

Webinaire "Application des méthodes de machine learning en hydrologie"

Prévision saisonnière du rendement des céréales au Maroc à l'aide de données multi-sources et de méthodes de machine learning

El houssaine Bouras^{1,2}, Lionel Jarlan², Salah Er-Raki^{1,3}, Riad Balaghi⁴, Abdelhakim Amazirh³, Bastien Richard⁵, and Saïd Khabba^{3,6}

1 ProcEDE, Cadi Ayyad University, Marrakech, Morocco.

2 CESBIO, University of Toulouse, IRD/CNRS/UPS/CNES, Toulouse, France.

3 CRSA, University Mohammed VI Polytechnic (UM6P), Benguerir, Morocco.

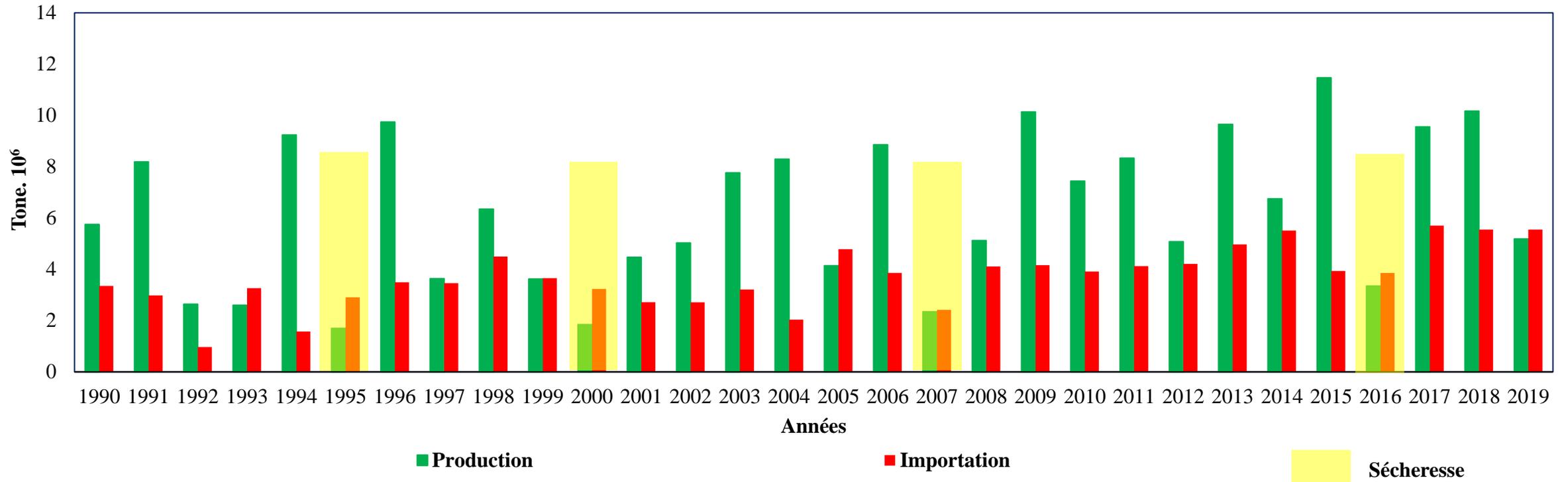
4 National Institute for Agronomic Research (INRA), Rabat, Morocco.

5 CIRAD, IRD, INRAE, Institut Agro, Montpellier, France.

6 LMFE,, Cadi Ayyad University, Marrakech, Morocco

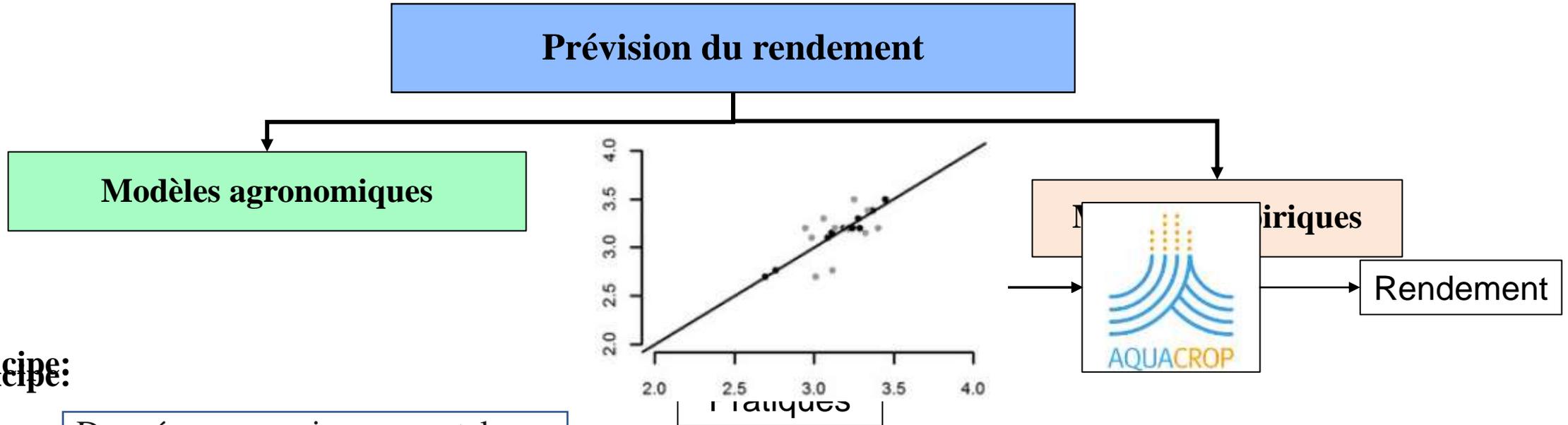
bouras.elhoussaine@gmail.com

Production et importation des céréales au Maroc (1990-2019)

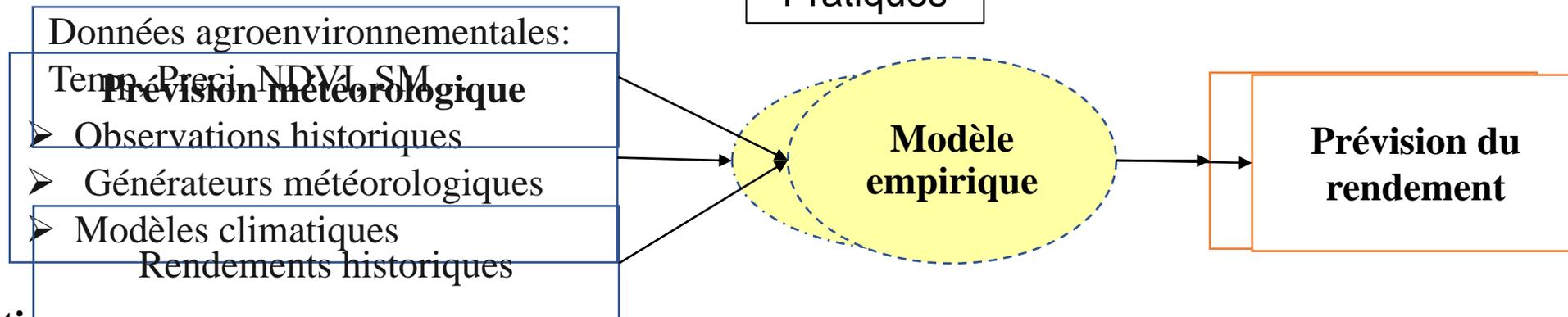


- Forte variabilité interannuelle de la production des céréales.
- Importations : environ 3.5 millions de tonnes (ONICL, 2021).
- Changement climatique (CC) → Plus forte variabilité de la production → Instabilité du marché agricole mondial (Lobell et al 2008).

Prévision saisonnière du rendement des cultures



Principe:

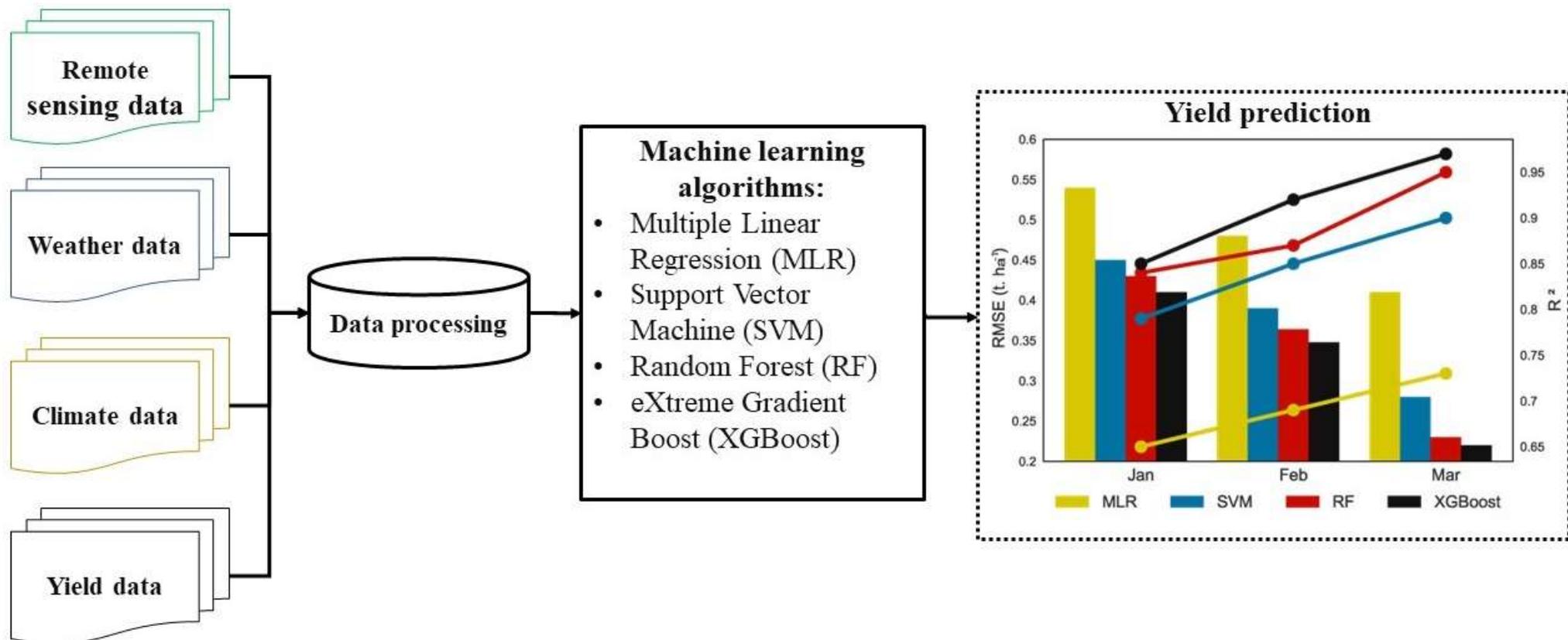


Limitations :

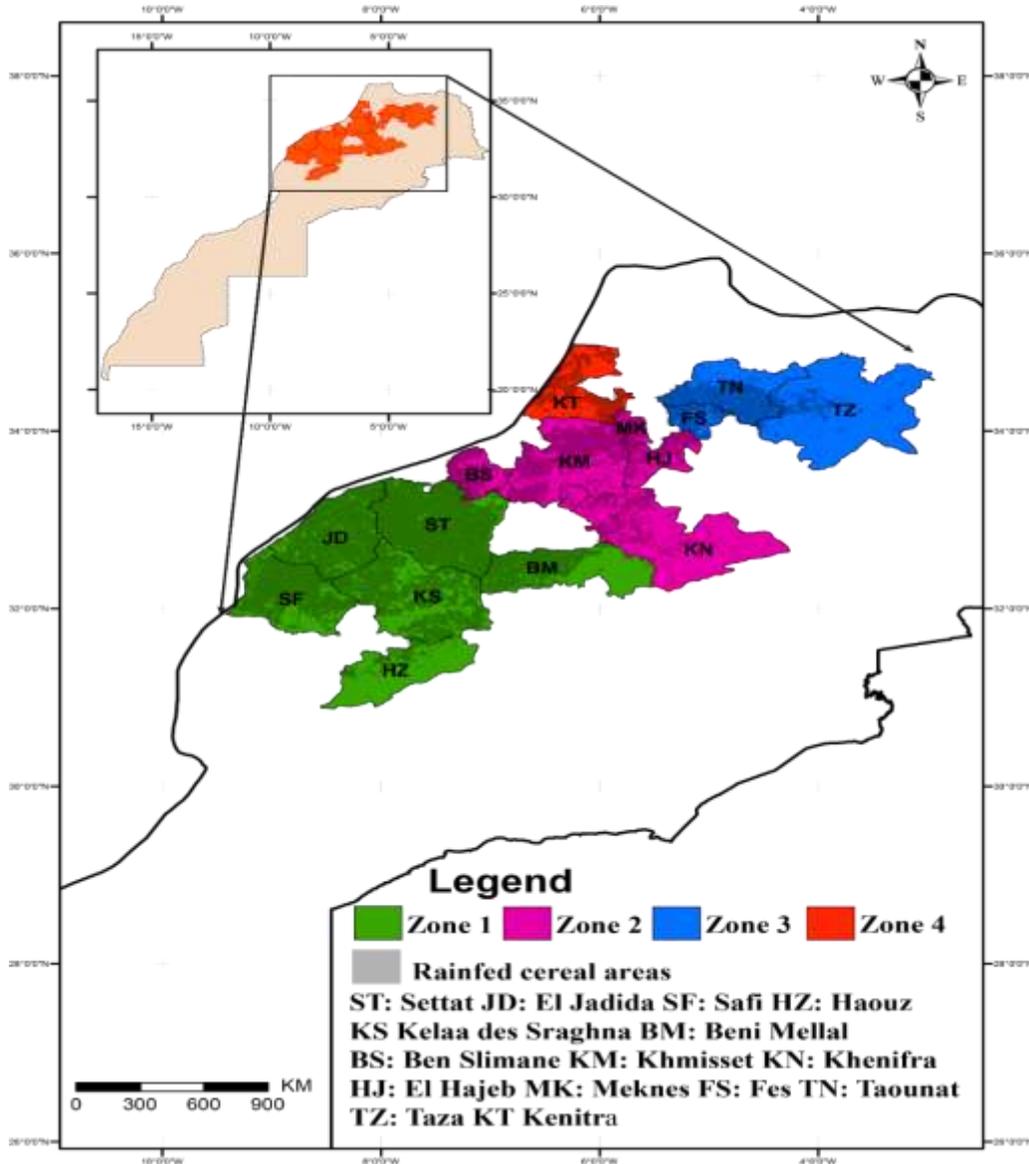
- Nombre important de données d'entrée.
- Incertitude liée aux prévisions météorologiques.
- Quantité et qualité des données
- Utilisé dans un contexte de prévision opérationnelle des rendements

Objectif

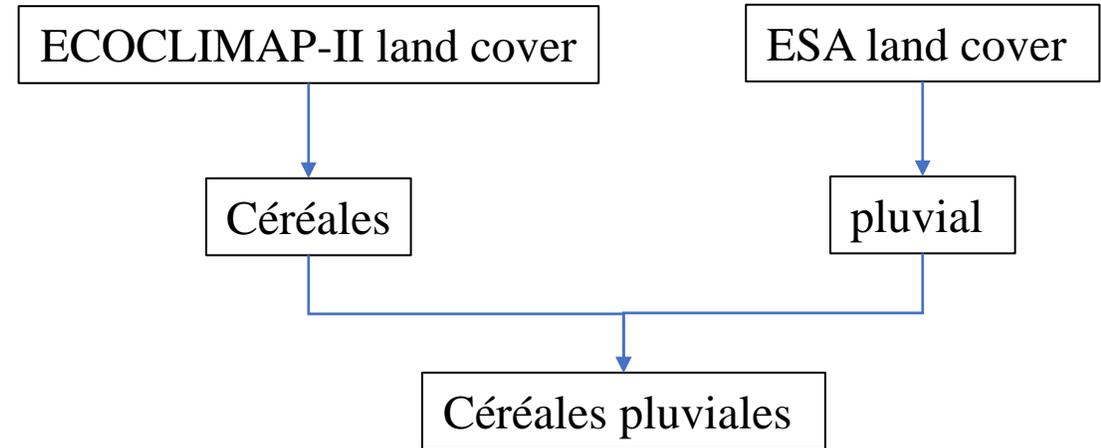
Développer des modèles empiriques pour la prévision précoce du rendement des céréales en utilisant des données multi-sources et le Machine Learning.



☐ Zone d'étude



☐ Identification des zones céréalières pluviales



☐ Cadre général d'étude

- Période d'étude 2000-2017.
- 15 provinces → +90% de la production nationale.

Statistiques sur le rendement des céréales, les précipitations moyennes et la température par province

	Province	Rendement (t/ha)	Cumul des précipitations (mm)	Température (°C)
Zone 1	Settat (ST)	1.3	420.3	15.1
	El Jadida (JD)	1.9	419.5	15.6
	Beni Mellal (BM)	1.5	555.1	11.8
	Kelaa Sraghna (KS)	1.0	340	15.9
	Safi (SF)	0.7	355.2	15.2
	Haouz (HZ)	0.9	381.1	9.9
	Total	1.2	411.9	13.9
Zone 2	Ben Slimane (BS)	1.9	518	14.7
	Khemisset (KM)	1.5	589.9	13.7
	Meknes (MK)	2.2	671.7	14.6
	El Hajeb (HJ)	2.1	688.7	12.2
	Khenifra (KN)	1.8	525.6	9.7
	Total	1.9	598.78	12.98
Zone 3	Fes (FS)	1.7	667.1	14.8
	Taounat (TN)	1.5	900.9	13.7
	Taza (TZ)	1.3	471.9	12.4
	Total	1.5	680.0	13.6
Zone 4	Kenitra (KT)	2.1	748.6	15.4

➤ Les rendements sont faibles par rapport au rendement optimal dans toutes les provinces (7 t/ha)

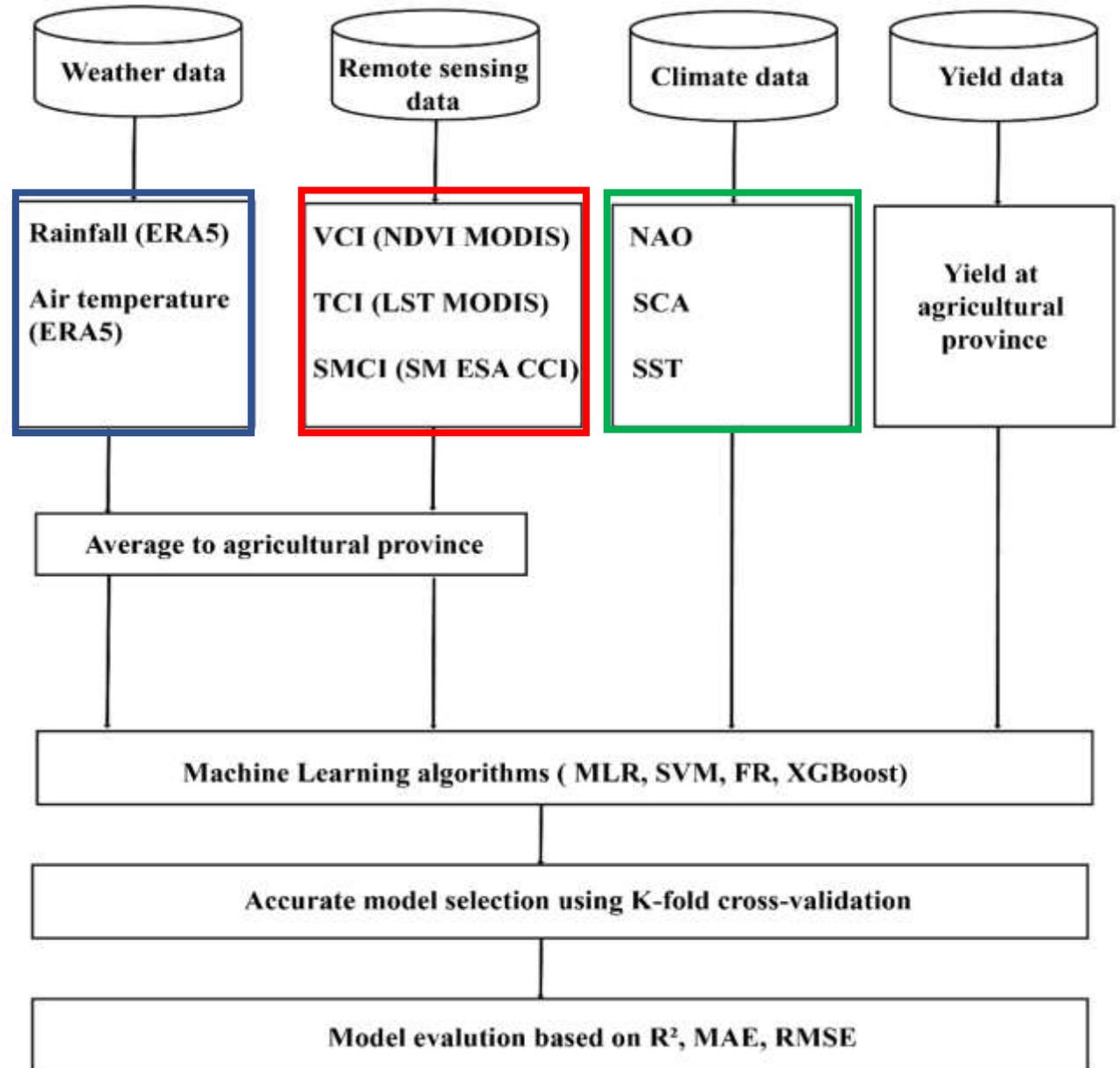
- Les statistiques officielles sur les rendements sont fournies par la Direction de la Stratégie et des Statistiques (DSS)
- Précipitations et Temperature (ERA5)

Prédicteurs du rendement et période de l'année où ces variables sont prises en compte

	Prédicteurs	Période de l'année	Publication
Données Météo	Temperature de l'air	Décembre-Janvier	Balaghi et al., 2008
	Pluie	Octobre-Novembre et Janvier-Mars	
Indices Climatiques	Oscillation Nord Atlantique (NAO)	Décembre	Jarlan et al., 2014
	Scandinavian pattern (SCA)	Janvier	
	Atlantic Tripole	Février	
	Atlantic Niño	Octobre	
Indices de sécheresse par télédétection	VCI	Février-Avril	Bouras et al., 2020
	TCI	Janvier-Mars	
	SMCI	Octobre-Novembre	

Matériels et Méthodes

- Modèle linéaire (MLR)
- Trois modèles non-linéaires: SVM, RF et XGBoost.



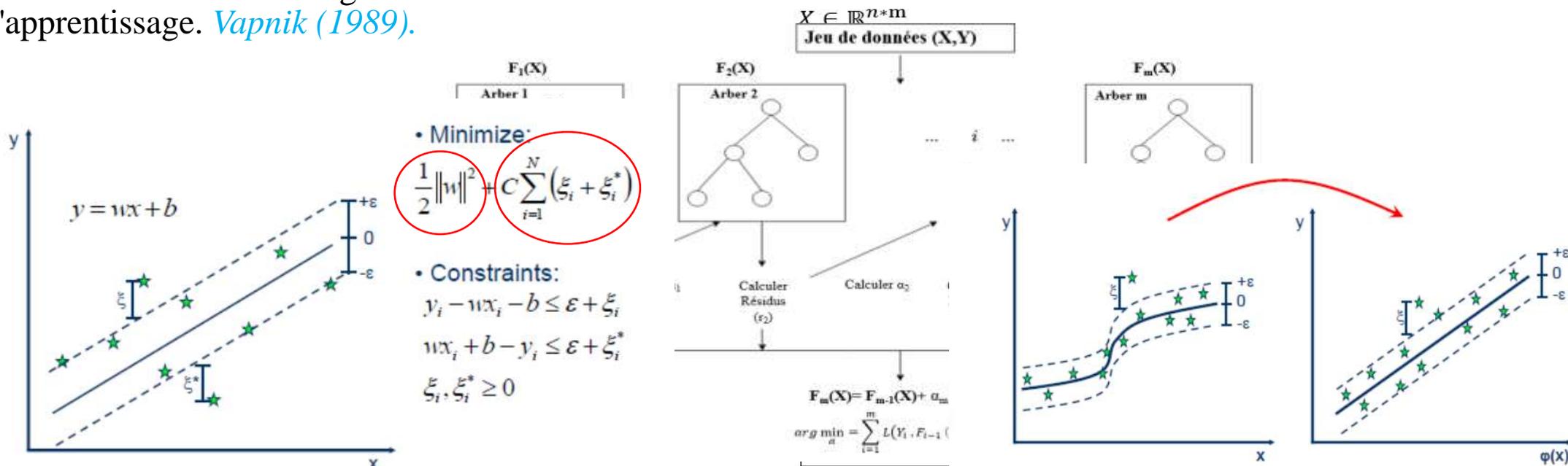
Les algorithmes de Machine Learning

Support Vector Machine (SVM)

Ensemble des arbres de régression constants en série d'apprentissage. *Vapnik (1989)*.

Fonction noyau

Projection des données d'origine dans un nouvel espace à haute dimension



Expérience 1

- Identification des meilleurs prédicteurs (Indices de sécheresse, météo et climat)

Expérience 2

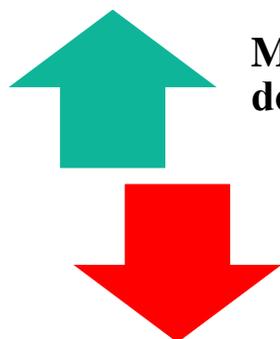
- Evaluation de la performance des modèles en fonction de la précocité (Janvier, Février et Mars)

Expérience 3

- Test de la performance pratique des modèles

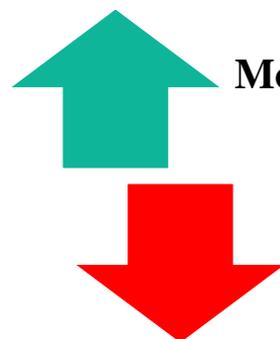
Identification de la meilleure combinaison des données d'entrée

Inputs data	Modèle	RMSE (t. ha ⁻¹)	MAE (t. ha ⁻¹)	R ²
Indices de sécheresse de télédétection	MLR	0.66	0.57	0.67
	SVM	0.54	0.43	0.78
	RF	0.46	0.35	0.80
	XGBoost	0.45	0.34	0.81
Indices de sécheresse de télédétection et données météo.	MLR	0.46	0.39	0.72
	SVM	0.40	0.31	0.80
	RF	0.34	0.24	0.84
	XGBoost	0.37	0.25	0.86
Indices de sécheresse de télédétection, données météo et indices climatiques	MLR	0.41	0.31	0.75
	SVM	0.25	0.21	0.88
	RF	0.22	0.19	0.92
	XGBoost	0.20	0.16	0.95



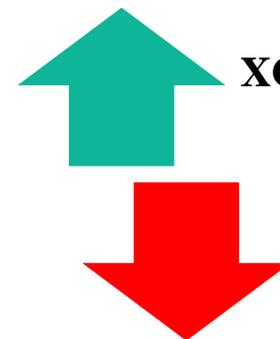
Modèles avec multi données d'entrée

Modèles avec une seule donnée



Modèles non linéaires

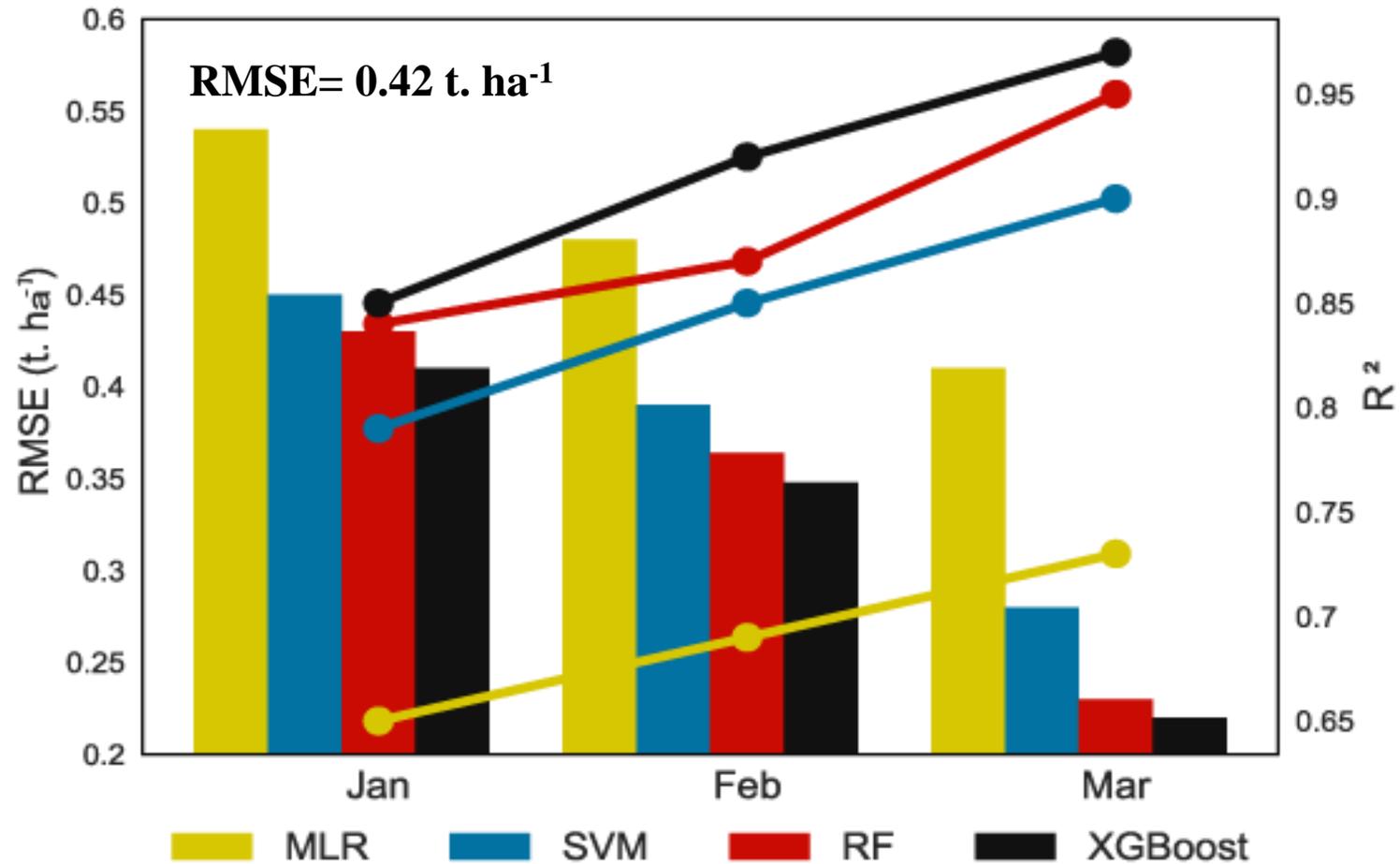
Modèle linéaire



XGBoost

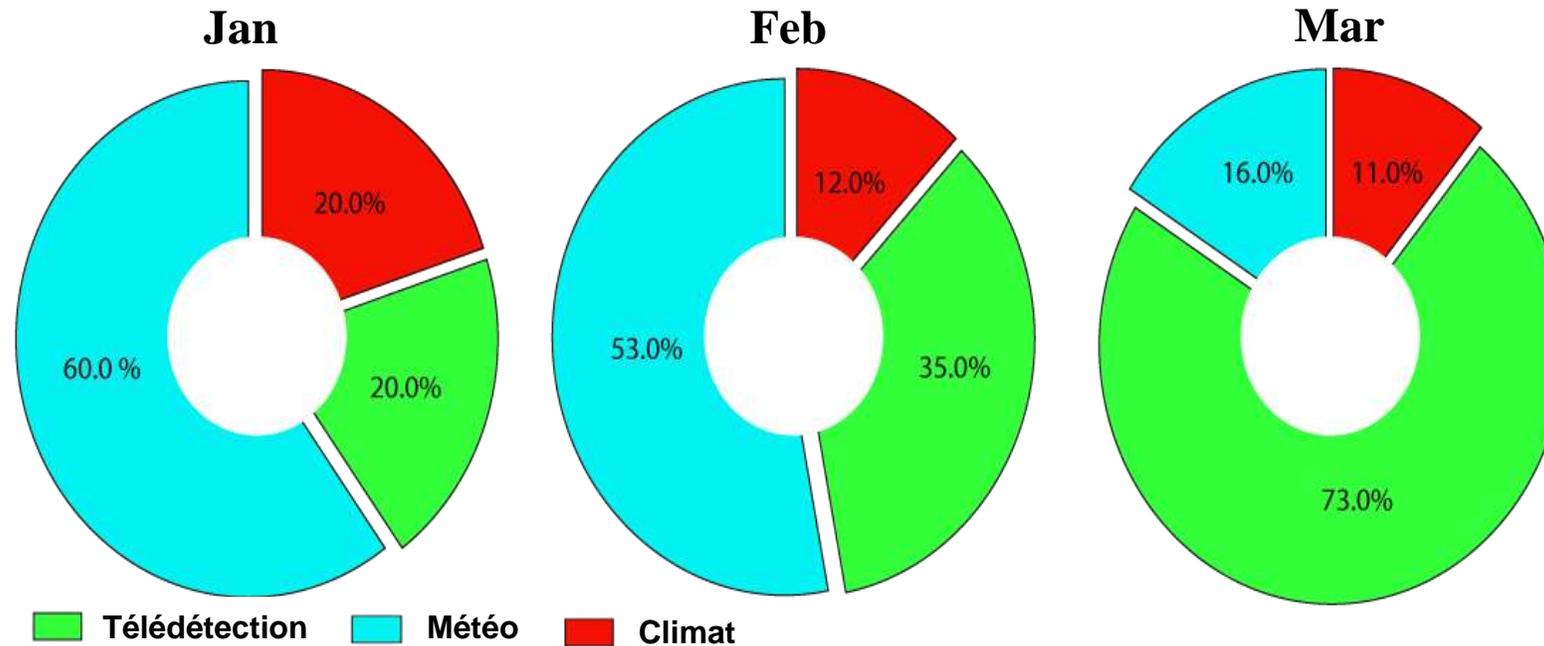
RF, SVM, MLR

Métriques statistiques en fonction de la précocité de prévision



- Augmentation de R^2 de **35 %** de Janvier à Mars.
- Diminution de RSME de **44 %** de Janvier à Mars.

Importance des variables

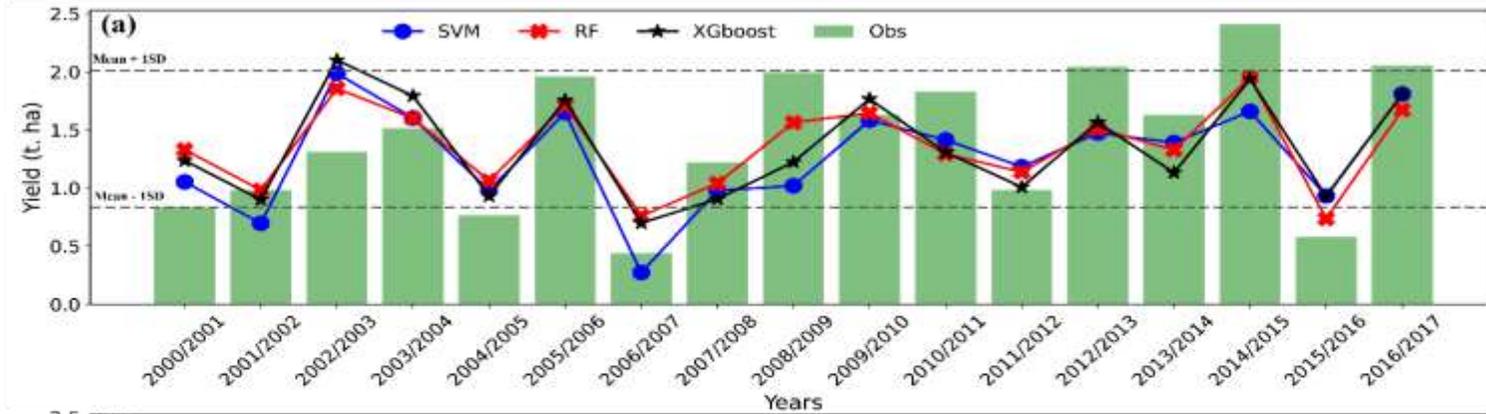


- Les données météorologiques dominent en Janvier (Pluie début de la saison)
- Les données météorologiques et les indices climatiques sont les prédictes clés pour la prévision précoce (80%)
- Les indices de sécheresse de télédétection dominent en Mars (Forte corrélation entre NDVI et le Rendement)

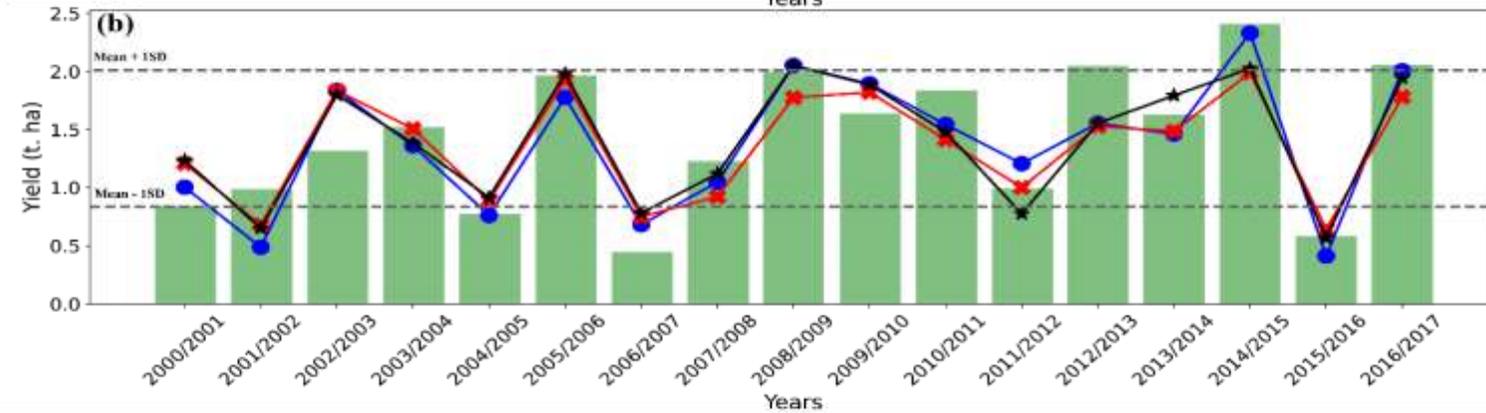
Performance pratique des modèles développés

- La tendance du rendement est bien simulée.
- 88% de la variabilité du rendement à l'échelle nationale en Janvier.

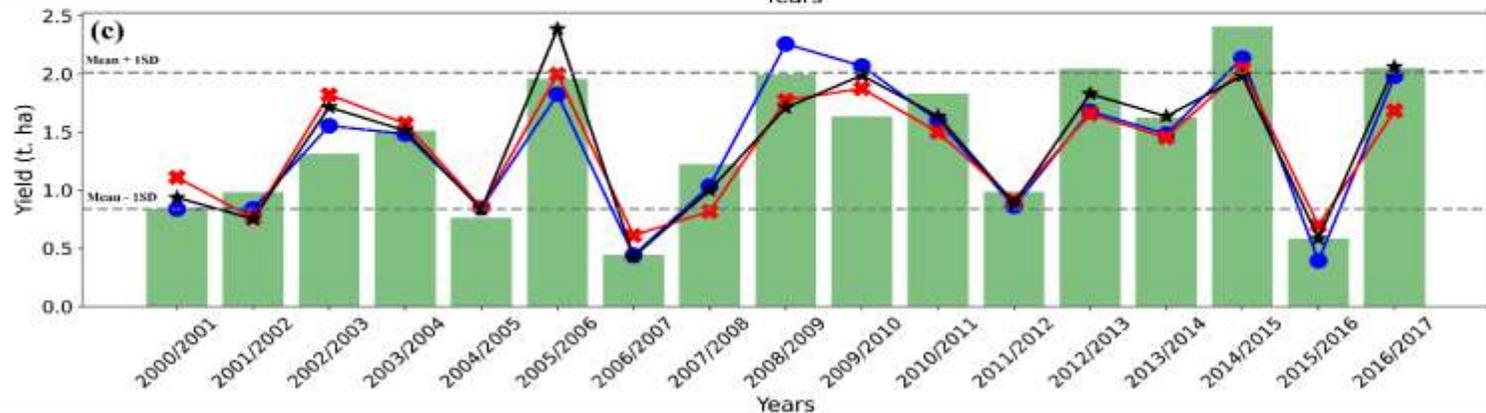
Jan



Feb



Mar



- La combinaison entre les données multi-sources (indices de sècheresse par télédétection, données météo et indices du climatiques) permet d'obtenir des meilleures performances de prévision du rendement (**$R^2 = 0.88$ en janvier**).
- Les données météorologiques et les indices climatiques sont les prédicteurs clés pour la prévision précoce.
- L'intégration des données multi-sources dans des algorithmes de machine learning → Outil d'aide à la décision pour la planification des importations.

- Prises en compte des autres variables qui influencent le rendement → Pratiques agricoles (date de semis , fertilisation) et type du sol...
- Tester les modèles agronomiques alimentés par les prévisions météorologiques saisonnières.
- Utiliser les modèles hybrides (empiriques + agronomiques) → Prédiction du rendement.



remote sensing



Article

Cereal Yield Forecasting with Satellite Drought-Based Indices, Weather Data and Regional Climate Indices Using Machine Learning in Morocco

El houssaine Bouras ^{1,2,*}, Lionel Jarlan ², Salah Er-Raki ^{1,3}, Riad Balaghi ⁴, Abdelhakim Amazirh ³, Bastien Richard ⁵ and Saïd Khabba ^{3,6}

<https://doi.org/10.3390/rs13163101>

Merci pour votre attention

Webinaire "Application des méthodes de machine learning en hydrologie"

Prévision saisonnière du rendement des céréales au Maroc à l'aide de données multi-sources et de méthodes de machine learning

El houssaine Bouras^{1,2}, Lionel Jarlan², Salah Er-Raki^{1,3}, Riad Balaghi⁴, Abdelhakim Amazirh³, Bastien Richard⁵, and Saïd Khabba^{3,6}

1 ProcEDE, Cadi Ayyad University, Marrakech, Morocco.

2 CESBIO, University of Toulouse, IRD/CNRS/UPS/CNES, Toulouse, France.

3 CRSA, University Mohammed VI Polytechnic (UM6P), Benguerir, Morocco.

4 National Institute for Agronomic Research (INRA), Rabat, Morocco.

5 CIRAD, IRD, INRAE, Institut Agro, Montpellier, France.

6 LMFE,, Cadi Ayyad University, Marrakech, Morocco

bouras.elhoussaine@gmail.com