



Prévision saisonnière du rendement en blé au Maroc à l'aide de données multi-sources et des approches de machine learning

El houssaine Bouras^{1,2}, Lionel Jarlan², Salah Er-Raki^{1,3}, Clément Albergel⁴, Riad Balaghi⁵, and Saïd Khabba^{3,6}

1 ProcEDE, Université Cadi Ayyad, Marrakech, Maroc.

2 CESBIO, Université de Toulouse, IRD/CNRS/UPS/CNES, Toulouse, France.

3 CRSA, Université Mohammed VI Polytechnique (UM6P), Benguerir, Maroc.

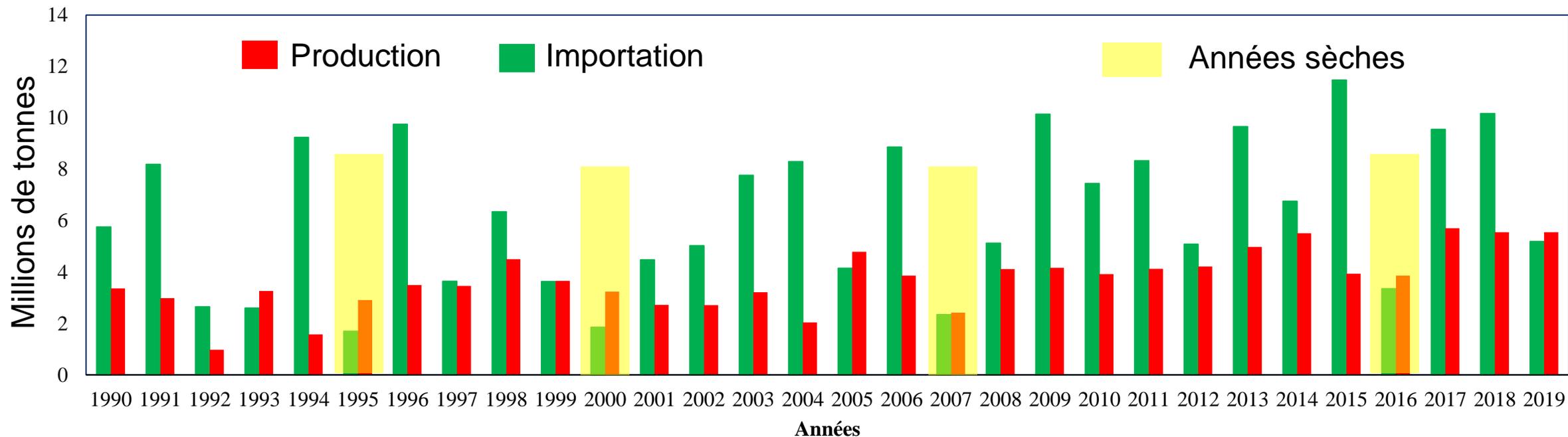
4 CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

5 Institut National de la Recherche Agronomique (INRA), Rabat, Maroc.

6 LMFE, Université Cadi Ayyad, Marrakech, Maroc.

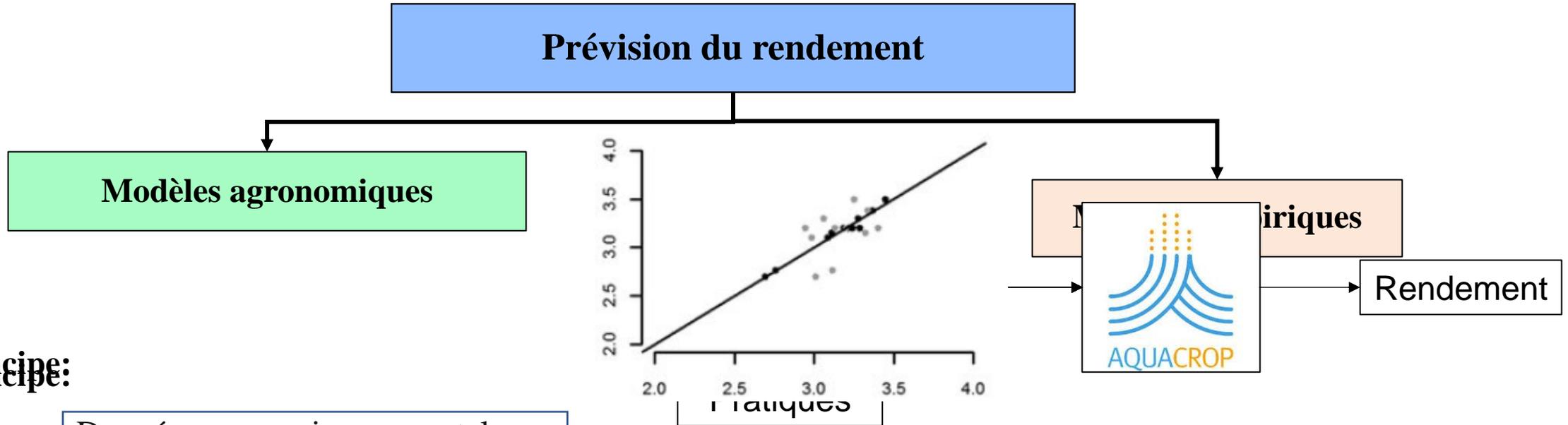


Production et importation des céréales au Maroc (1990-2019)

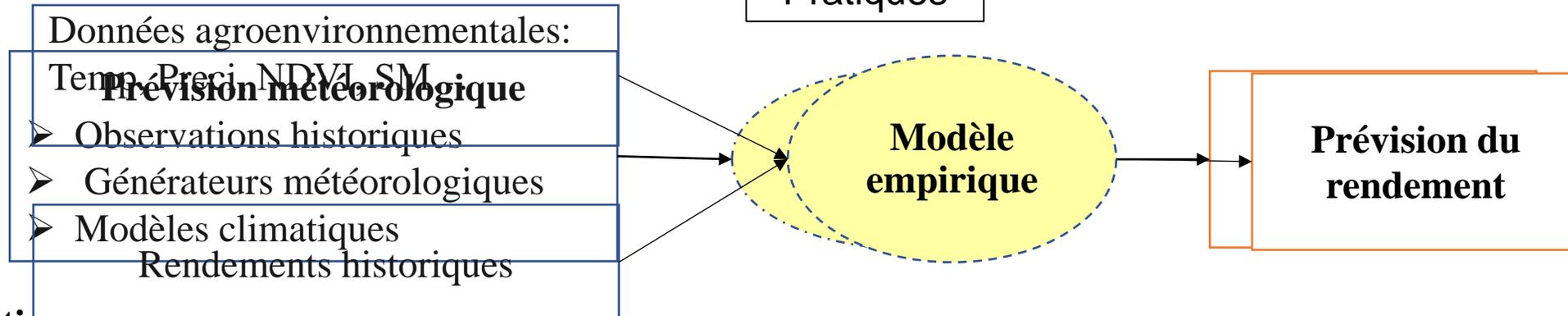


- Production encore dominée par le pluvial (60/40 environ)
- Forte variabilité interannuelle de la production des céréales.
- Importations : environ 3.5 millions de tonnes (ONICL, 2021).
- Changement climatique (CC) → Plus forte variabilité de la production

Prévision saisonnière du rendement agricole



Principe:

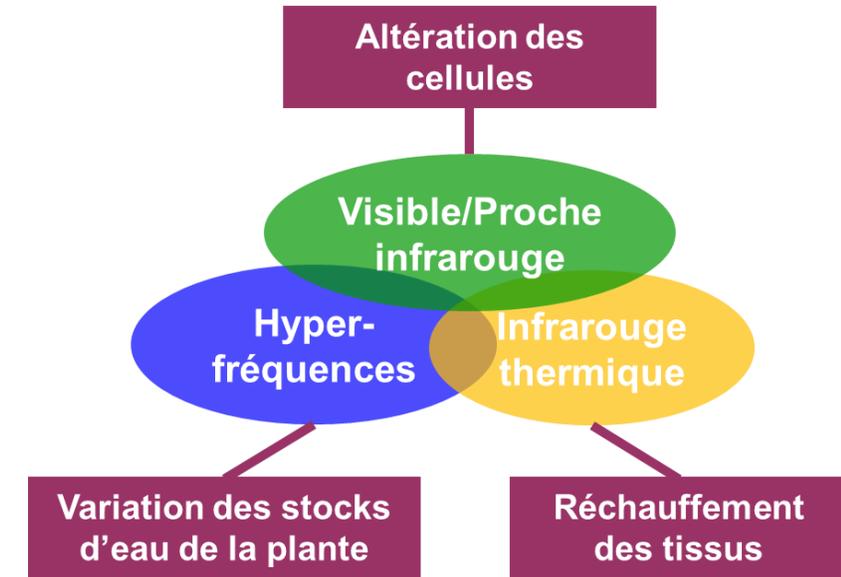


Limitations :

- 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

Les indices de sécheresses dérivés de la télédétection et lien avec le rendement du blé

- Une grande diversité d'indices dérivés des 3 grands domaines de longueur d'onde
- Complémentarité d'information pour caractériser la sécheresse agronomique

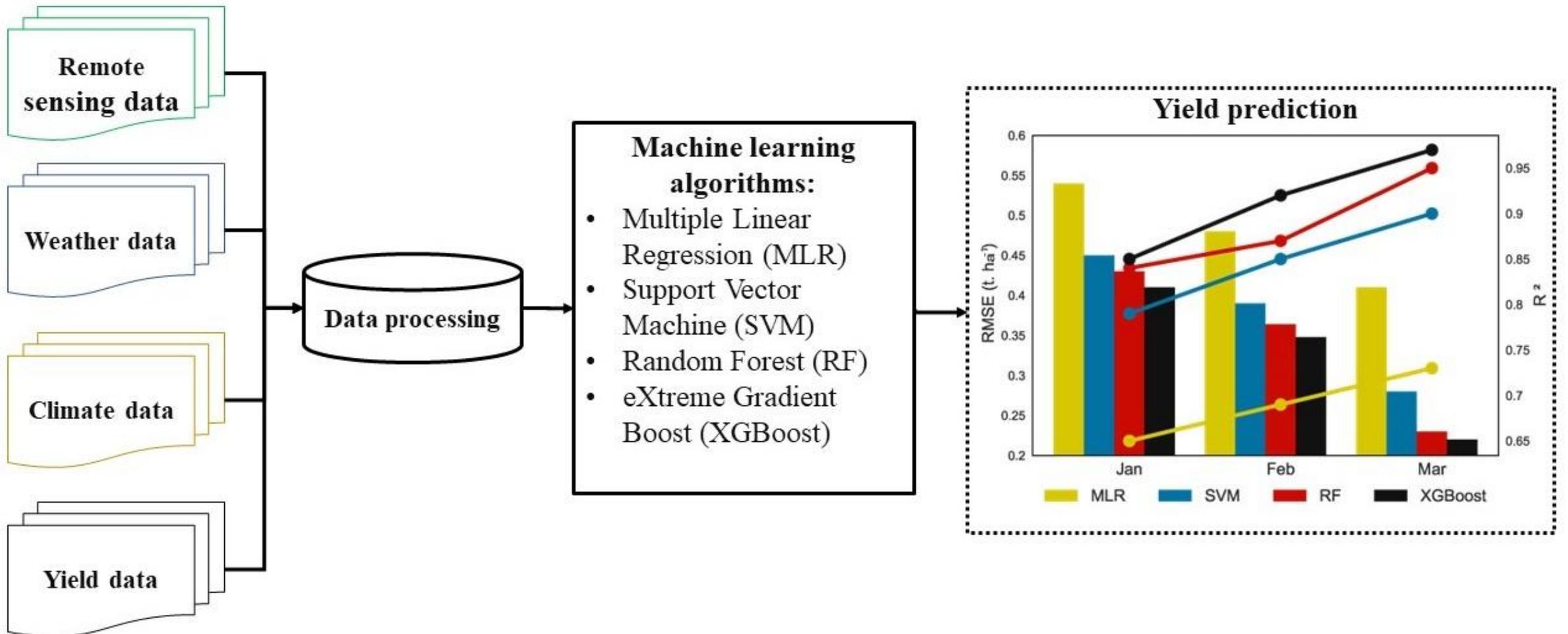


Etude des liens entre indices de sécheresses et rendement du blé au Maroc (Bouras et al., 2019)

	Emergence	Tillering	Elongation	Booting	Anthesis
VCI				++	+++
TCl time		++			
TCl space					++
VHI			++	++	++
SWI	++				

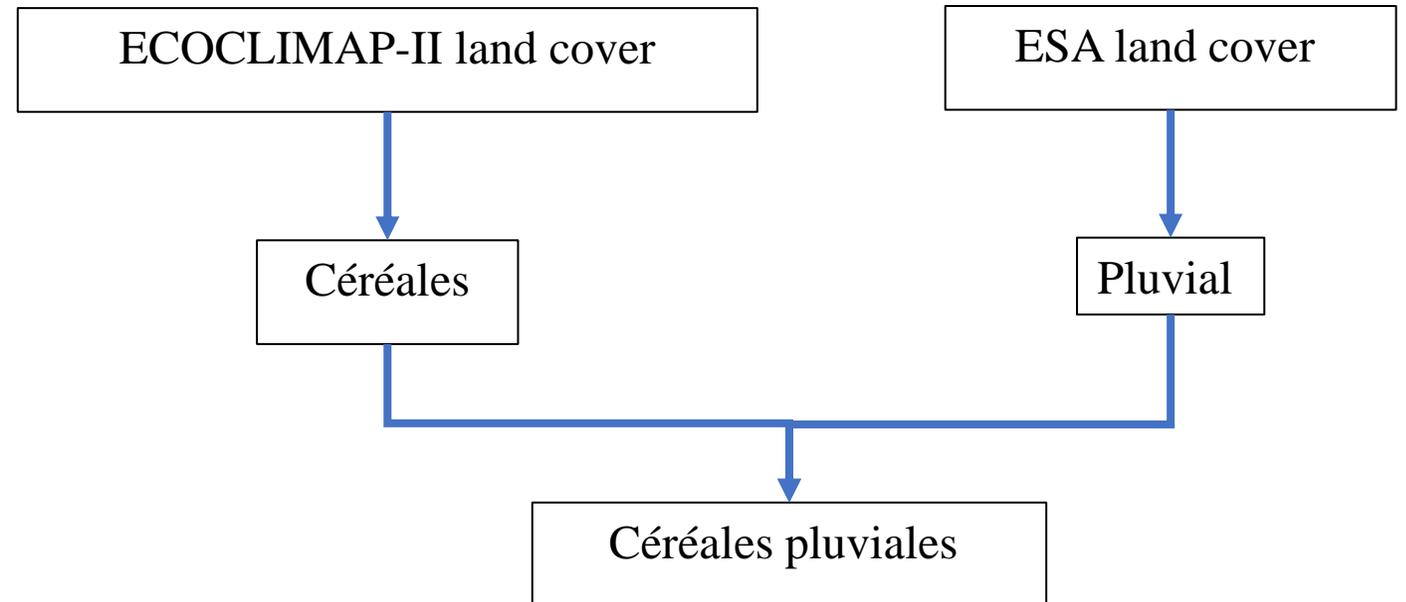
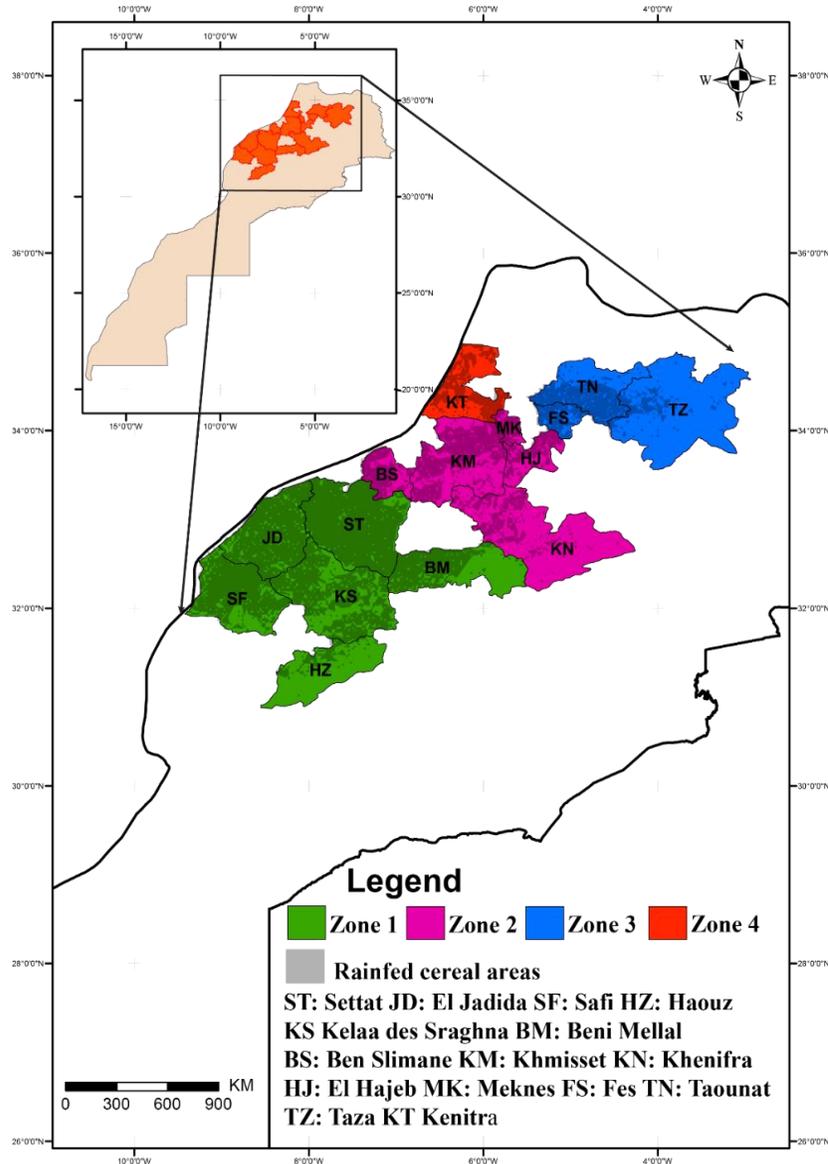
Objectif

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



Région d'étude et données

- Période d'étude : 2000-2017
 - 15 provinces les plus productives: +90% de la production nationale
- Identification des zones de productions de blé pluvial



Les données de rendements en blé sur le Maroc

	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
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

- Données: statistiques officielles de la Direction de la Stratégie et des Statistiques (DSS) Ministère de l'agriculture.

- Précipitations et Temperature : ERA5

➤ Faibles rendements par rapport à l'optimal (7 t/ha)

Les indices de sècheresse dérivés de la télédétection

Vegetation Condition Index

$$VCI = \frac{NDVI_{(i)} - NDVI_{min}}{NDVI_{max} - NDVI_{min}} * 100$$

MODIS MOD13A2
(16-days, 1 km, 2000-2017)

Altération des
cellules

Visible/Proche
infrarouge

Hyper-
fréquences

Infrarouge
thermique

Variation des stocks
d'eau de la plante

Réchauffement
des tissus

$$VHI_i = \alpha VCI + (1 - \alpha) TCI$$

Optimisation

Soil water index

$$SWI(t_n) = \frac{\sum_i^n SSM(t_i) e^{-\frac{t_n - t_i}{T}}}{\sum_i^n e^{-\frac{t_n - t_i}{T}}}$$

ASCAT SWI 10
(Copernicus, 10 day,
0.25°, 2007-2017)

Temperature Condition Index

$$TCI = \frac{LST_{max} - LST_i}{LST_{max} - LST_{min}} * 100$$

MODIS MOD11A1
(daily, 1 km, 2000-2017)

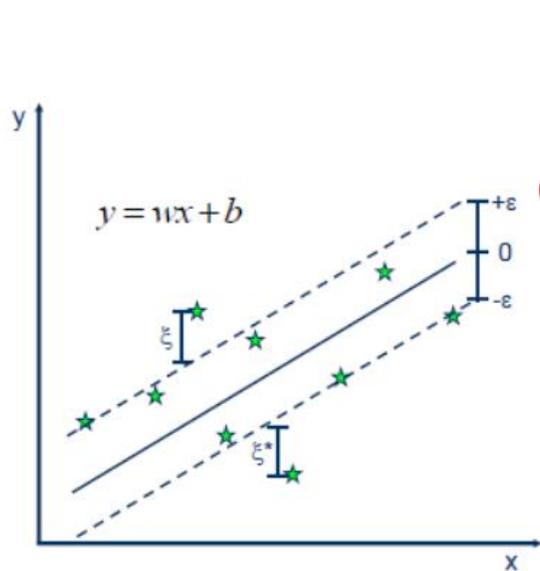
Les prédicteurs potentiels

	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	

Les algorithmes utilisés

Support Vector Machine (SVM)

Trouver une fonction entre la variable dépendante et les prédicteurs en minimisant les erreurs de régression à constantes positives. Ensemble des arbres de régression construits en série d'apprentissage. *Vapnik (1989)*.



• Minimize:

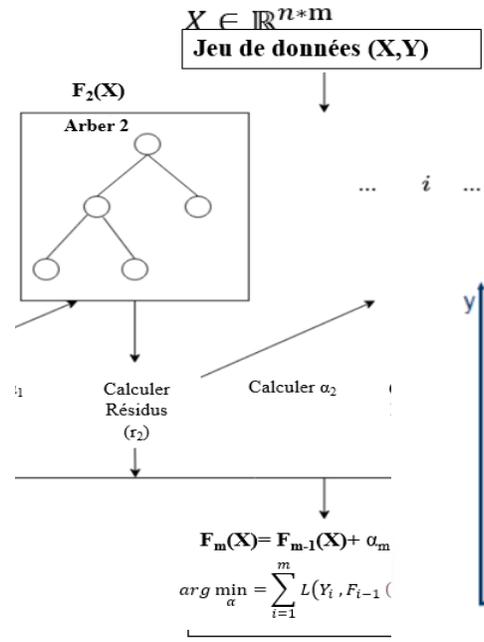
$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*)$$

• Constraints:

$$y_i - wx_i - b \leq \epsilon + \xi_i$$

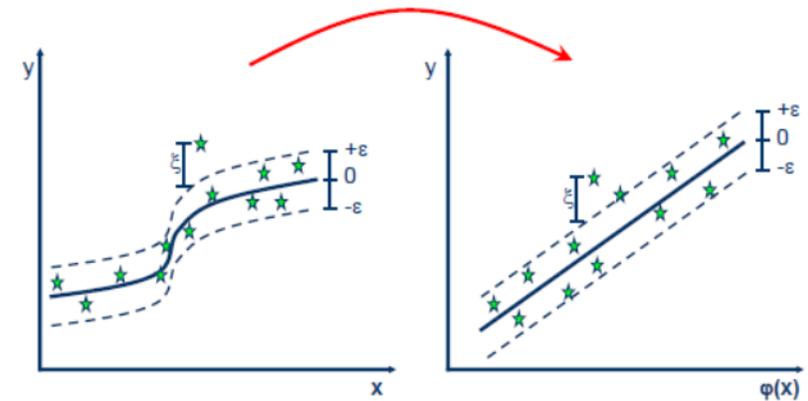
$$wx_i + b - y_i \leq \epsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0$$



Fonction noyau

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



Démarche

Expérience 1

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

Expérience 2

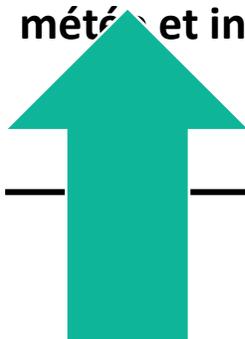
- Evaluation de la performance des modèles en fonction de la précocité (Janvier, Février et Mars)

Expérience 3

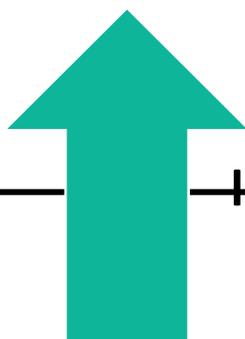
- Test de la performance pratique des modèles

Quels sont les meilleurs prédicteurs ?

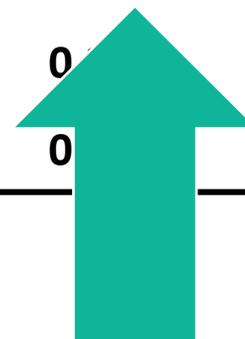
Inputs data	Modèle	RMSE (t. ha ⁻¹)	MAE (t. ha ⁻¹)	R ²
Indices de sécheresse de télédétection	MLR	0.66	0.57	0.67
	SVM	0.54	0.43	0.78
	RF	0.46	0.35	0.80
	XGBoost	0.45	0.34	0.81
Indices de sécheresse de télédétection et données météo.	MLR	0.46	0.39	0.72
	SVM	0.40	0.31	0.80
	RF	0.34	0.24	0.84
	XGBoost	0.37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



Modèles multi jeux de données

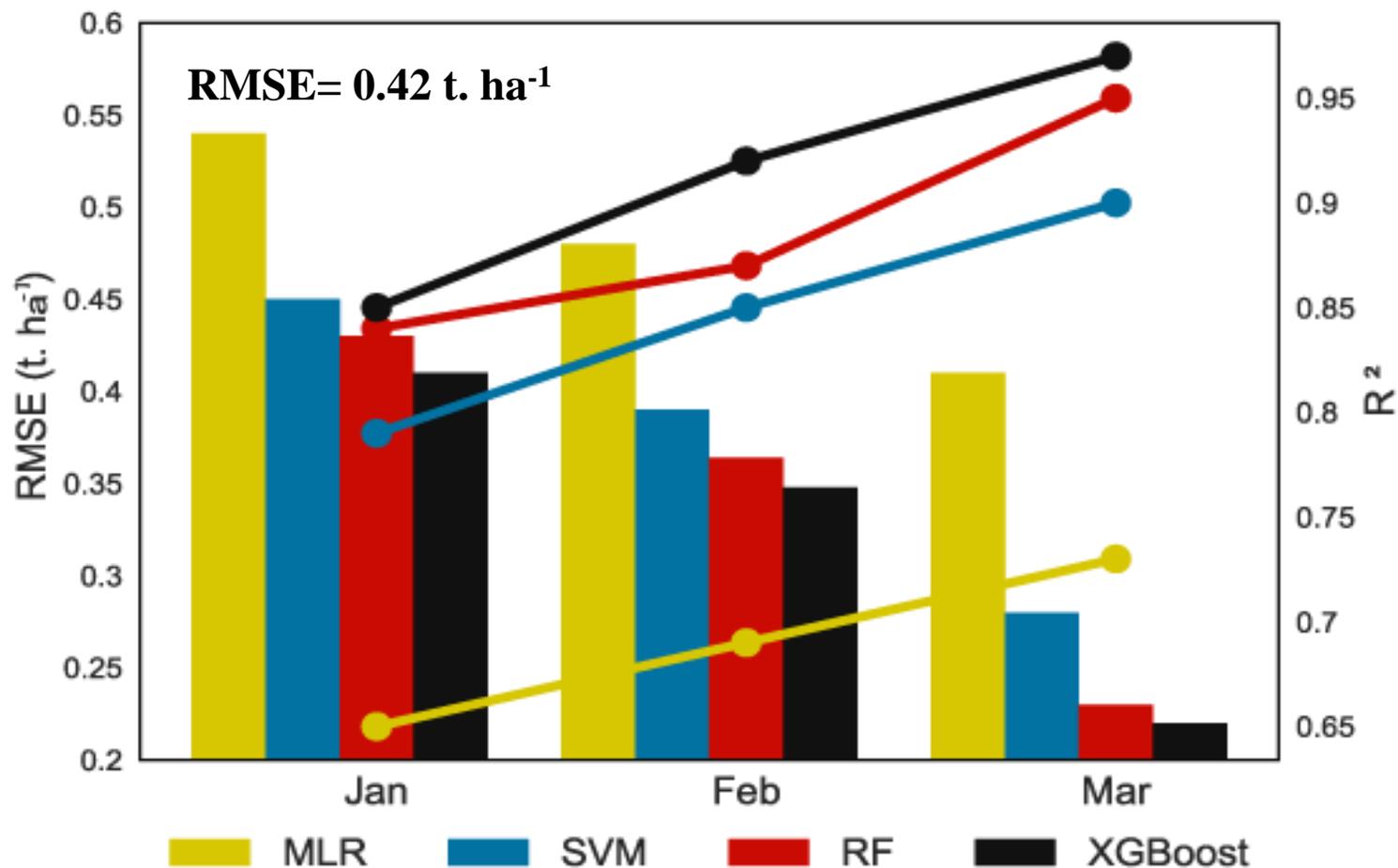


Modèles non linéaires



XGBoost

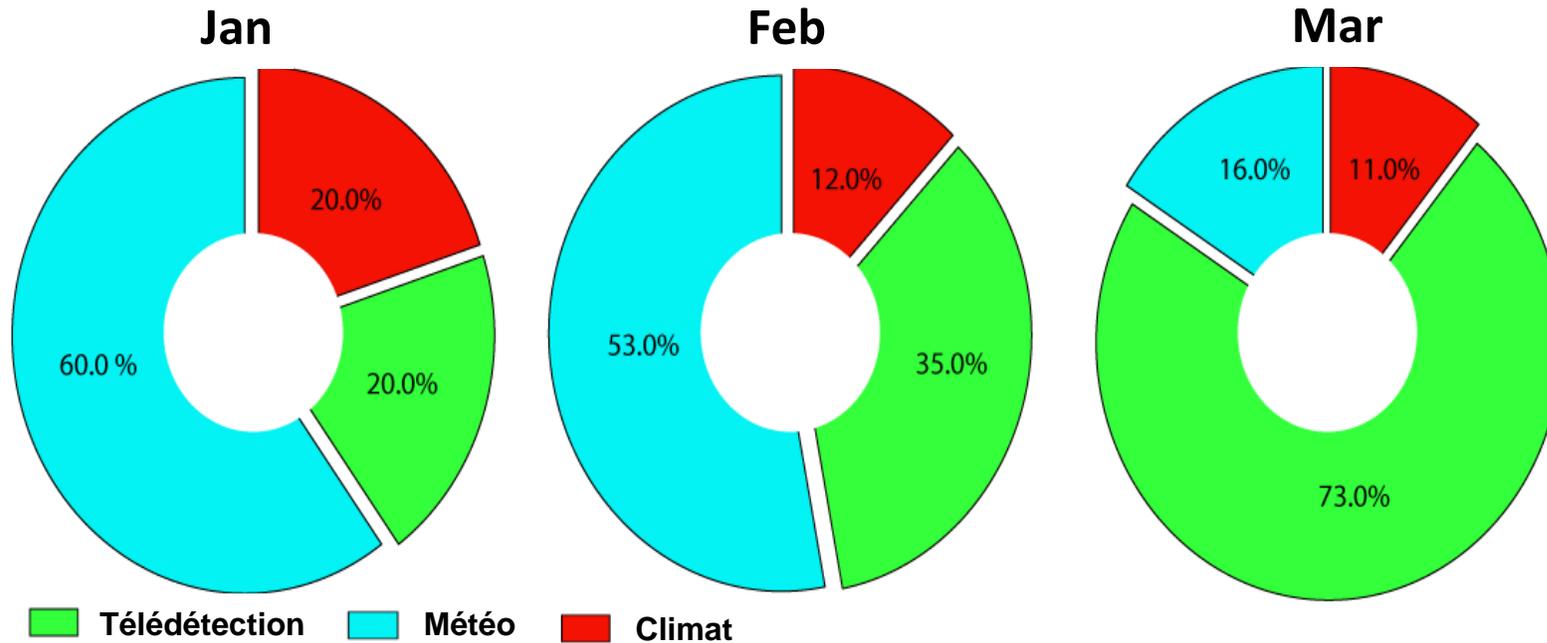
Performance des modèles en fonction de la précocité



Augmentation de R² de **35 %** de Janvier à Mars.

Diminution de RSME de **44 %** de Janvier à Mars.

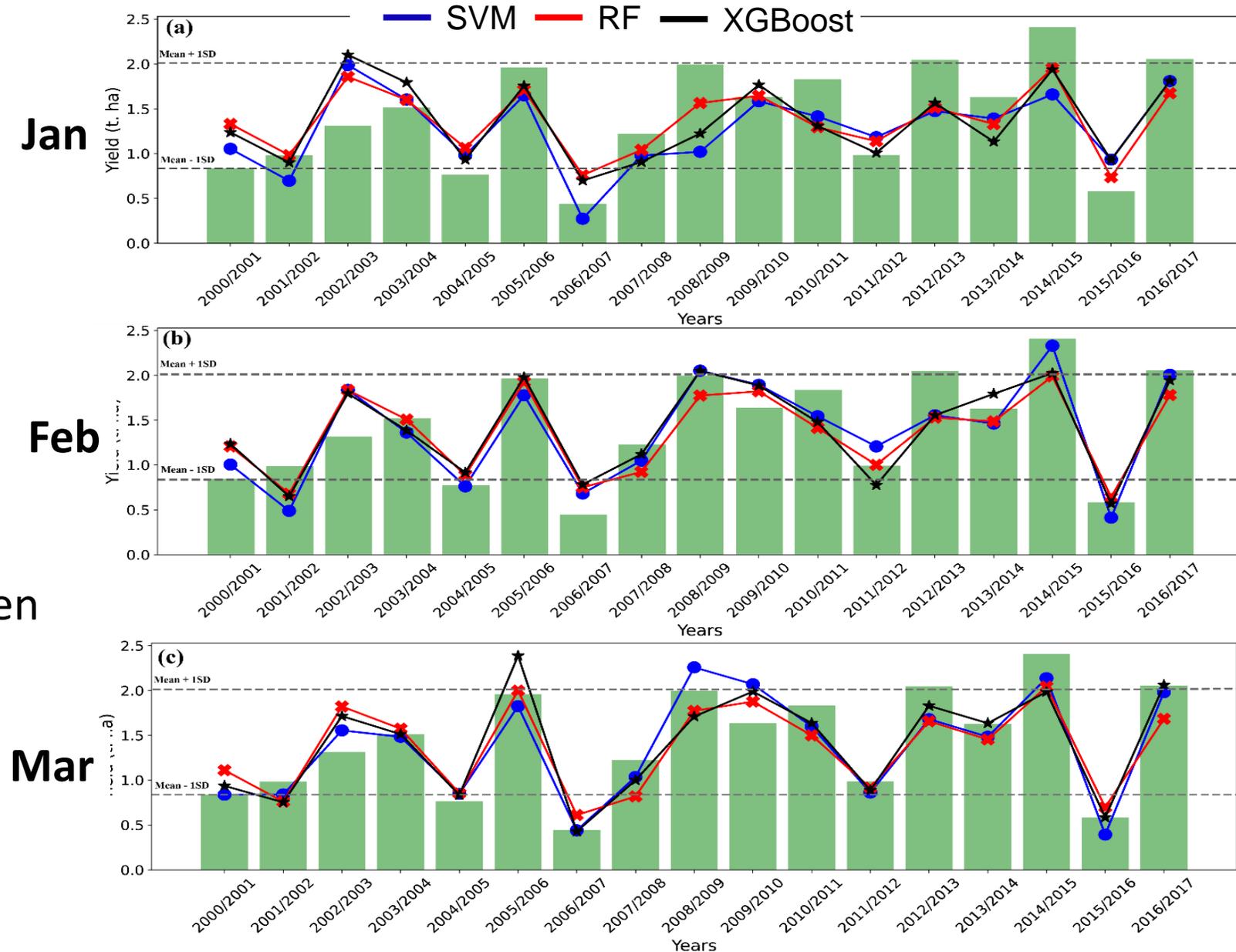
Poids des différents jeux de données en fonction de la précocité de la prévision



- Les données météorologiques dominant en Janvier / Importance des précipitations à l'émergence
- Les données météorologiques et les indices climatiques sont les prédicteurs clés pour la prévision précoce (80%)
- Les indices de sécheresse de télédétection dominant en Mars (forte corrélation entre NDVI et le Rendement)

Performance pratique des modèles

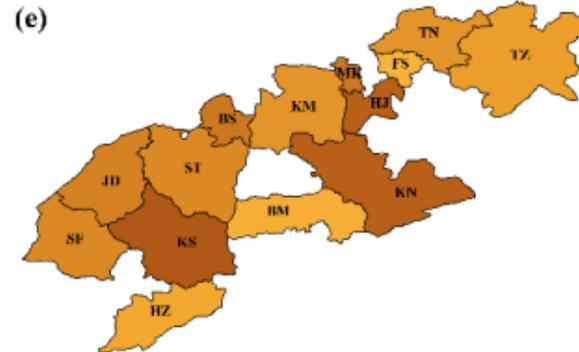
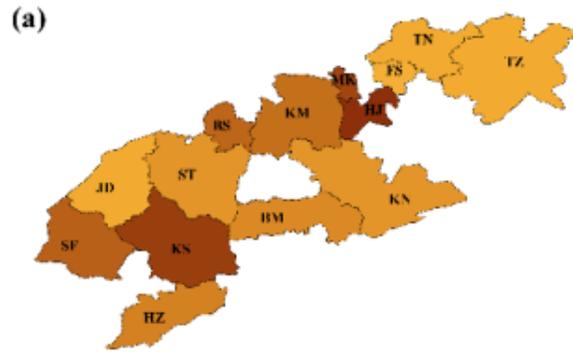
- Méthode leave-one-out
- Tendance du rendement bien simulée
- 88% de la variabilité du rendement est bien reproduite en Janvier.



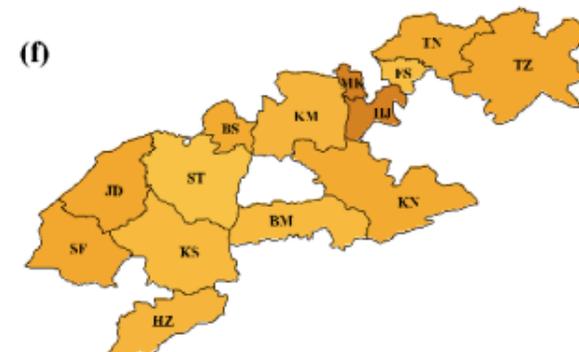
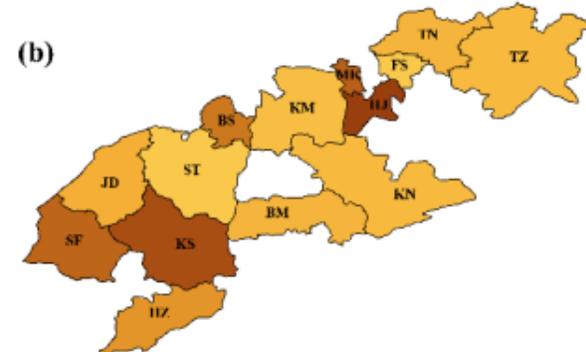
Modèle global versus modèles par provinces

Global

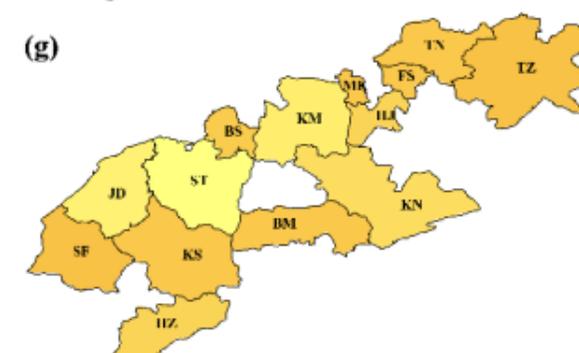
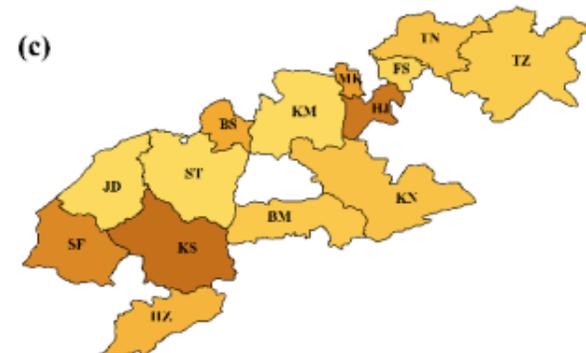
Province



- Meilleures performances des modèles construits par province (amélioration de 4 à 32% des RMSE)



- Meilleures performances sur les provinces céréalières le long de la côte Atlantique (El Jadida JD, Settat ST and Khmisset KM)



- RMSE < 200 kgs/ha pour certains modèles

RMSE (t. ha⁻¹) 0,17 0,35 0,55

Conclusions

- Meilleures performances avec la combinaison de données multi-sources
- Données météorologiques et les indices climatiques importantes pour la prévision très précoce
- Certains modèles atteignent le seuil préconisés de 200 kg/ha en RMSE
- Quelques perspectives d'amélioration
 - Prise en compte des types de sol et des pratiques (date de semis)
 - Ajuster la période d'intégration des différents prédicteurs sur les stades phénologiques
 - Considérer d'autres observations (radar, fluorescence ...etc...)
 - Descente d'échelle possible avec les données de la constellation copernicus

Bouras E., Jarlan L., Er-Raki S., Albergel C., Richard B., Balaghi R., Khabba S., 2020, Linkages between rainfed cereal production and agricultural drought through remote sensing indices and a Land Data Assimilation System: a case study in Morocco., *Remote Sensing*, 12(24), 4018; <https://doi.org/10.3390/rs12244018>.

El Houssaine Bouras, Lionel Jarlan, Salah Er-Raki, Riad Balaghi , Abdelhakim Amazirh, Bastien Richard and Saïd Khabba, 2021, Cereal Yield Forecasting in Morocco Combining Satellite Drought-Based Indices, Weather Data and Regional Climate Indices Using Machine Learning, *Remote Sens.* 2021, 13(16), 3101; <https://doi.org/10.3390/rs13163101>.