

WP3 : Nouvelles technologies d'acquisition et d'exploitation des données massives multi sources

T3.1. Collecte des données disponibles auprès de l'observatoire

Les données prévues à collecter dans le cadre du projet GEANTech concernent trois bassins versants : le bassin de Souss, le bassin de Tensift et le bassin d'Oum Errabia. Les échelles spatiales ciblées sont les échelles des processus de prise de décision allant de la parcelle au secteur irrigué pour la planification de l'eau d'irrigation, au bassin pour la planification intégrée du bassin hydraulique. En plus des mesures expérimentales spécifiques au projet GEANTech, qui ne sont pas en cours réalisées faute du retard de l'acquisition des équipements scientifiques, la collecte de certaines données a reposé sur d'autres projets scientifiques spécifiques de chaque bassin versant (voir en bas).

- **Bassin de Tensift**

A nos jours, la réalisation des travaux du projet est reposée sur la base de données et le dispositif expérimental du Laboratoire Mixte International LMI-TREMA (Télédétection et Ressources en Eau en Méditerranée semi-Aride, <https://www.lmi-trema.ma/observatoire/>) basé à l'université Cadi Ayyad à Marrakech. A cette base de données, s'ajoutent des données récemment collectées dans le cadre de différents projets de recherche (ASSIWAT de l'UM6P, PRIMA-IDEWA de l'UCAM) sur un site d'oliviers situé dans le verger d'Agdal autour de la ville de Marrakech. Dans ce site, deux parcelles irriguées par deux systèmes le système goutte à goutte (superficiel et enterré) ont été équipées. Un système d'Eddy covariance complet a été installé sur chaque parcelle pour le suivi de différentes composantes du continuum Sol-Plante-Atmosphère (voir § T 3.3).

Une autre expérience récente (Jan-Juin, 2023) sur l'irrigation et la fertilisation déficitaires a été conduite sur la culture de quinoa. 10 traitements de différente quantité d'irrigation et de de fertilisation ont été mises en place dans le site pilote Chichaoua (Figure1).

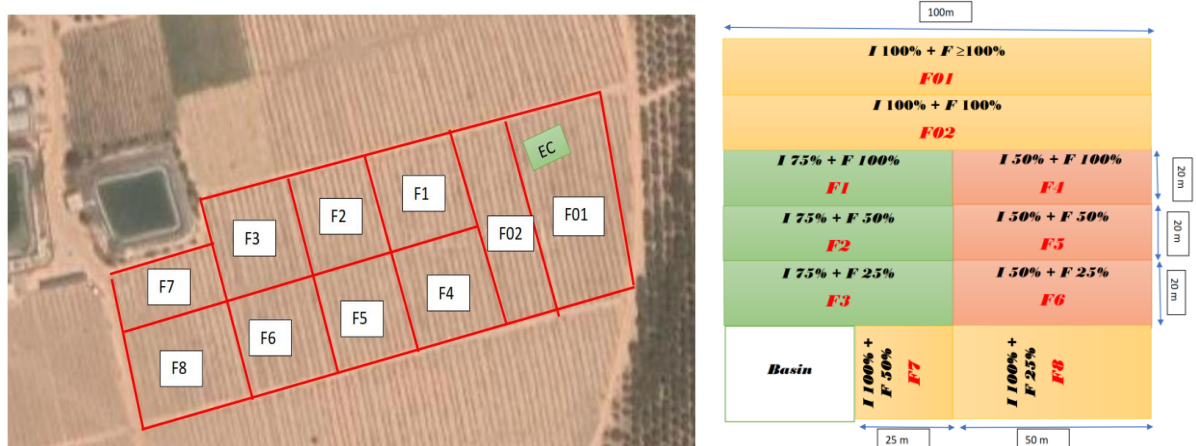


Figure 1: Différentes parcelles de quinoa suivies de différents traitements de l'irrigation et de la fertilisation.

Sur chaque traitement (figure 2), la quantité d'irrigation et de fertilisation apportée sont différentes et décrites comme suit

✓ **Fertilisation:**

- ❖ La fertilisation par NPK au début de l'expérience est apportée dans chaque parcelle comme suit :
 - F100%=Fertilisation 100% =2kg/ligne/100m
 - F50%= Fertilisation 50%=1kg/ligne/100m
 - F25%= Fertilisation 25%=0.5kg/ligne/100m
- ❖ En milieu de la saison agricole, nous avons appliqué uniquement la fertilisation azotée, et la quantité de fertilisation a été ajustée en fonction des traitements (25 %, 50 %, 100 %)
- ❖ La différence entre F01 et F02 est qu'il y a un excès de fertilisation dans la parcelle F01 par rapport à F02.

✓ **Irrigation :**

I100%=Irrigation 100%, cela signifie que ce traitement a reçu le besoin maximum d'eau en irrigation (pas de stress)

I75%= Irrigation 75%. Ce traitement a reçu une quantité d'irrigation de 25% moins par rapport au traitement I100%. Un léger stress hydrique a été appliqué.

I50%= Irrigation 50%. Ce traitement a reçu la moitié de la quantité d'irrigation appliquée au traitement I100%. Un fort stress hydrique a été appliqué.



Figure 2: Quelques photos de quinoa pendant différentes phases de développement.

Différentes mesures relatives au sol et à la végétation ont été collectées sur chaque parcelle.

✓ **Mesures relatives au développement des plantes:**

Les données collectées (figure 3) chaque semaine sont :

- ✓ Contenu en eau du sol
- ✓ Indice de Végétation par Différence Normalisée (NDVI)
- ✓ Photos hémisphériques pour évaluer la couverture végétale
- ✓ Images thermiques pour suivre les variations de température de la surface
- ✓ Prélèvement d'échantillons de la plante avant et après l'application d'azote en vue d'analyses minérales
- ✓ Biomasse végétale
- ✓ Rendement final de la culture



Figure 3: Différentes mesures effectuées: a) Mesure du Contenu en Eau du Sol, b) Séchage de la Biomasse au Four, c) Mesure de la Couverture Végétale, d) Mesures de l'Indice de Végétation (NDVI).

✓ **Mesures liées aux Analyses des Échantillons de Sol :**

En complément de la collecte de données de végétation, nous avons collecté des échantillons de sol et de végétation avant et après l'application de la fertilisation. Ces analyses ont fourni des informations précieuses pour comprendre les caractéristiques du sol et leur impact sur la culture. Les éléments analysés comprenaient :

- ✓ pH du sol
- ✓ Conductivité du sol
- ✓ Texture du sol
- ✓ Teneur totale en azote

- ✓ Quantité d'éléments présents dans le sol
- ✓ Teneur totale en azote dans la végétation

- **Bassin de Souss**

Runoff prediction

In the following, we make predictions of rainfall-runoff using several models concerning Machine learning such as XGboost, Random Forest and SVM, as well as Deep learning models. For the dataset, we have used hourly data extract from Era5-land between the years 2000 and 2022 of three different areas, namely, Aoulouz, Taroudant, and Aitmelloul.

1. Data

We have used hourly Era5-land data of the region of Taroudant between the years 2000 and 2022. The whole dataset comprised +200K samples, and was used according to the following table:

	TRAINING SET	VALIDATION SET	TEST SET
SAMPLES	112909	48390	40325

The precipitation-runoff plot for each area is illustrated in the following:

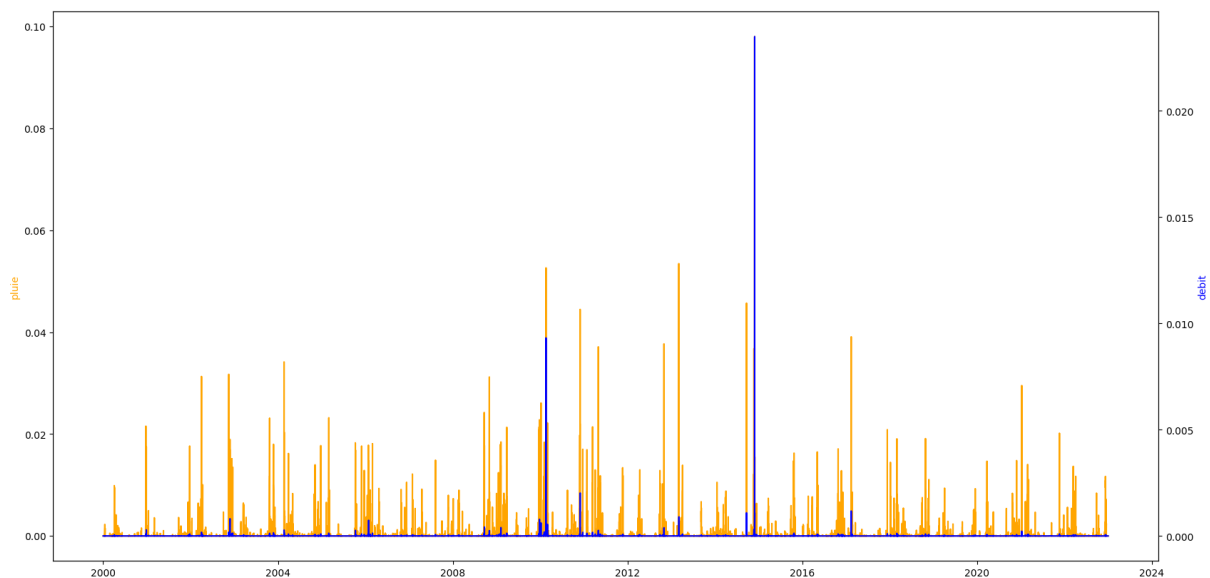


Figure 2: precipitation-runoff plot for Aoulouz

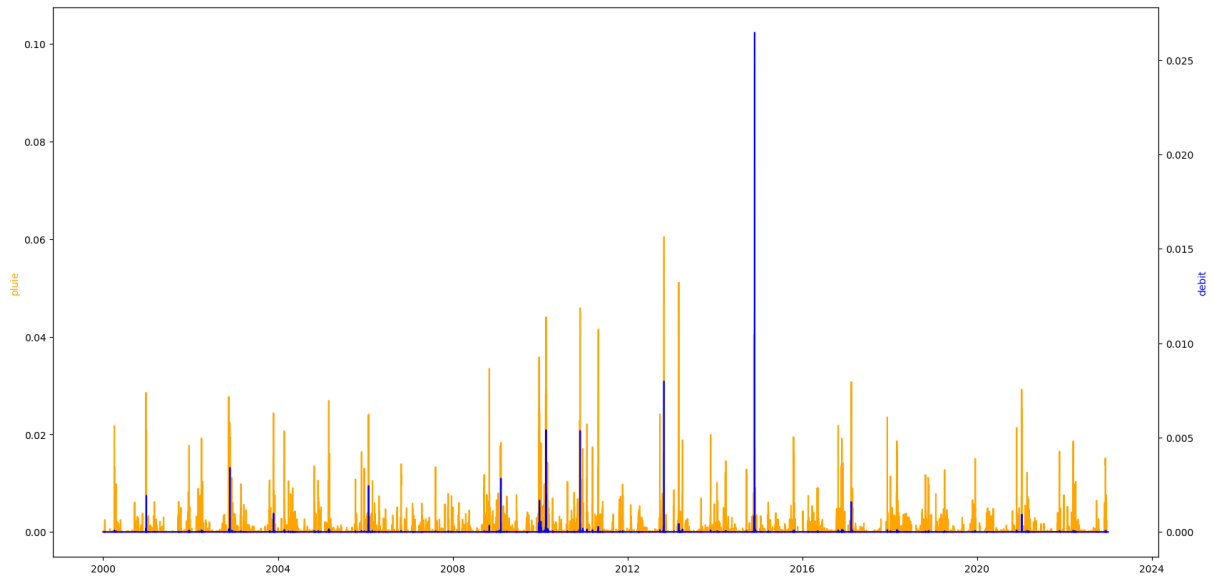


Figure 3: precipitation-runoff plot for Taroudant

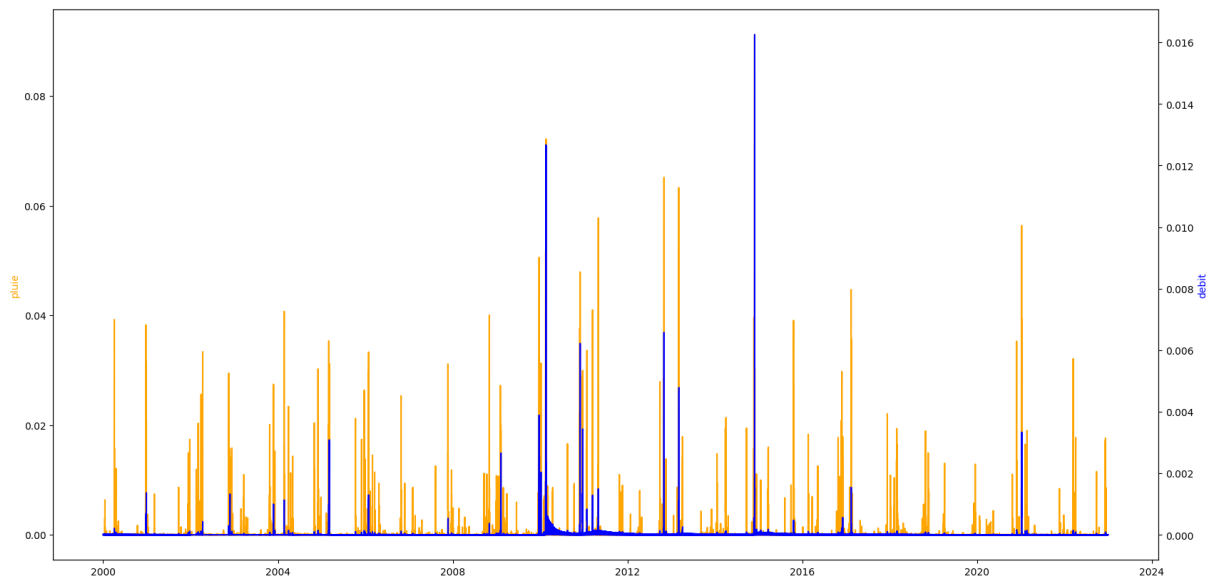


Figure 4: precipitation-runoff plot for Aitmelloul

2. Machine learning models

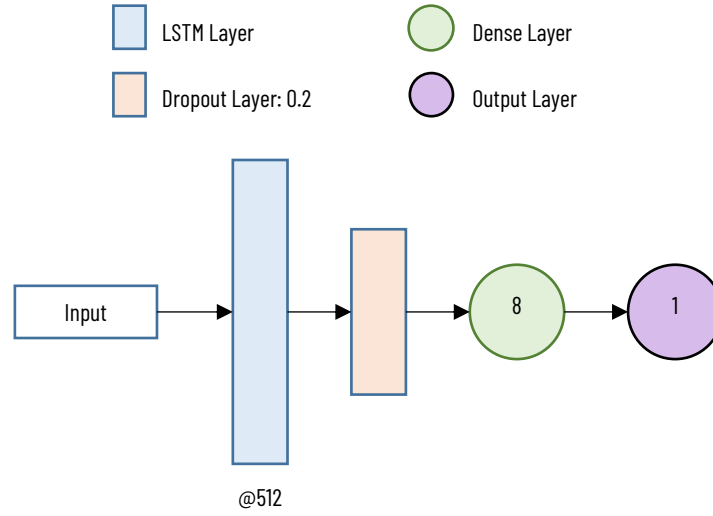
In this section, we have performed this regression task using four different machine learning models, namely:

- Random forest
- Decision tree
- XGBoost

3. Deep learning models

- a. **LSTM Architecture 1:** The first architecture is composed of an LSTM layer with 512 neurons, followed by a dropout layer to avoid overfitting, a dense layer with 8 neurons and a Relu activation function, and lastly, a regression layer. Figure 5 illustrates the architectures.

Figure 5: The first LSTM architecture used (LSTM 1)



- b. **LSTM architecture 2:** This architecture is similar yet simpler than the previous one. While the previous architecture's LSTM layer has 512 neurons, this architecture only has 16. The rest of the layers have the same composition.

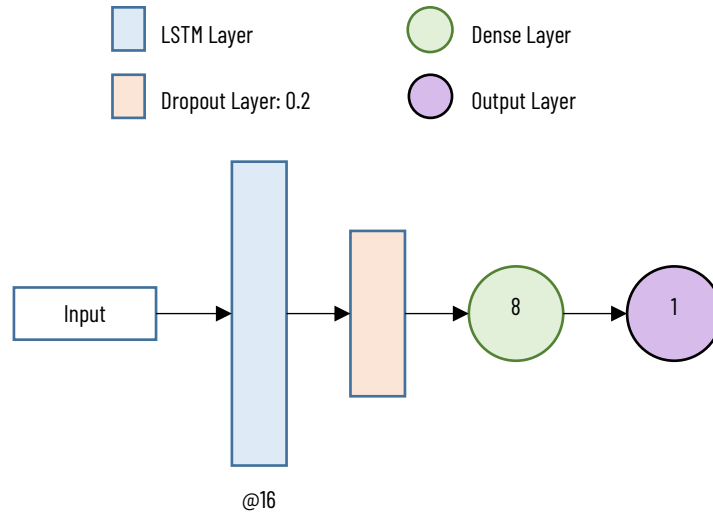


Figure 6: The second LSTM architecture used (LSTM 2)

- c. **Bidirectional LSTM:** Another approach is the usage of bidirectional LSTM. We have used the same LSTM 2 architecture, however, instead of a normal LSTM layer, we use a bidirectional LSTM layer.
- d. **CNN:** We have also tried to use the 1-dimensional CNN for this task. The architecture is the same as LSTM 2, we just changed the LSTM layer into a Conv1D layer.

4. Experiments

We have conducted several experiments related to the different zones of interest. Geographically speaking, the three zones are connected via a river that spans the whole length between Aoulouz, passing through Taroudant, and going through Aitmelloul. Hence, it is plausible to say that the precipitation and/or run-off of the zones close to the headwater (the starting point of the river) may affect the precipitation and/or run-off the zones close to the estuary (the point where the river meets the sea).

For this reason, we conducted the following 5 different experiments:

- Aoulouz Run-off prediction using precipitation values.
- Taroudant Run-off prediction using precipitation values
- Aitmelloul Run-off prediction using precipitation values
- Taroudant Run-off prediction using precipitation values and Aoulouz run-off values.
- Aitmelloul Run-off prediction using precipitation values and Taroudant run-off values.

We have trained our DL models for 100 epochs with a batch size of 256. The training progress for each model is illustrated in the corresponding figures, and the results are described in the corresponding tables in terms of R2score.

a. ***Experiment 1: Aoulouz Run-off prediction using precipitation values.***

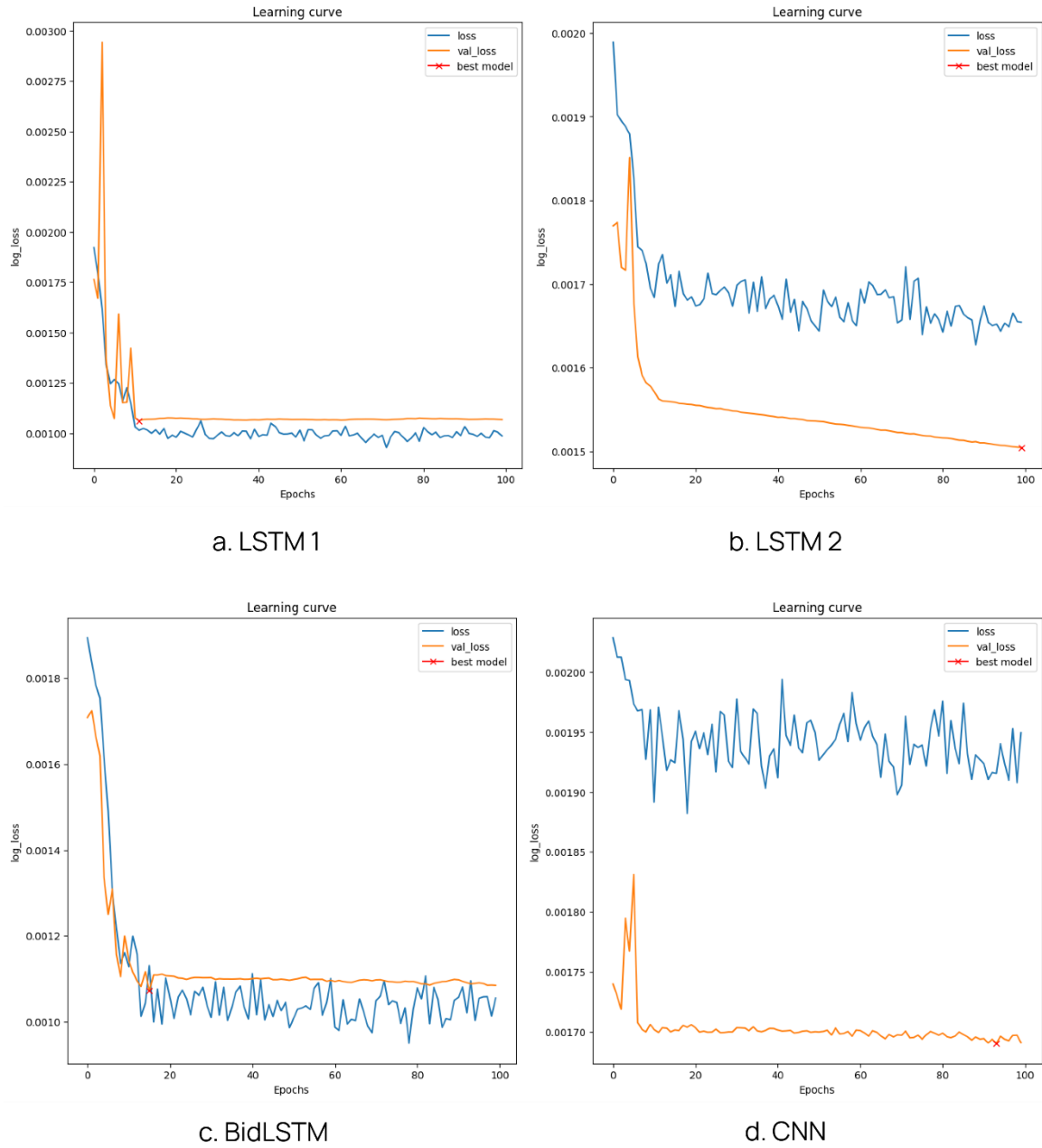


Figure 7: Learning progress of the Deep learning models

	LSTM 1	LSTM 2	BID LSTM	CNN	RANDOM FOREST	DECISION TREE	XGBOOST
TRAINING	0.9328	0.6971	0.9293	0.5708	0.9816	0.9954	0.9988
VALIDATION	0.9006	0.6976	0.8936	0.5707	0.8591	0.8201	0.8178
TEST	0.9116	0.7090	0.9074	0.5726	0.9185	0.8945	0.8943

b. Experiment 2: Taroudant Run-off prediction using precipitation values.

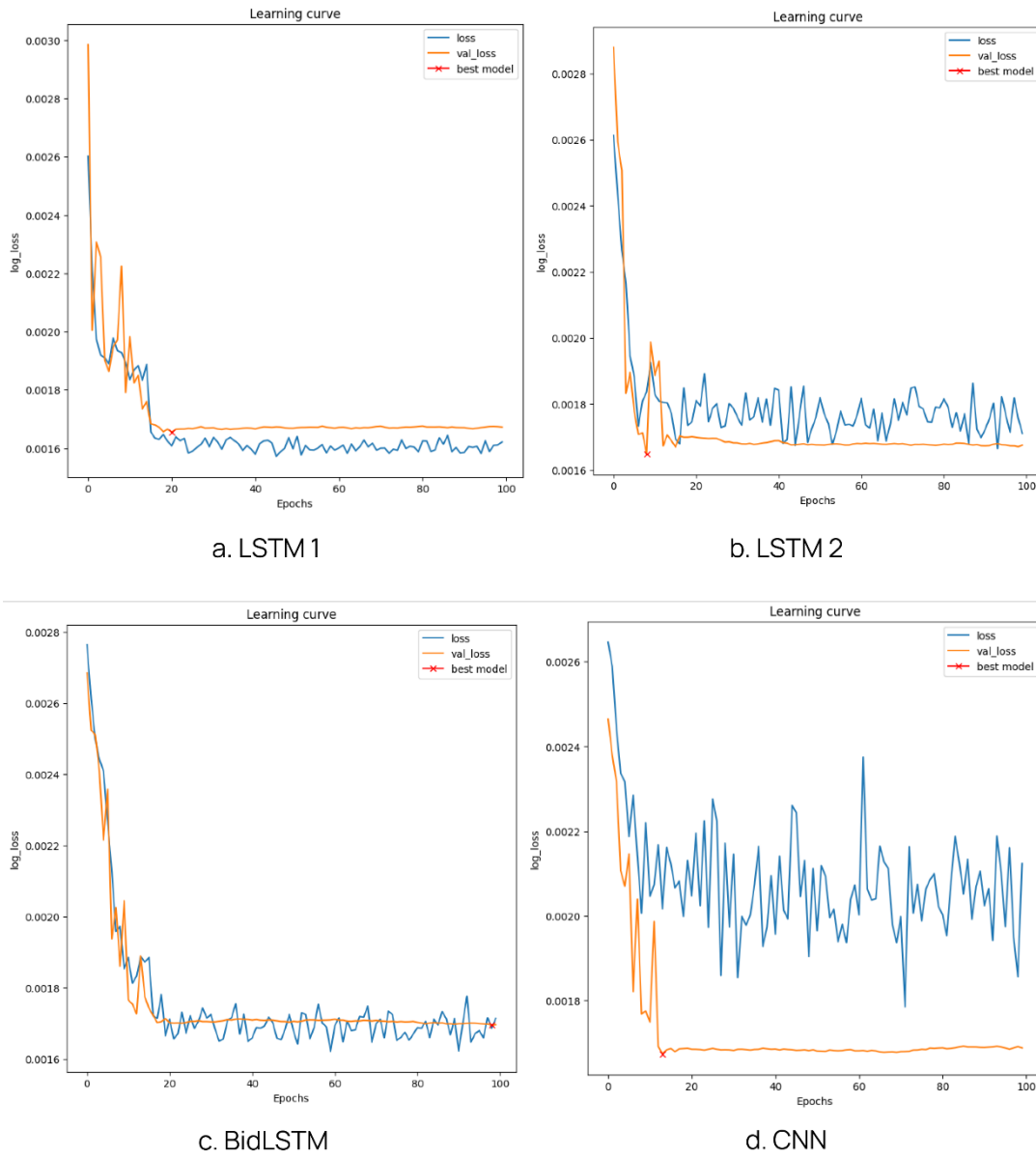


Figure 8: Learning progress of the Deep learning models

	LSTM 1	LSTM 2	BID LSTM	CNN	RANDOM FOREST	DECISION TREE	XGBOOST
TRAINING	0.9270	0.9244	0.9262	0.8922	0.9728	0.9831	0.9944
VALIDATION	0.9048	0.9099	0.9088	0.8983	0.8991	0.8801	0.8715
TEST	0.9294	0.9346	0.9340	0.9201	0.9321	0.9170	0.9137

c. **Experiment 3: Aitmelloul Run-off prediction using precipitation values.**

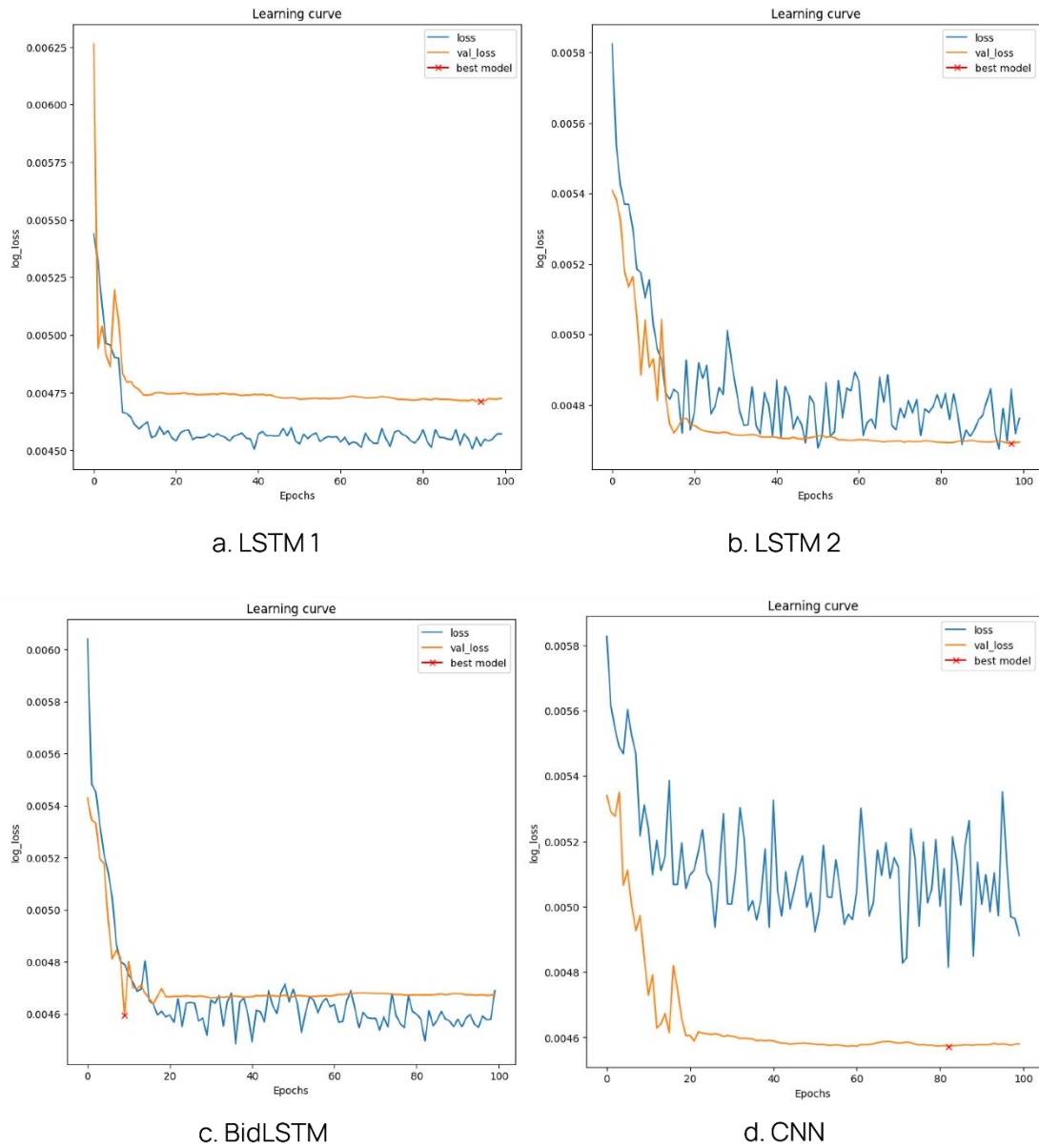


Figure 9: Learning progress of the Deep learning models

	LSTM 1	LSTM 2	BID LSTM	CNN	RANDOM FOREST	DECISION TREE	XGBOOST
TRAINING	0.8723	0.8686	0.8743	0.8508	0.9400	0.9485	0.9781
VALIDATION	0.8393	0.8503	0.8456	0.8408	0.8228	0.7452	0.7191
TEST	0.8816	0.8905	0.8873	0.8750	0.8960	0.8570	0.8511

- d. **Experiment 4: Taroudant Run-off prediction using precipitation values and Aoulouz run-off values.**

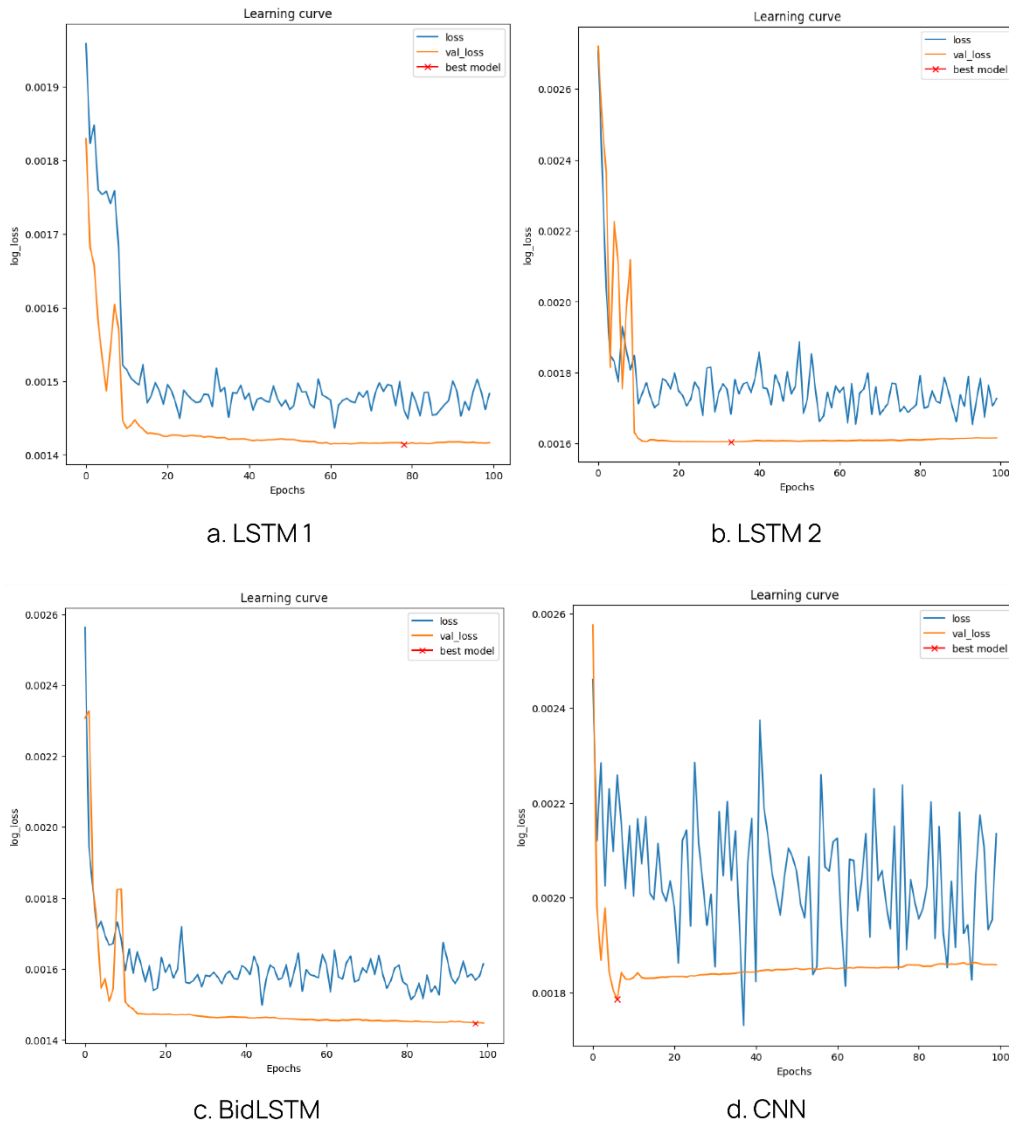


Figure 10: Learning progress of the Deep learning models

	LSTM 1	LSTM 2	BID LSTM	CNN	RANDOM FOREST	DECISION TREE	XGBOOST
TRAINING	0.9624	0.9586	0.9609	0.9240	0.9935	0.9985	0.9997
VALIDATION	0.9624	0.9575	0.9621	0.9222	0.9705	0.9541	0.9620
TEST	0.9715	0.9764	0.9754	0.9415	0.9782	0.9754	0.9747

- e. **Experiment 5: Aitmelloul Run-off prediction using precipitation values and Taroudant run-off values.**

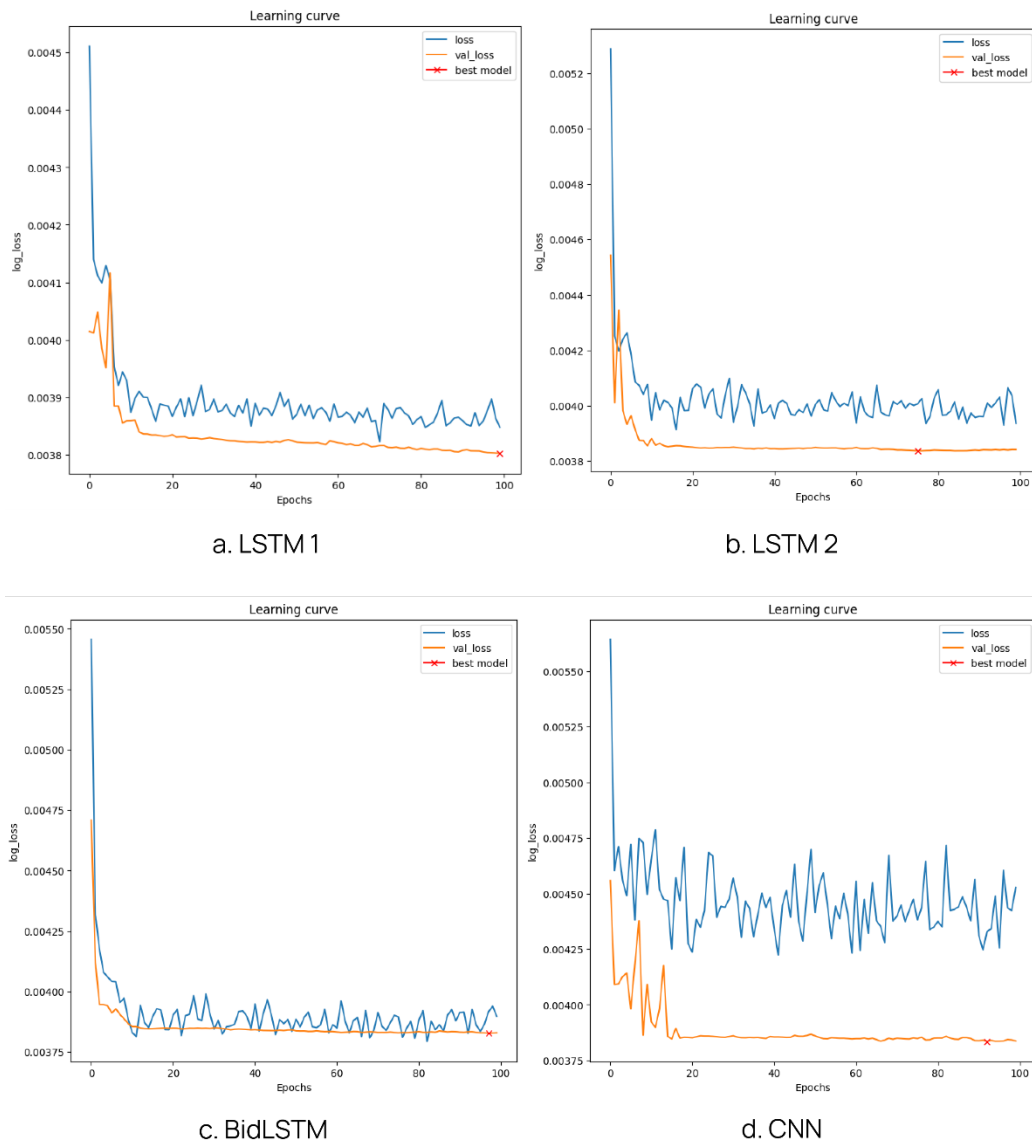


Figure 11: Learning progress of the Deep learning models

	LSTM 1	LSTM 2	BID LSTM	CNN	RANDOM FOREST	DECISION TREE	XGBOOST
TRAINING	0.9376	0.9390	0.9431	0.9357	0.9819	0.9874	0.9945
VALIDATION	0.9373	0.9379	0.9385	0.9322	0.9486	0.9240	0.9440
TEST	0.9524	0.9539	0.9575	0.9503	0.9765	0.9685	0.9793

5. Comparison

To have a better look at the impact of our results, we put a side-by-side view of the R2 metric values that correspond to the different experiments related to the zones Taroudant and Aitmelloul.

LSTM 1	LSTM 2	BID LSTM	CNN	RANDOM FOREST	DECISION TREE	XGBOOST
-----------	-----------	-------------	-----	------------------	------------------	---------

TAROUDANT	0.9294	0.9346	0.9340	0.9201	0.9321	0.9170	0.9137
AOULOZ-TAROUDANT	0.9715	0.9764	0.9754	0.9415	0.9782	0.9754	0.9747

	LSTM 1	LSTM 2	BID LSTM	CNN	RANDOM FOREST	DECISION TREE	XGBOOST
AITMELLOUL	0.8816	0.8905	0.8873	0.8750	0.8960	0.8570	0.8511
TAROUDANT- AITMELLOUL	0.9524	0.9539	0.9575	0.9503	0.9765	0.9685	0.9793

In the two tables, it is clear that the experiment that takes into consideration the two zones at once gave better results compared to the other one. In case of Taroudant, the R2 score increased by 4.36% up to **0.9747** compared to 0.9346. As for Aitmelloul, the results were even higher with an R2 score of 0.9793 with an increase of 8.33%.

From another view, the following is a summary of the experiments where we declare the best model for each experiment along with its R2 score.

EXPERIMENT	MODEL	R2 SCORE
AOULOZ	RF	0.9185
TAROUDANT	LSTM 2	0.9346
AITMELLOUL	RF	0.8960
AOULOZ-TAROUDANT	RF	0.9782
TAROUDANT-AITMELLOUL	XGBoost	0.9793

Conclusion

Based on the conducted experiments, we can extract the following insights:

- Using the total precipitation for run-off prediction has given good results in the testing dataset with an R2 score of above 90% in most cases, and up to 97% in the best case.
- In this study, we have considered the runoff of the zones close to the source of the river as an input feature to the models. Including the geographical context of the area of interest into the prediction task, has increased the performance of the models in predicting the run-off up to 8.33%.
- The explanation for the increase of the performance could be attributed to the fact that runoff is an accumulation of the runoff of previous zones added to the precipitation. However, a simple addition would not give the exact results, hence the impact of various other factors, such as meteorological factors, and/or infiltration.
-

T3.2. Exploitation des données satellites multi-capteurs et multi-résolution

Les données satellites Sentinel 1 et 2 du programme européen COPERNICUS ont été utilisées pour l'estimation de l'évapotranspiration de la culture du blé à travers des relations établies entre le

coefficient cultural (K_c) et des paramètres dérivés du sentinel 1 et 2 tels que le rapport de polarisation (PR), la cohérence interférométrique (ρ) et le NDVI. Pour cela nous avons dérivé les valeurs de K_c à partir des mesures de l'évapotranspiration par le système d'Eddy covariance installé sur deux champs du blé pendant deux saisons agricole. La figure 4 présente l'évolution temporelle du coefficient cultural (K_c) ainsi que les différents paramètres dérivés du Sentinel 1 et 2. Il est clair que l'évolution du PR et de NDVI est similaire à celle du K_c et suit le développement de la culture, alors la cohérence interférométrique (ρ_{vv} et ρ_{vh}) varie inversement de ces paramètres. Suite à ce constat, nous avons développé des corrélations entre K_c et les différents autres paramètres (Figure 5). Ces corrélations sont utilisés après pour estimer l'évapotranspiration (ET) par l'équation de la FAO-56 ($ET = K_c * ET_0$) où ET_0 est l'évapotranspiration de référence. La comparaison entre ET simulée et celle mesurée est illustrée par la figure 6.

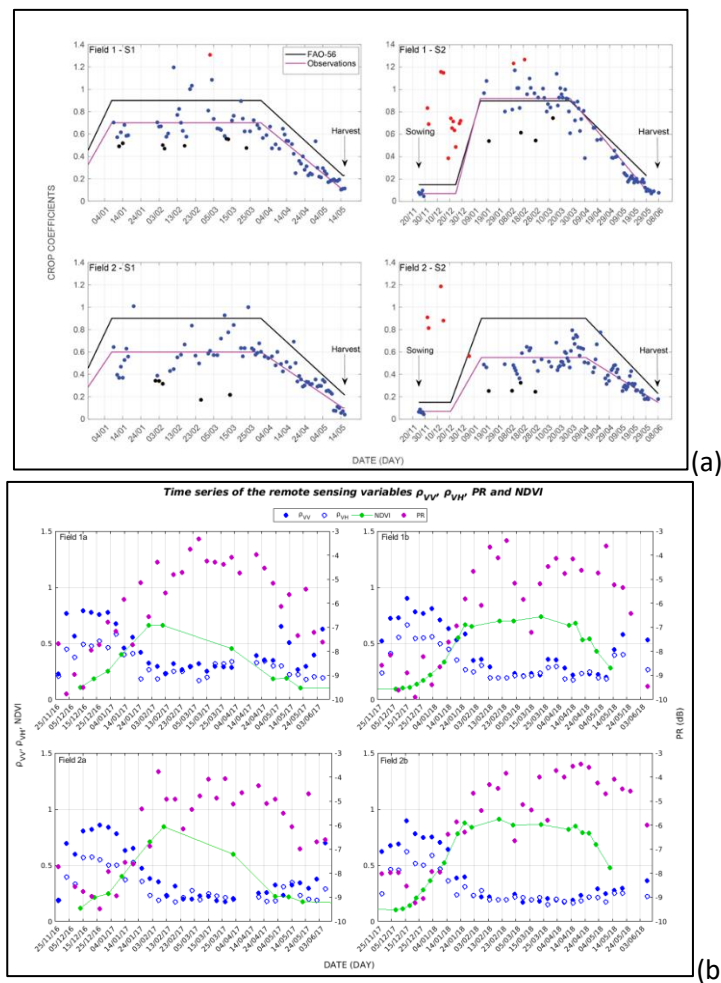


Figure 4. Evolution temporelle du coefficient cultural (a) et des paramètres dérivés du sentinel 1 et 2 (b) ; rapport de polarisation (PR), la cohérence interférométrique (ρ_{vv} et ρ_{vh}) et le NDVI au cours de la saison agricole du blé.

Relationships between K_{cb} and satellite data

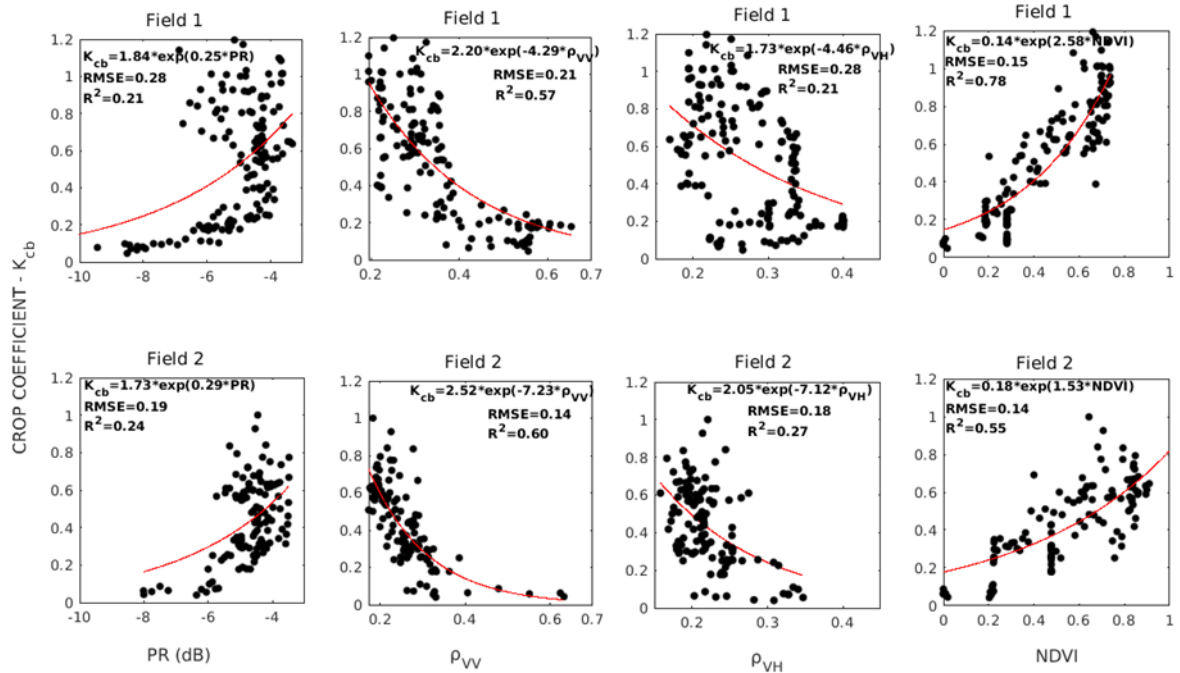


Figure 5. différentes corrélations entre le coefficient cultural et les différents paramètres dérivés du sentinel 1 et 2 ; rapport de polarisation (PR), la cohérence interférométrique (ρ_{VV} et ρ_{VH}) et le NDVI au cours de la saison agricole du blé.

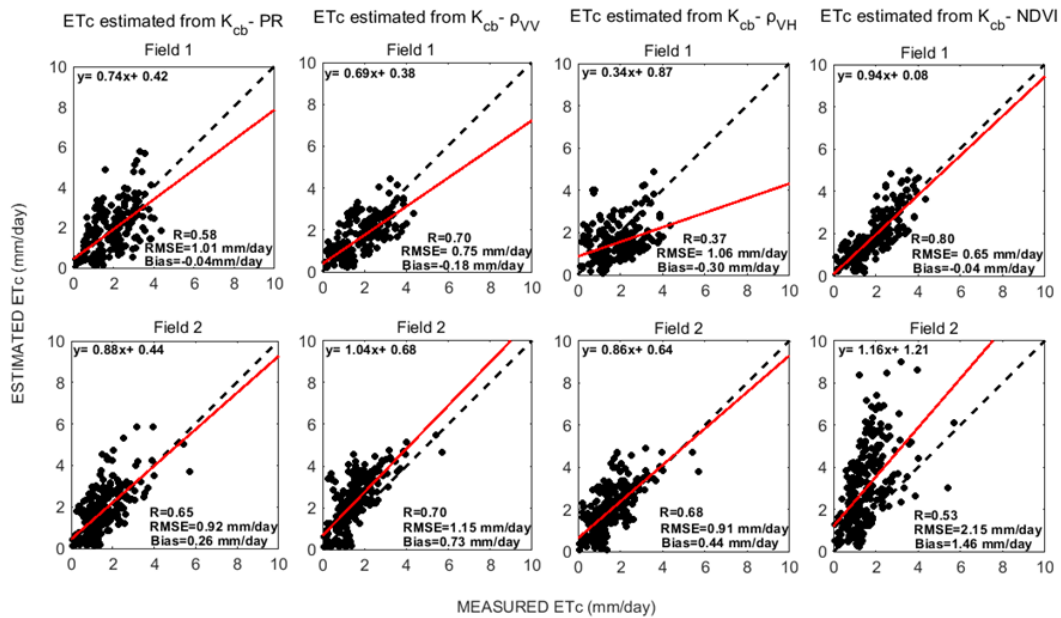


Figure 6. Comparaison entre l'évapotranspiration mesurée et celle estimée à partir de différentes relations, K_c -PR, K_c - ρ_{VV} , K_c - ρ_{VH} and K_c -NDVI.

T3.3. Installation des dispositifs expérimentaux de suivi du continuum Sol-Plante-Atmosphère

Sur la parcelle de référence (F01) de quinoa (voir figure 1) qui a reçu la quantité maximale en irrigation et en fertilisation, nous avons installé un système d'Eddy Covariance (figure 7) pour le suivi des principaux termes des bilans hydrique et énergétique au pas de temps demi-horaire : flux conductifs, radiatifs et convectifs dans le système sol-plante-atmosphère.

Le même dispositif expérimental a été installé sur les deux parcelles du site d'oliviers d'Agdal pour contrôler toutes les composantes du bilan hydrique et thermique notamment la composante évapotranspiration qui est un terme important pour la gestion de l'eau d'irrigation.

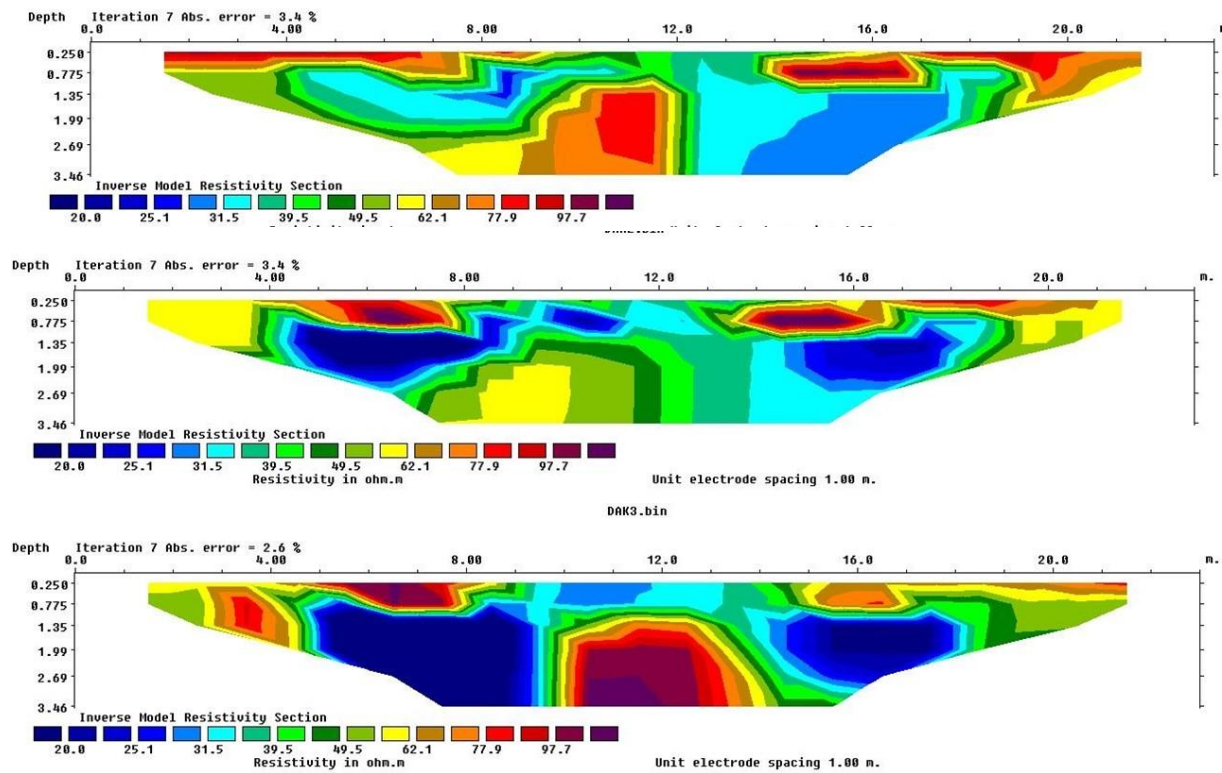
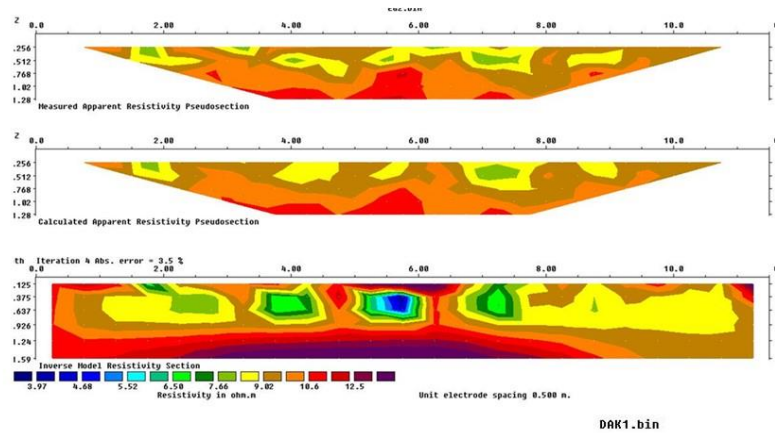


Figure 7: Installation du système d'Eddy covariance pour la mesure de différentes composantes du bilan hydrique.



Figure 8: Installation du système d'Eddy covariance pour la mesure de différentes composantes du bilan hydrique.

Techniques de Tomographie pour l'acquisition des données d'humidité du sol en relation avec l'irrigation



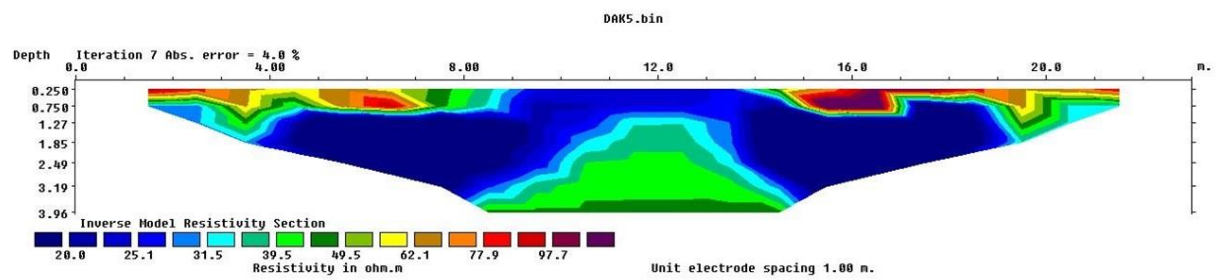
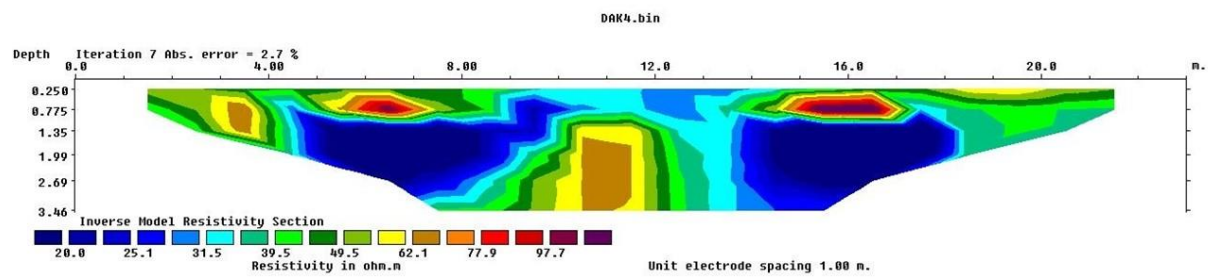


Image drone de la Parcelle expérimentale