

Proposition de sujet d'alternance 1A 2023-2024

22 septembre 2023

Laboratoire : LIS (UMR 7020)

Titre du sujet : **Irrigation intelligente** - Prédiction de série temporelle pour l'estimation de l'humidité dans le sol.

Encadrants : Christophe Roman (Maitre de conférences),
Guillaume Graton (Maitre de conférences).

Localisation : Domaine universitaire de Saint Jérôme,
Avenue Escadrille Normandie Niemen,
13397 MARSEILLE Cedex 20.

Mails : christophe.roman@lis-lab.fr,
guillaume.graton@lis-lab.fr.

Cadre

Ce sujet d'alternance est associé à un projet en partenariat avec une entreprise qui donne lieu actuellement à une thèse CIFRE. La problématique est celle de l'irrigation intelligente.

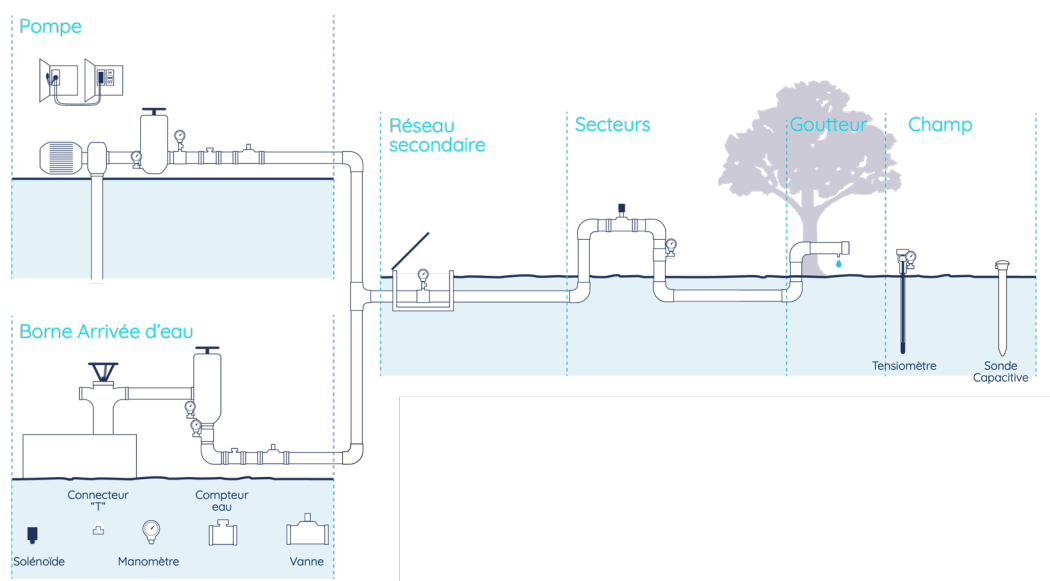


FIGURE 1 – Illustration d'un système d'irrigation.

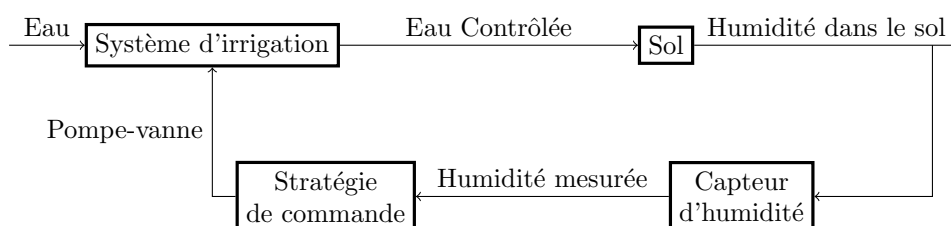


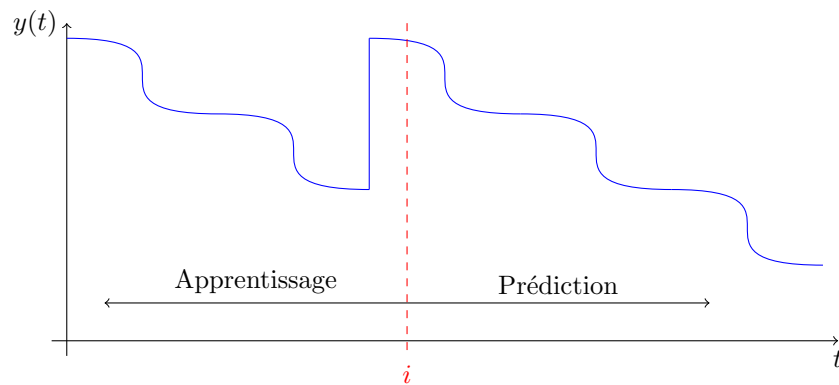
FIGURE 2 – Schéma du système d'irrigation automatique

Contexte

L'eau est une ressource précieuse, son utilisation de façon optimale est fondamentale pour notre civilisation. Dans le cadre de l'irrigation, il y a deux points limites. Soit, l'irrigation est trop importante et une partie de la ressource n'est pas utilisée par la plante. Soit, l'irrigation n'est pas suffisante, ceci peut aller jusqu'à la perte de la plante et par conséquent le gâchis de toute la quantité d'eau fournie précédemment à la plante. Un des objectifs de la thèse CIFRE est d'optimiser les plannings d'irrigations. Pour arriver à cet objectif, il est nécessaire d'avoir un modèle permettant de prédire l'humidité dans le sol. La propagation de l'eau dans le sol est un phénomène complexe, de plus certaines perturbations peuvent avoir lieu (rosée, pluie ...) qui sont difficilement détectables et prévisibles. Finalement, le système évolue dans le temps, par exemple plus la plante est mature, plus elle absorbe d'eau. Dès lors, la dynamique de l'humidité dans le sol se trouve modifiée.

C'est sur ce sujet de **modèle de prédiction** que d'alternance repose.

Concrètement on veut prédire l'humidité en prenant en compte le comportement passé du système. Pour prédire cette humidité, nous allons créer un modèle qui va apprendre du passé pour prédire l'état futur. Dans la suite $y(t)$ représente l'humidité à l'instant t :



La méthode développée pour l'instant est basée sur un modèle type ARX (AutoRegressive Xéogenous) commuté. L'intérêt a été de pouvoir expliquer la construction de ce modèle par des considérations physiques et l'apprentissage de ce type de modèle est relativement simple et performant. De nombreuses améliorations néanmoins sont possibles.

Cependant, il est intéressant de comparer ses performances avec les modèles plus classiques de la littérature. La performance ne prendra pas en compte que l'erreur de prédiction, mais aussi le temps de calcul nécessaire à l'apprentissage. Nous avons en vue

- Les réseaux de neurones à travers les LSTM “Long Short - Term Memory” qui sont un type de réseau de neurones récurrent.
- Les modèles de forêt aléatoire avec des apprentissages de type XGboost.
- Les machines à vecteur de support, qui bien qu'utilisées massivement pour la classification, peuvent être aussi utilisé pour la régression.
- Des améliorations de la 1^{er} méthode avec de la logique floue.

Détails sur les LSTM

Les réseaux de neurones récurrents ont une structure autorégressive avec une fonction d'activation. Certaines structures particulières permettent d'ajouter un effet mémoire sur le modèle. Les LSTM ont une structure permettant d'avoir une mémoire long terme et une mémoire court terme. La structure mathématique est la suivante :

$$f(t) = \sigma_f(W_f x(t) + U_f h(t-1) + b_f), \quad (1)$$

$$i(t) = \sigma_i(W_i x(t) + U_i h(t-1) + b_i), \quad (2)$$

$$o(t) = \sigma_o(W_o x(t) + U_o h(t-1) + b_o), \quad (3)$$

$$\tilde{c}(t) = \sigma_c(W_c x(t) + U_c h(t-1) + b_c), \quad (4)$$

$$c(t) = f(t) \cdot c(t-1) + i(t) \cdot \tilde{c}(t), \quad c(0) = 0 \quad (5)$$

$$h(t) = o(t) \cdot \sigma_h(c(t)), \quad h(0) = 0. \quad (6)$$

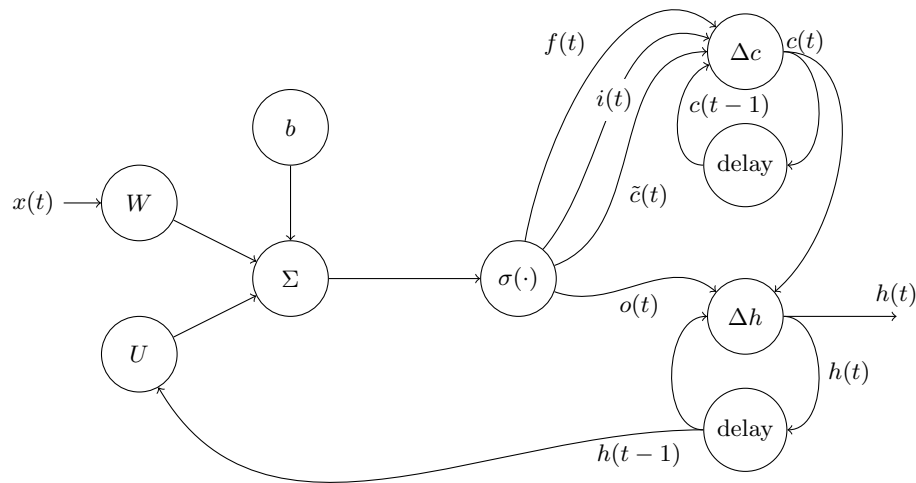


FIGURE 3 – Schéma des interconnexions d'un LSTM

Précision sur le stage

On dispose de données de capteur, certains relevant l'humidité dans le sol à certaines profondeurs, et d'autres relevant la pression d'eau envoyée en irrigation. L'idée est donc d'arriver à modéliser le système en se basant sur les données réelles issues des capteurs. Pour l'instant, plusieurs variations de modèles ont été proposées, avec des résultats acceptables. Il est néanmoins possible de les améliorer.

Objectif de l'alternance

L'alternance concerne le développement et la validation de modèle de *machine learning* pour la prédiction de série temporelle à appliquer au domaine de l'irrigation.

Un des objectifs de l'alternance est de développer d'autres modèles (on pense par exemple au modèle classique utilisé en *time series forecasting* et en apprentissage, à savoir réseau de Neurones, forêt aléatoire ... déjà cités plus haut).

Des modèles codés sous *Python* fonctionnant seront fournis, ainsi qu'un ensemble de données.

Attente de l'alternant

Nous attendons de l'alternant du sérieux, des compétences de base en mathématique et informatique par rapport à son niveau d'étude. Le travail impliquera du développement de code sous Python et une compréhension mathématique des objets et méthodes utilisées.

Validation pour mise en ligne ECM

IDM