sur un tensiomètre
2. Prédiction de sécheresse avec uniquement les données d'un tensiomètre
3. Prédiction de sécheresse avec plus d'un tensiomètre et des données complémentaires comme les conditions météorologiques, les coordonnées GPS et les temps d'irrigation
4. Définition du temps d'irrigation optimal
5. Analyse de l'intérêt de faire un modèle générique versus un modèle par client
6. Partitionnement IA embarquée vs IA sur cloud

3. Problématique académique relevée par le LIS

La problématique posée par Telaqua est à deux niveaux. Le premier concerne la détection d'anomalies sur les signaux issus des capteurs de Telaqua et utilisés pour réaliser l'irrigation automatique. Le second est une problématique de contrôle du système afin de réaliser une irrigation automatique et optimale des champs en faisant un pronostic sur le temps restant avant sécheresse. Ceci permettra ainsi de déterminer le planning d'irrigation optimal qui prend en considération les besoins de tous les secteurs ainsi que la ressource eau. Ces deux niveaux rentrent dans les thèmes forts de l'équipe PECASE du LIS, entre autres : détection automatique de fautes, diagnostic, pronostic, et contrôle. Ceci soit par des approches statistiques et des approches faisant appel à de l'intelligence artificielle soit par des approches orientées modélisation physique.

Le premier point relève des problèmes de diagnostic et de détection de défaut capteurs, au niveau des tensiomètres. Cette thématique a fait l'objet de plusieurs études dans la littérature et s'inscrit dans le domaine de la validation des données capteur. L'étude de la détection de fautes sur les tensiomètres pourra être augmentée par du diagnostic. En effet il est possible par rapport aux données de déterminer à quelles causes sont liées les défauts (e.g. perte de colonne d'eau, batterie faible, câble défectueux) par des techniques de traitement de l'information, de redondance matérielle et de redondance analytique tout en respectant les capacités de calculs disponibles. Et si les données sont suffisamment riches, il est possible d'étudier la déviation résiduelle pour s'orienter vers le pronostic, c'est-à-dire prédire quand un défaut arrivera. Le diagnostic automatique et le pronostic sont à ce jour des thèmes de recherche très actifs.

Dans le cas de la prédiction de sécheresse, on parlera aussi de pronostic de dégradation du système. Les différentes approches de pronostic peuvent être décomposées en fonction de leur applicabilité et

de leur rentabilité en trois familles : les approches à base de modèles, les approches guidées par les données et les approches basées sur l'expérience. Dans le cadre de cette thèse nous proposons d'étudier et d'exploiter les approches à base de modèles et celles guidées par les données. Les modèles développés dans la littérature sur la filtration de l'eau dans les sols non saturés semblent être une première piste dans la création et la validation d'un modèle à même de décrire l'évolution de l'humidité dans le sol. Néanmoins ces modèles (Équation de Richards et théorie de la percolation) ne disposent pas des caractéristiques intéressantes pour la prédiction et le contrôle optimal. En particulier les simulations de l'équation de Richards ont tendance à être instables, ce qui constitue un verrou scientifique important pour notre utilisation à savoir la prédiction. La deuxième piste est l'utilisation de modèles génériques issus du domaine de l'intelligence artificielle comme par exemple les systèmes autorégressifs, les réseaux de neurones récurrents ou les réseaux de neurones convolus. Les approches fondées sur les données exigent que les informations issues des capteurs soient suffisantes en qualité et en quantité pour évaluer l'état actuel du sol. L'évolution de l'indicateur de dégradation est alors réalisée à l'aide d'une méthode statistique. On distingue trois classes d'approches : le pronostic par analyse de tendance, par apprentissage et par estimation d'état.

Le travail sur un seul capteur permettra de valider ou de réfuter certains modèles/méthodes grâce aux données réelles. Ceci effectué, l'extension multi-capteurs et multi-entrées (irrigation et conditions météorologiques, etc.) constitue un problème distribué. Distribué en surface où le maillage des capteurs est connu par les données GPS et potentiellement distribué en profondeur. En effet, il peut être très intéressant de caractériser le taux de pénétration de l'eau dans le sol. L'analyse, la simulation et le contrôle (optimal) de ce type de problème peut s'avérer très délicat et est encore une thématique de recherche très active, dont une grande partie des travaux ne sont que théoriques. Dans le cas des approches par modèles, soit les modèles considérés seront suffisamment proches de l'existant pour qu'une simple adaptation des méthodes soit suffisante, soit il faudra en développer de nouveaux. Dans le cas des approches par données, l'explosion en termes de quantités données et leur corrélation peuvent avoir un effet dramatique sur les méthodes mises en œuvre. Il faudra proposer des mises en œuvre innovantes afin de pallier ce problème.

Ces travaux fondamentaux permettront aussi en accord avec les méthodes validées de sélectionner entre un modèle général adaptatif ou un modèle adapté pour chaque exploitation. De façon similaire le choix entre une implémentation cloud ou embarquée sera associé à la complexité des stratégies mises en œuvre et des calculs numériques nécessaires.

Il est ainsi prévu de découper le travail de thèse en cinq grandes phases :

1. Analyse de l'existant et caractérisation des problématiques
2. Proposition et validation des nouvelles approches permettant d'améliorer les performances du diagnostic de défauts capteurs et sa localisation
3. Proposition de méthode de pronostic du temps restant avant la sécheresse du sol
4. Proposition d'une planification optimale de l'irrigation des champs
5. Intégration de l'outil de diagnostic et pronostic dans l'environnement de Telaqua et intégration dans les systèmes automatisés