

الجمهورية الجزائرية الديمقراطية الشعبية  
وزارة التعليم العالي والبحث العلمي



Département des Sciences Agronomiques

N°...../SNV/2024

## THÈSE

Présentée par

**Boulaaras Hakima**

Pour l'obtention du diplôme de

**DOCTORAT 3<sup>ème</sup> CYCLE**

**Filière : Sciences Agronomiques**

**Spécialité : Amélioration de la Production Végétale**

THÈME

**Application de l'image numérique dans l'évaluation  
des rendements des cultures céréalières en milieu  
semi-aride**

Soutenu publiquement le : 13 / 11 / 2024

DEVANT LE JURY

**Président :** M. Hafsi Miloud

Pr. Université Ferhat Abbas Sétif 1

**Directeur :** M. Bouregaa Tarek

MCA. Université Ferhat Abbas Sétif 1

**Examineur :** M. Mekhlouf Abdelhamid

Pr. Université Ferhat Abbas Sétif 1

**Examineur :** M. Rouabhi Amar

MCA. Université Ferhat Abbas Sétif 1

**Examinatrice :** Me. Labad Ryma

MCA. ENSA El Harrach, Alger

## ***Dédicaces***

***À mes chers parents, pour leur amour infini, leur soutien et leur sacrifice sans limites tout au long de ce parcours académique.***

***À mes frères Ahmed, Walid, Mahdi, et Fères.***

***À mes sœurs Noura, Fouzia, et Nawel.***

***À mon mari Mousaab et mon fils Ouais.***

***À mes neveux Mouhamed, Mouad et Mouhamed Djoud.***

***À tous les membres du laboratoire LADPVA***

***A mes amis, sans exception et à tous mes camarades du  
Département  
des Sciences Agronomiques.***

***À l'Algérie.***

***Hakima***

# ***Remerciements***

***Je remercie Dieu tout puissant qui m'a donné la force et la foi pour mener à bien ce modeste travail de recherche***

***Je tiens à remercier:***

***Monsieur Bouregaa Tarek; pour m'avoir encadré et dirigé tout au long de la réalisation de ce travail avec une grande rigueur scientifique, pour sa parfaite disponibilité et ses précieux conseils.***

***M. Hafsi Miloud; professeur à l'Université Ferhat Abbas de Sétif 1, pour m'avoir fait l'honneur de présider le jury de soutenance. Je lui exprime toute ma reconnaissance.***

***M. Mekhlouf Abdelhamid; professeur à l'Université Ferhat Abbas de Sétif 1, de m'avoir fait l'honneur d'accepter de faire partie du jury et d'examiner ce travail.***

***M. Rouabhi Amar; docteur à l'Université Ferhat Abbas de Sétif 1, d'avoir accepté d'examiner ce travail. Qu'il trouve ici mon profond respect et ma parfaite reconnaissance.***

***M<sup>me</sup>. Labeled Ryma; docteur à l'ENSA d'El Harrach, je voudrais qu'elle trouve ici mes remerciements les plus vifs d'avoir voulu examiner mon travail et de faire partie du jury.***

***Je tiens à exprimer mes sincères remerciements au cadre de l'Institut Technique des Grandes Cultures (ITGC) de Sétif : pour leur collaboration précieuse et leur soutien tout au long de cette étude. Leur expertise, leur disponibilité et leur engagement ont été essentiels pour la réalisation de ce travail de recherche.***

# Table des matières

ملخص

Résumé

Abstract

Liste des abréviations

Liste de tableaux

Liste des figures

Liste des annexes

Les travaux dirigés dans le cadre de cette thèse

Introduction

Générale ..... 01

## PARTIE I : Synthèse Bibliographique

### Chapitre I : Aperçu générale sur les céréales (Blé dur, Blé tendre et l'orge)

1. Généralité sur les cereals.....	06
1.1. Définition.....	06
1.2. Historique.....	06
1.3. Taxonomie.....	06
2. Cycle de développement.....	07
2.1. Période végétative.....	07
2.2. Période reproductrice.....	08
3. Exigences des cereals.....	09
3.1. Exigences pédoclimatique.....	09

3.2. Exigences en nutriments.....	10
4. Importance des céréales et leurs distributions.....	12
4.1. Dans le monde.....	12
4.1.1. Production.....	12
4.2. En Algérie.....	15
4.2.1. Répartition géographique.....	15
4.2.2. Place et importance des céréales en Algérie.....	16
4.2.3. Production.....	16
5. Les contraintes de production des céréales en Algérie.....	19

## **Chapitre II : Notions fondamentale sur la télédétection**

1. Méthodes d'évaluation du rendement des céréales.....	23
1.1. Méthodes traditionnelles.....	23
1.2. Méthodes innovantes.....	24
1.2 .1. Télédétection.....	24
1.2.1.1. Définition.....	24
1.2.1.2. Principe de la télédétection.....	25
1.2.1.3. Processus de la télédétection.....	29
1.2.1.4. Vecteur et capteur .....	30
2. Technologie d'imagerie numérique .....	31
2.1. Résolutions d'image numérique.....	31
2.1.1. Spatiale.....	31
2.1.2. Spectrale.....	32

2.1.3. Temporelle.....	32
2.1.4. Radiométrie.....	33
2.2. Types de capteurs d'image.....	33
2.2.1. Photographie aérienne.....	33
2.2.2. Drones.....	33
2.2.3. Satellites.....	34
3. Application de l'imagerie numérique dans évaluation du rendement des cultures.....	37
3.1. Le spectroradiomètre imageur à résolution moyenne ( MODIS).....	37
3.2. Les indices de végétations.....	38
3.2.1. L'indice de végétation par différence normalisée (NDVI).....	39
3.3. Méthodes de prévision des rendements des cultures à partir des images satellitaires...	39
3.3.1. Prévision du rendement des cultures avec la méthode statistique.....	40
3.3.1.1. Les approches linéaires.....	40
3.3.1.2. Les approches non linéaires.....	41

## **PARTIE II : Matériel et Méthodes**

### **Chapitre I : Présentation des zones d'étude**

Introduction.....	43
1. La wilaya de Sétif.....	43
1.1. Cadre géographique.....	43
1.2. Caractéristiques agro-pédologiques.....	44
1.3. Les ressourceshydrauliques.....	45
1.4. Les potentialités agricoles.....	45
1.5. Production végétale.....	45

1.6. Production animale.....	46
1. 7. Caractéristiques climatiques.....	47
1.7.1. Pluviométrie.....	47
1.7.2. Température.....	49
1.7.3. Diagramme ombrothermique.....	50
2. L'échelle communale (commune de Sétif) .....	51
2.1. Situation géographique.....	51
2.2. Caractéristiques climatiques.....	52
2.2.1. Pluviométrie.....	52
2.2.2. Température.....	54
2.2.3. Diagramme ombrothermique.....	55
3. L'échelle parcellaire.....	56
3.1. Situation géographique.....	56
3.2. Caractéristiques physico-chimiques du sol.....	57
3.3. Caractéristiques climatiques.....	58
3.3.1. Pluviométrie.....	58
3.3.2. Température.....	59
3.3.3. Diagramme ombrothermique.....	60
4. Le matériel végétal.....	61
5. Itinéraire technique et conduite de la culture.....	62
5.1. La préparation du sol.....	62
5.2. Fumure de fond.....	63

5.3. Semis.....	63
5.4. Fumure azotée.....	63
5.5. Désherbage chimique.....	63
5.6. Récolte.....	63

## **Chapitre II : Données et Outils utilisés**

1. Objectif.....	64
2. Les produits satellitaires MODIS.....	64
2.1. Choix de la source de données de télédétection.....	64
2.2. L'indice de végétation par différence normalisée ou indice de Tucker (NDVI).....	65
2.3. Données MODIS-NDVI.....	66
3. Calcul de la fraction de couverture verte de la canopée (FGCC).....	66
3.1. L'application "Canopeo".....	66
3.2. Fonctionnement de l'application.....	67
3.3. Calcul du FGCC au niveau des parcelles expérimentales.....	68
4. Estimation de la biomasse.....	69
5. Données agro-météorologiques.....	71

## **Chapitre (III) : Approche de prédiction des rendements céréaliers**

1. Analyse et outils statistique des résultats.....	74
1.1. Analyse par regression linéaire.....	74
1.1.1. Régression linéaire simple.....	74
1.1.2. Régression linéaire multiple.....	74
1.2. Analyse par régression non linéaire.....	75
1.2.1. Le logiciel Orange3.....	75

1.2.1.1. Présentation et principe de fonctionnement du logiciel Orange3.....	75
--	----

### **PARTIE III : Résultats et Discussion**

#### **Chapitre (I) : Analyses des paramètres agro-météorologiques**

1. Analyses des rendements céréaliers.....	78
1.1. Au niveau parcellaire : les parcelles expérimentales de l'ITGC.....	78
1.2. Au niveau de la commune de Sétif.....	79
1.3. Au niveau de la wilaya de Sétif.....	80
2. Analyses de la variation saisonnière et interannuelle du NDVI.....	82
2.1 Au niveau parcellaire.....	82
2.2. Au niveau communal.....	83
2.3. Au niveau wilaya.....	84
2.4. Relation entre NDVI et le précipitation.....	84

#### **Chapitre (II) : Prédiction des rendements céréaliers**

I) Prédiction du rendement des céréales avec des données expérimentales de court terme (2020-2022) en utilisant le FGCC, le NDVI et la biomasse végétale au niveau parcellaire	88
1. Prédiction des rendements par le FGCC (Fractional Green CanopyCover).....	88
1.1. Vérification de la performance des modèles.....	89
2. Prédiction des rendements par le NDVI.....	90
2.1. Vérification de la performance des modèles.....	92
3. Prédiction des rendements en utilisant la biomasse sèche.....	93
3.1. Vérification de la performance des modèles.....	94
II) Prédiction des rendements à partir des données historiques de moyen terme (2002-2022).....	96
1. Prédiction des rendements au niveau parcellaire.....	96
1.1. Prédiction des rendements en utilisant le NDVI.....	96

1.1.1. Vérification de la performance des modèles.....	97
1.2. Prédiction des rendements en utilisant le NDVI et les précipitations.....	99
1.2.1. Vérification de la performance des modèles.....	101
1.3. Prédiction des rendements avec les algorithmes d'apprentissage (Machine learning)...	102
1.3.1 Entraînement des modèles en utilisant tous les variables au niveau parcellaire.....	102
1.3.2. Prédiction des rendements en utilisant les dix meilleurs variables.....	104
1.3.3. Evaluation de la prédiction des meilleurs modèles.....	107
2. Prédiction des rendements au niveau de la commune de Sétif.....	108
2.1. Prédiction des rendements en utilisant le NDVI.....	108
2.1.1. Vérification de la performance des modèles.....	110
2.2. Prédiction des rendements en utilisant le NDVI et les précipitations.....	112
2.2.1. Vérification de la performance des modèles.....	113
2.3. Prédiction des rendements avec les algorithmes d'apprentissage.....	115
2.3.1. Entraînement des modèles en utilisant tous les variables au niveau communal.....	116
2.3.2. Prédiction des rendements en utilisant les dix meilleurs variables.....	116
2.3.3. Evaluation de la prédiction des meilleurs modèles.....	120
3. Prédiction des rendements au niveau de la wilaya de Sétif.....	121
3.1. Prédiction des rendements par le NDVI.....	121
3. 1.1. Vérification de la performance des modèles créés.....	123
3.2. Prédiction des rendements en utilisant le NDVI et les précipitations.....	125
3.2.1. Vérification de la performance des modèles.....	126
3.3. Prédiction des rendements avec les algorithmes d'apprentissage.....	128

3.3.1. Entraînement des modèles en utilisant tous les variables au niveau wilaya.....	128
3.3.2. Prédiction des rendements en utilisant les dix meilleurs variables.....	129
3.3.3. Vérification de la performance des modèles.....	132
Discussion générale.....	134
Conclusion générale et Perspectives.....	140
Références bibliographiques .....	143
Annexes	

## ملخص

يشكل اليوم تغير المناخ العالمي تهديدا كبيرا للأمن الغذائي، مما يؤثر على إنتاج المحاصيل الزراعية. ولمواجهة هذا التحدي، من الضروري وضع نظام تشغيلي للتنبؤ بمحصول الحبوب في بداية الموسم الزراعي. إن ظهور تقنيات الاستشعار عن بعد في النظم الزراعية الدقيقة جعل التنبؤ بمحصول الحبوب ممكنا، وبالتالي توفير معلومات قيمة عن التغيرات الزمنية والمكانية في ظروف إنتاج الحبوب في الأراضي الزراعية الكبيرة وصغيرة المساحة. الهدف من هذا الدراسة هو تقييم مدى فعالية استخدام معطيات MODIS-NDVI و FGCC التي تم الحصول عليها من خلال تطبيق Canopeo® إضافة إلى الكتلة الحيوية الخضراء للتنبؤ بمحصول الحبوب من القمح والشعير من 2020 إلى 2022 على مستوى الحقل من ناحية، ومن ناحية أخرى تقييم إمكانية استخدام قيم MODIS-NDVI وحدها (1)، والجمع بين بيانات NDVI مع كميات الأمطار (2)، واستخدام نماذج التعلم الآلي (3)، للتنبؤ بمحصول الحبوب (القمح والشعير) ما بين 2002 و 2022 على ثلاث مستويات: الحقول التجريبية لـ ITGC، بلدية وولاية سطيف. و من خلال النتائج التي تم الحصول عليها يتضح أن FGCC يقدم أفضل أداء في تقييم إنتاجية القمح والشعير من بين الطرق الثلاث التي تمت دراستها على مستوى ITGC ( $R^2$  يتراوح من 0.78 إلى 0.79، RMSE بين 0.051 و 0.107، MSE من 0.004 إلى 0.018)، بالإضافة إلى أن الاستخدام المشترك لبيانات NDVI مع كميات الأمطار يقدم نتائج التنبؤ الأكثر دقة ( $R$  من 0.64 إلى 0.77، RMSE بين 0.277 إلى 0.343، MSE من 0.092 إلى 0.246). تشير هذه النتائج إلى أن تطبيق Canopeo® أثبت أنه وسيلة سريعة، موثوقة وفعالة لتقدير إنتاج حبوب القمح والشعير في قطع الأراضي الصغيرة. ومن ناحية أخرى، فإن الاستخدام المشترك لـ NDVI مع هطول الأمطار يكون أكثر فعالية ودقة مقارنة بالطرق الثلاث المستعملة. أظهرت نماذج التعلم الآلي أداءً جيدًا في توقع إنتاج القمح، خاصة نموذج " Stochastic Gradient Descent" الذي حقق قيم  $R^2$  تبلغ 0.651 و 0.783 وقيم RMSE تبلغ 0.368 طن/هكتار و 0.231 طن/هكتار على مستوى البلدية والولاية، على التوالي. بينما، بالنسبة لشعير، أظهر النموذج "Tree" أداءً جيدًا على مستوى الحقل ( $R^2 = 0.672$ ،  $R^2 = 0.289$  RMSE طن/هكتار) والولاية ( $R^2 = 0.390$ ،  $RMSE = 0.151$  طن/هكتار). ومع ذلك، سجل نموذج "Gradient Boosting" أداءً أقل على مستوى البلدية ( $R^2 = 0.358$ ،  $RMSE = 0.375$  طن/هكتار).

**الكلمات المفتاحية:** مردود، الحبوب، FGCC، الكتلة الحيوية، الاستشعار عن بعد، NDVI، التعلم الآلي.

## Résumé

Aujourd'hui, le changement climatique mondial constitue une menace importante pour la sécurité alimentaire, affectant le rendement des cultures. Pour faire face à ce défi, il est essentiel de mettre en place un système opérationnel de prévision du rendement céréalier en début de saison agricole. L'avènement des technologies de télédétection dans les systèmes agricoles de précision a rendu possible la prévision des rendements céréaliers, fournissant ainsi des informations précieuses sur les variations temporelles et spatiales des conditions de production céréalières sur les terres cultivées à grande et à petite échelle. Le but de cette recherche était d'une part, évaluer le potentiel de l'utilisation des données MODIS-NDVI, FGCC obtenue de l'application Canopeo® et de la biomasse sèche pour prédire les rendements céréaliers du blé et de l'orge entre 2020 à 2022 au niveau parcellaire. D'autre part évaluer le potentiel d'utilisation des valeurs MODIS-NDVI seul (1), combinaison des données NDVI avec les précipitations (2), et l'utilisation des modèles d'apprentissage automatique (3), pour prédire le rendement des céréales (blé et orge) entre 2002 et 2022 dans trois échelles d'étude : les parcelles expérimentales, la commune et la wilaya de Sétif. Les résultats obtenus montrent que FGCC présente les meilleures performances d'évaluation des rendements de blé et d'orge parmi les trois méthodes étudiées au niveau parcellaire, avec ( $R^2$ = variait de 0.78 à 0.79, RMSE= entre 0.051 et 0.107, MSE= de 0.004 à 0.018), ainsi que l'utilisation combinée du NDVI au données de pluviométrie présente les résultats de prédiction les plus précis avec ( $R$ = de 0.64 à 0.77, RMSE= entre 0.277 à 0.343, MSE= de 0.092 à 0.246 ). Ces résultats indiquent que l'application Canopeo® s'est avérée être une méthode rapide et fiable pour estimer le rendement en grains de blé et d'orge sur des petits parcelles. En revanche, l'utilisation combinée du NDVI avec les précipitations est plus efficace et précise que les trois autres méthodes utilisées. Les modèles de machine learning ont montré des bonne performances de prédiction pour le rendement de blé, notamment "Stochastic Gradient Descent" avec des  $R^2$  de 0.651 et 0.783 et des RMSE de 0.368t/ha et 0.231t/ha, au niveau communal et wilaya, respectivement. Tandis que pour l'orge, le modèle 'Tree' a montré des bonnes performances au niveau parcellaires ( $R^2$ =0.672, RMSE=0.289t/ha) et wilaya ( $R^2$ =0.390, 0.151t/ha). Cependant, le modèle "Gradient Boosting" a enregistré des performances moindre au niveau communale ( $R^2$ =0.358, RMSE=0.375t/ha).

**Mots clé :** Rendement, Céréales, FGCC, Biomasse, Télédétection, NDVI, Apprentissage automatique.

## Abstract

Today, global climate change poses a significant threat to food security, affecting crop yields. To face this challenge, it is essential to put in place an operational system for forecasting cereal yield at the start of the agricultural season. The advent of remote sensing technologies in precision agricultural systems has made the prediction of grain yields possible, providing valuable information on temporal and spatial variations in grain production conditions on large- and small-scale croplands. The aim of this research was, on the one hand, evaluate the potential of using MODIS-NDVI, FGCC data obtained from the Canopeo® application and biomass to predict cereal yields of wheat and barley between 2020 to 2022 at field level. On the other hand, evaluate the potential for using MODIS-NDVI values alone (1), combination of NDVI data with precipitation (2), and the use of machine learning models (3), to predict cereal yield (wheat and barley) between 2002 and 2022 in three study levels: the ITGC experimental plots, the commune and the wilaya of Sétif. The results obtained show that FGCC presents the best performance in evaluating wheat and barley yields among the three methods studied at the ITGC plot level with ( $R^2$ = varied from 0.78 to 0.79, RMSE= between 0.051 and 0.107, MSE= from 0.004 to 0.018), as well as the combination of NDVI and rainfall data presents the most accurate prediction results with ( $R$ = from 0.64 to 0.77, RMSE= between 0.277 to 0.343, MSE= from 0.092 to 0.246). These results indicate that the Canopeo® application has proven to be a rapid and reliable method for estimating wheat and barley grain yield on small plots. On the other hand, the combined use of NDVI with precipitation is more efficient and precise than the other three methods used. Machine learning models have shown good prediction performance for wheat yield, particularly "Stochastic Gradient Descent," with  $R^2$  values of 0.651 and 0.783 and RMSE values of 0.368 t/ha and 0.231 t/ha at the communal and wilaya levels, respectively. Meanwhile, for barley, the 'Tree' model exhibited good performance at the parcel level ( $R^2=0.672$ , RMSE=0.289 t/ha) and wilaya level ( $R^2=0.390$ , RMSE=0.151 t/ha). However, the "Gradient Boosting" model recorded lower performance at the communal level ( $R^2=0.358$ , RMSE=0.375 t/ha).

**Keywords:** Yield, cereals, FGCC, Biomass, Remote sensing, NDVI, Machine learning.

## Liste des abréviations

**%** : pourcent

**°C**: degrés Celsius

**µm** : Micromètre

**aSPI** : Indice de précipitation agricole standardisé

**AVHRR**: Advanced Very High Resolution Radiometer

**CNES** : Centre National d'Études Spatiales

**DSA** : Direction des Services Agricoles

**eRDI** : L'indice de sécheresse de reconnaissance efficace

**ETP** : L'évapotranspiration potentielle

**FAO**: Food and Agriculture Organization

**FAOSTAT**: Food and Agriculture Organization Corporate Statistical Database

**FGCC** : Couverture fractionnée du couvert vert

**g** : gramme

**GLAM**: Global Agricultural Monitoring

**GTS** : Système Global de télécommunications

**ha** : Hectare

**INRA** : Institut National de la Recherche Agronomique

**IR** : l'infrarouge

**ITGC** : l'Institut Technique de Grandes Cultures

**Kg** : kilogramme

**Km** : kilomètres

**KNN** : k-Nearest Neighbors

**LACIE** : Large Area Crop Inventory Experiment

**m** : Mètres

**MODIS** : Spectroradiomètre Imageur à Résolution Moyenne

**NASA** : National Aeronautics and Space Administration

**NDVI** : Indice de végétation par différence normalisée

**nm**: Nanomètre

**NOAA**: National Oceanic and Atmospheric

**ONS** : Office National des Statistiques

**P** : Précipitations

**PIR** : Canal Proche infra-rouge

**PVI** : Perpendicular Vegetation Index

**R** : canal du Rouge

**SVM** : Support Vector Machine

**t** : Tonnes

**T**: Température

**TSAVI**: Transformed Soil Adjusted Vegetation Index

**U.S.D.A**: Département Américain de l'Agriculture

**UV**: Ultraviolets

**VI** : Indices de végétation

## Liste des tableaux

<b>Tableau 1 :</b> Caractéristiques des principaux satellites opérationnels en 2017, offrant des images dans le domaine optique avec des résolutions spatiales élevées, moyennes et basses, notamment pour le suivi de la végétation Bégué et <i>al.</i> (2018).	<b>36</b>
<b>Tableau 2 :</b> Principaux production végétales (DPSB, 2021).	<b>46</b>
<b>Tableau 3:</b> Principaux production animales (DPSB, 2021).	<b>46</b>
<b>Tableau 4 :</b> Résultats d'analyses physico-chimiques du sol des parcelles expérimentales de l'ITGC-Sétif.	<b>57</b>
<b>Tableau 5:</b> tableau des statistiques descriptif des rendements de blé et d'orge au niveau des parcelles de l'ITGC entre 2002 et 2022.	<b>78</b>
<b>Tableau 6:</b> tableau des statistiques descriptif des rendements de blé et d'orge au niveau des de la commune de Sétif entre 2002 et 2022.	<b>80</b>
<b>Tableau 7:</b> tableau des statistiques descriptif des rendements de blé et d'orge au niveau des parcelles de l'ITGC entre 2002 et 2022.	<b>81</b>
<b>Tableau 08 :</b> Résultats de performances des modèles au niveau parcellaire en utilisant le FGCC.	<b>90</b>
<b>Tableau (09) :</b> Résultats de performances des modèles au niveau parcellaire en utilisant le NDVI durant la période 2020-2022.	<b>93</b>
<b>Tableau (10) :</b> Résultats de performances des modèles au niveau parcellaire en utilisant la biomasse sèche.	<b>95</b>
<b>Tableau (11).</b> Résultats de performances des modèles de prédiction du rendement au niveau parcellaire, en utilisant le NDVI.	<b>99</b>
<b>Tableau (12) :</b> Résultats de performances des modèles de régression linéaire multiple associant le NDVI et la pluviométrie au niveau parcellaire.	<b>102</b>
<b>Tableau (13) :</b> Résultats d'analyse de performance des modèles de prédiction des	<b>103</b>

rendements de blé et d'orge en utilisant toutes les variables au niveau parcellaire

**Tableau (14):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant les dix meilleurs variables au niveau parcellaire **105**

**Tableau (15) :** Résultats de performances des modèles au niveau de la commune de Sétif en utilisant le NDVI. **112**

**Tableau (16) :** Résultats de performances des modèles au niveau de la commune de Sétif en utilisant le NDVI et les précipitations. **114**

**Tableau (17):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant toutes les variables au niveau de la commune de Sétif **116**

**Tableau (18) :** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant les dix meilleurs variables au niveau de la commune de Sétif **117**

**Tableau (19).** Résultats de performances des modèles au niveau de la wilaya de Sétif en utilisant le NDVI. **125**

**Tableau (20):** Résultats de performances des modèles le NDVI et les précipitations au niveau de la wilaya de Sétif **127**

**Tableau (21):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant tout les variables au niveau de la wilaya de Sétif **129**

**Tableau (22):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant les dix meilleurs variables au niveau de la wilaya de Sétif **130**

## Liste des figures

<b>Figure 1:</b> La taxonomie de quelques céréales (Zibouche et grimes, 2016).....	<b>7</b>
<b>Figure 2 :</b> Production, utilisation et stocks mondiale des céréales (FAO, 2023).....	<b>12</b>
<b>Figure 3 :</b> Marché mondial du blé (USDA, 2023)	<b>13</b>
<b>Figure 4 :</b> Les cinq premiers pays producteurs de blé (FAO, 2023)	<b>14</b>
<b>Figure 5 :</b> Marché mondial de l'orge (USDA, 2023)	<b>14</b>
<b>Figure 6 :</b> Les cinq premiers pays producteurs de l'orge (FAO, 2023)	<b>15</b>
<b>Figure 7:</b> Evolution du rendement de blé (gauche) et de l'orge (à droite) en Algérie (USDA, 2023)	<b>18</b>
<b>Figure 8:</b> Evolution des importations de blé et d'orge en Algérie (FAOstat, 2023)	<b>19</b>
<b>Figure 9 :</b> Carré de rendement (Paden)	<b>24</b>
<b>Figure 10 :</b> principe de base de la télédétection (Martignac, 2006)	<b>26</b>
<b>Figure 11 :</b> Représentation schématique d'une onde électromagnétique (Bonn et Rochon, 1992).	<b>27</b>
<b>Figure 12 :</b> Le spectre électromagnétique (Bonn et Rochon, 1992).	<b>28</b>
<b>Figure 13 :</b> Signatures spectrales, caractéristiques des principaux types de surfaces.	<b>29</b>
<b>Figure 14 :</b> Le processus de télédétection.	<b>30</b>
<b>Figure 15:</b> Drones à rotors multiples.	<b>34</b>
<b>Figure 16 :</b> Situation géographique de la wilaya de Sétif.	<b>44</b>
<b>Figure 17:</b> Carte de répartition des pluies de la wilaya de Sétif (Rouabhi, 2017)	<b>48</b>
<b>Figure 18 :</b> Précipitations moyennes mensuelles au niveau de la wilaya de Sétif entre 2002 et 2022.	<b>48</b>
<b>Figure 19 :</b> Précipitations annuelle au niveau de la wilaya de Sétif entre 2002 et 2022.	<b>49</b>

<b>Figure 20:</b> Températures moyennes mensuelles au niveau de la wilaya de Sétif entre 2002 et 2022.	<b>50</b>
<b>Figure 21:</b> Températures moyennes annuelles au niveau de la wilaya de Sétif entre 2002 et 2022	<b>50</b>
<b>Figure 22 :</b> Diagramme ombrothermique de la wilaya de Sétif entre 2002 et 2022.	<b>51</b>
<b>Figure 23 :</b> Localisation géographique de la commune de Sétif (Beldjazia, 2022).	<b>52</b>
<b>Figure 24 :</b> Précipitations moyennes mensuelles enregistrées au niveau de la commune de Sétif de 2002 à 2022.	<b>53</b>
<b>Figure 25 :</b> Précipitations annuelle au niveau de la commune de Sétif entre 2002 et 2022.	<b>53</b>
<b>Figure 26:</b> Températures moyennes mensuelles au niveau de la commune de Sétif entre 2002 à 2022.	<b>54</b>
<b>Figure 27:</b> Températures moyennes annuelles au niveau de la commune de Sétif entre 2002 à 2022.	<b>55</b>
<b>Figure 28 :</b> Diagramme ombrothermique de la commune de Sétif entre 2002 et 2022.	<b>55</b>
<b>Figure 29 :</b> situation géographique de l'ITGC de Sétif.	<b>56</b>
<b>Figure 30:</b> Localisation géographique des parcelles expérimentales de l'ITGC.	<b>56</b>
<b>Figure 31 :</b> Précipitations moyennes mensuelles enregistrées au niveau d'ITGC (Sétif) entre 2002 et 2022.	<b>58</b>
<b>Figure 32 :</b> Précipitations annuelle enregistrées au niveau d'ITGC (Sétif) entre 2002 et 2022.	<b>59</b>
<b>Figure 33 :</b> Températures moyennes mensuelles au niveau d'ITGC (Sétif) entre 2002 et 2022.	<b>60</b>
<b>Figure 34:</b> Températures moyennes annuelles au niveau d'ITGC (Sétif) entre 2002 et 2022.	<b>60</b>
<b>Figure 35:</b> Diagramme ombrothermique au niveau de l'ITGC-Sétif- entre (2002-2022).	<b>61</b>
<b>Figure 36 :</b> Canopeo® application.	<b>67</b>

<b>Figure 37</b> : Application Canopeo (a), Image originale (b), Image de FGCC après traitement par Canopeo® (c).	<b>69</b>
<b>Figure 38</b> : collection de la biomasse (a), échantillons placées dans des sacs identifiés (b), échantillons soumises à un séchage (c), mesure du poids secs des échantillons (d).	<b>70</b>
<b>Figure 39</b> : Méthodologie de prévision des rendements céréaliers avec le logiciel Orange3.	<b>77</b>
<b>Figure 40</b> : Variabilité temporelle des rendements (blé et orge) de 2002 à 2022 au niveau des parcellaire.	<b>79</b>
<b>Figure 41</b> : Variabilité temporelle des rendements (blé et orge) de 2002 à 2022 au niveau de la commune de Sétif	<b>80</b>
<b>Figure 42</b> : Variabilité temporelle des rendements (blé et orge) de 2002 à 2022 au niveau de la wilaya de Sétif.	<b>81</b>
<b>Figure 43</b> : Variabilité saisonnière et interannuelle du NDVI de blé et d'orge au niveau parcellaire entre 2002 et 2022.	<b>83</b>
<b>Figure 44</b> : Variabilité saisonnière et interannuelle du NDVI de blé et d'orge au niveau de la commune de Sétif entre 2002 et 2022	<b>84</b>
<b>Figure 45</b> : Variabilité saisonnière et interannuelle du NDVI de blé et d'orge au niveau de la wilaya de Sétif entre 2002 et 2022	<b>85</b>
<b>Figure 46</b> : Corrélation entre le NDVI moyen (de blé et d'orge) et la pluviométrie cumulée (mm) au niveau des parcelles expérimentales de l'ITGC (2002 à 2022)	<b>86</b>
<b>Figure 47</b> : Corrélation entre le NDVI moyen (de blé et d'orge) et la pluviométrie cumulée (mm) au niveau de la commune de Sétif (2002 à 2022)	<b>86</b>
<b>Figure 48</b> : Corrélation entre le NDVI moyen (de blé et d'orge) et la pluviométrie cumulée (mm) au niveau de la wilaya de Sétif (2002 à 2022)	<b>87</b>
<b>Figure 49</b> : Evolution du coefficient de détermination entre le rendement en grains et le FGCC pour les deux céréales (orge, blé) de 2020 à 2022 au niveau des parcelles de l'ITGC.	<b>88</b>
<b>Figure 50</b> : Comparaison des rendements observés et prédit selon le modèle créé pour le blé et l'orge en utilisant FGCC au niveau parcellaire.	<b>90</b>
<b>Figure (51)</b> : Evolution du coefficient de détermination entre le rendement en	<b>91</b>

grains et le NDVI pour les deux céréales (orge, blé) de 2020 à 2022 au niveau parcellaire.

**Figure (52).** Comparaison des rendements observés et prédit selon le modèle créé pour le blé et l'orge en utilisant le NDVI au niveau parcellaire. **92**

**Figure (53) :** Evolution du coefficient de détermination entre le rendement en grains et la biomasse sèche pour les deux céréales (orge, blé) de 2020 à 2022 au niveau des parcelles de l'ITGC. **93**

**Figure (54) :** Corrélation entre les rendements observés et prédits selon le modèle créé pour le blé et l'orge en utilisant la biomasse sèche au niveau parcellaire **95**

**Figure (55) :** Evolution du coefficient de déterminations entre le rendement en grains et le NDVI des deux céréales (orge et blé), couvrant la période de 2002-2022 au niveau parcellaire **96**

**Figure (56) :** Modèle de régression linéaire entre le rendement (de blé et d'orge) et le NDVI, le 13 mars couvrant la période 2002-2022, au niveau parcellaire. **97**

**Figure (57).** Corrélation entre les rendements observés et prédit du blé en utilisant le NDVI au niveau parcellaire **98**

**Figure (58).** Corrélation entre les rendements observés et prédit de l'orge en utilisant le NDVI au niveau parcellaire **98**

**Figure (59) :** Evolution du coefficient de détermination dans le cas de la combinaison du NDVI et de la pluviométrie dans la prévision du rendement des deux céréales (orge et blé) au niveau parcellaire. **100**

**Figure (60) :** Corrélation entre les rendements prédit et observé du blé au niveau parcellaire en utilisant un modèle de régression multiple (NDVI et pluviométrie) entre 2002-2022 **101**

**Figure (61) :** Corrélation entre les rendements prédit et observé de l'orge au niveau parcellaire en utilisant un modèle de régression multiple (NDVI et pluviométrie) entre 2002-2022 **102**

<b>Figure (62)</b> : Les dix meilleurs variables de prédiction du rendement du blé au niveau parcellaire.	<b>106</b>
<b>Figure (63)</b> : Les dix meilleurs variables de prédiction du rendement de l'orge au niveau parcellaire	<b>107</b>
<b>Figure (64)</b> : Rendements prédit et observé du blé au niveau parcellaire obtenus par le modèle "KNN" entre 2002-2022	<b>108</b>
<b>Figure (65)</b> : Rendements prédit et observé d'orge au niveau parcellaire obtenus par le modèle "Tree" entre 2002-2022	<b>108</b>
<b>Figure (66)</b> : Evolution du coefficient de détermination entre le rendement et le NDVI pour les deux céréales (orge et blé), couvrant la période 2002 -2022 dans la commune de Sétif	<b>109</b>
<b>Figure (67)</b> : Modèle de régression linéaire entre le rendement et le NDVI le 08 Mai pour le blé et le 30 Avril pour l'orge au niveau de la commune de Sétif.	<b>110</b>
<b>Figure (68)</b> : Corrélation entre les rendements observés et prédits selon le modèle élaboré pour le blé en utilisant le NDVI au niveau de la commune de Sétif	<b>111</b>
<b>Figure (69)</b> : Corrélation entre les rendements observés et prédits selon le modèle élaboré pour l'orge en utilisant le NDVI au niveau de la commune de Sétif	<b>111</b>
<b>Figure (70)</b> : Evolution du coefficient de détermination dans le cas de la combinaison du NDVI et de la pluviométrie dans la prévision du rendement des deux céréales (orge et blé) au niveau communale	<b>112</b>
<b>Figure (71)</b> : Corrélation entre les rendements prédits et observés du blé au niveau de la commune de Sétif en utilisant le NDVI et la pluviométrie de la période 2002-2022.	<b>114</b>
<b>Figure (72)</b> : Corrélation entre les rendements prédits et observés de l'orge au niveau de la commune de Sétif en utilisant le NDVI et la pluviométrie de la période 2002-2022.	<b>114</b>
<b>Figure (73)</b> : Les dix meilleurs variables de prédiction du rendement du blé au	<b>119</b>

niveau communale.

<b>Figure (74) :</b> Les dix meilleurs variables de prédiction du rendement de l'orge au niveau communale	<b>120</b>
<b>Figure (75) :</b> Comparaison entre les rendements du blé observés et prédits par le modèle "Stochastic Gradient Descent" au niveau de la commune de Sétif	<b>121</b>
<b>Figure (76) :</b> Comparaison entre les rendements d'orge observés et prédits par le modèle "Gradient Boosting" au niveau de la commune de Sétif	<b>121</b>
<b>Figure (77) :</b> Evolution du coefficient de détermination entre le rendement en grains et le NDVI pour les deux céréales (orge et blé) couvrant la période de 2002 à 2022 dans la wilaya de Sétif	<b>122</b>
<b>Figure (78) :</b> Modèle de régression linéaire entre le rendement et le NDVI, le 30 Avril pour le blé et le 08 Mai pour l'orge au niveau de la wilaya de Sétif.	<b>123</b>
<b>Figure (79) :</b> Corrélacion entre les rendements observés et prédits du blé en utilisant le NDVI	<b>124</b>
<b>Figure (80) :</b> Corrélacion entre les rendements observés et prédits de l'orge en utilisant le NDVI	<b>124</b>
<b>Figure (81) :</b> Evolution des coefficients de détermination entre le rendement, pour les deux céréales (orge et blé) en utilisant l'NDVI et la pluviométrie, couvrant la période de 2002- 2022 au niveau de la wilaya de Sétif.	<b>125</b>
<b>Figure (82):</b> Relation entre les rendements prédits et observés du blé au niveau de la wilaya de Sétif en utilisant le NDVI et la pluviométrie entre 2002 et 2022.	<b>127</b>
<b>Figure (83) :</b> Relation entre les rendements prédits et observés d'orge au niveau de la wilaya de Sétif en utilisant le NDVI et la pluviométrie entre 2002 et 2022	<b>127</b>
<b>Figure (84) :</b> Les dix meilleurs variables sélectionnés pour prédire le rendement du blé au niveau de la wilaya de Sétif	<b>131</b>
<b>Figure (85) :</b> Les dix variables les plus pertinentes pour prédire le rendement d'orge à l'aide d'Orange3 au niveau de la wilaya de Sétif	<b>132</b>

**Figure (86) :** Comparaison entre les rendements du blé, observés et prédits par le modèle "Stochastic Gradient Descent" au niveau de la wilaya de Sétif **132**

**Figure (87) :** Comparaison entre les rendements du blé, observés et prédits par le modèle "Tree" au niveau de la wilaya de Sétif **133**

## Liste des annexes

**Tableau (23)** : la moyenne de la production de blé et d'orge au niveau de la wilaya de Sétif entre 2002 et 20022 (DSA Sétif, 2023).

**Tableau (24)** : la moyenne de la production de blé et d'orge au niveau de la commune de Sétif entre 2002 et 20022 (DSA Sétif, 2023).

**Tableau (25)** : la moyenne de la production de blé et d'orge au niveau des parcelles expérimentales de l'ITGC de Sétif entre 2002 et 20022 (DSA Sétif, 2023).

## Les travaux réalisés dans le cadre de cette thèse

**-Boulaaras, H** et Bouregaa, T. (2024). Cereal yield forecasting in semi-arid region of Algeria using MODIS-NDVI. *Journal of Aridland Agriculture*, 10: 7-14.

[.https://doi.org/10.25081/jaa.2024.v10.856](https://doi.org/10.25081/jaa.2024.v10.856)

**-Boulaaras, H** et Bouregaa, T. (2022). Application de la télédétection dans la prédiction précoce du rendement d'orge (*Heurdeum Vulgare L.*) en zone semi-aride. Séminaire national : sur « l'agriculture et le développement durable en zone semi-arides (ADDZSA) », le 15 et 16 Novembre 2022, Souk ahras.

**-Boulaaras, H** et Bouregaa, T. (2022). Prédiction précoce du rendement de blé dur (*triticum durum L.*) par télédétection en zone semi-aride. Séminaire national : sur « la valorisation des ressources naturelles et de l'environnement », le 30 Mars 2022, Sétif.

# **Introduction**

## **Introduction**

La production agricole est fortement influencée par les conditions environnementales, sociales et économiques à l'échelle mondiale, notamment en ce qui concerne les cultures stratégiques et leur lien avec la sécurité alimentaire (Ruiter et *al.*, 2017).

Les céréales et leurs dérivés constituent les principales ressources alimentaires de l'humanité, en raison de leur source d'énergie et leur grande richesse en protéines. Principalement destinés à l'alimentation humaine (à hauteur de 75% de la production), elles assurent 15% des besoins énergétiques, elles servent également à l'alimentation animale (15% de la production) et à des usages non alimentaires (Feillet, 2000). Le blé, avec le riz et le maïs, est l'une des trois principales cultures vivrières mondiales (Cai et *al.*, 2019). Le blé tendre est l'une des céréales les plus importantes, assurant la nourriture de près de 40 % de la population mondiale (Liu et *al.*, 2020). Sans oublier l'orge qui se classe au quatrième rang mondial en termes de superficie cultivée et de quantité produite après le blé, le riz et le maïs (Geng et *al.*, 2022).

Au cours des dernières années, il y a eu une augmentation significative des événements météorologiques extrêmes, notamment des vagues de chaleur, des sécheresses et des périodes prolongées de précipitations intenses (Yan et *al.*, 2022), qui diminuent la production agricole (Lobell et *al.*, 2013) et qui devraient réduire la production céréalière mondiale de maïs et de blé de 3.8 % et 5.5 %, respectivement (Lobell et *al.*, 2011), ce qui entraînerait des perturbations dans les chaînes d'approvisionnement alimentaire à l'échelle mondiale, accentuant ainsi les défis en matière de sécurité alimentaire et de nutrition dans les régions du monde les plus vulnérables (Verschuur et *al.*, 2007).

En Algérie, les céréales jouent un rôle important dans les habitudes alimentaires de la population, englobant les activités de production et de transformation telles que la production de semoule et la boulangerie dans l'industrie alimentaire (Mohammed Ammar, 2014). La consommation des produits céréaliers en Algérie se situe à environ 180Kg tête/habitant/an (Mahmoud et *al.*, 2017). Le blé est la principale céréale cultivée, suivie par l'orge et l'avoine, bien que la superficie réellement récoltée soit souvent inférieure à celle cultivée, même lorsque les conditions climatiques sont favorables (Aidani, 2015).

La production agricole nationale est fortement dépendante de ses conditions climatiques, marquées par des variations annuelles des précipitations, la rareté de l'eau et des températures

élevées pendant les périodes de croissance des cultures, ces facteurs ont un impact négatif sur la production agricole des céréales (Mekhlouf et *al.*, 2012). En 2022, l'Algérie a importé 10.6 millions de tonnes de céréales. La majorité de ces importations ont été constituées de blé tendre, représentant près de 6.1 millions de tonnes, suivi du maïs avec 2.6 millions de tonnes (en baisse par rapport aux 4.8 millions de tonnes de la campagne précédente), du blé dur avec près de 1.4 million de tonnes et de 571 000 tonnes d'orge ( Algérie Eco, 2022).

L'Algérie figure parmi les pays menacés par le changement climatique, ce qui représente un risque pour la sécurité alimentaire, qui impacte négativement les activités agricoles de manière multiforme, affectant les revenus des ménages agricoles, les composés biophysiques de l'agriculture ainsi que la dynamique collective des territoires agricoles. Par conséquent, cela contribue à accroître la vulnérabilité à la pauvreté (Dugué, 2012). Dans les zones semi-arides des Hauts Plateaux, la culture du blé est confrontée à plusieurs contraintes biotiques et surtout abiotiques qui limitent fortement le potentiel de production (Fellahi et *al.*, 2019).

Dans ces circonstances, des prévisions précises, fiables et opportunes des rendements des céréales sont essentielles pour assurer la sécurité alimentaire nationale et internationale (Becker-Reshef et *al.*, 2020). L'estimation des rendements peuvent informer les agriculteurs sur l'efficacité de leurs pratiques agricoles (densité de semis, fertilisation, etc.), facilitant la détection précoce des problèmes liés aux facteurs de rendement des cultures. Cette approche contribue à améliorer les rendements et les revenus des agriculteurs (Kayad et *al.*, 2022). Malgré que les méthodes traditionnelles de comptage manuel des tiges (comptage du nombre de tiges, de tiges principales et de talles) sont précises pour les prévisions de rendement. Cependant, cette méthode est également laborieuse et prend du temps, ce qui fait que de nombreux agriculteurs n'évaluent pas leurs rendement de céréales (Goodwin, Lindsey., 2018), aussi elles s'avèrent insuffisantes pour prédire le rendement des cultures dans de grandes régions en raison de contraintes, telles que les limitations budgétaires, les contraintes de temps et le manque de main-d'œuvre qualifiée (Fang et *al.*, 2008).

Les avancées technologiques dans les capteurs satellitaires ont donné lieu au développement de la télédétection (remote sensing), qui est une science et une technique visant à recueillir des informations sur des objets terrestres à partir d'images satellitaire sans avoir besoin d'un contact direct (Sabins, 1987). Aujourd'hui la télédétection a démontré son efficacité en tant

que technologie dans l'agriculture de précision, en particulier pour le suivi des cultures (Seo et *al.*, 2019) et la prévision du rendement des cultures (Zhang et *al.*, 2020 ; Mashaba et *al.*, 2017). Il s'agit également d'une méthode non destructive, rentable, peu coûteuse (Son et *al.*, 2014) et rapide pour la modélisation de la croissance des cultures (Seo et *al.*, 2019). Les données de télédétection peuvent couvrir une vaste zone pour récupérer les informations spectrales en temps réel pendant la période de croissance des cultures (Idrissi et *al.*, 2022).

Les indices de végétation (VI) représentent l'une des techniques de télédétection les plus utilisées. Ils sont calculés à partir de différentes longueurs d'onde capturées par un capteur et servent à évaluer divers aspects de la végétation, tels que la couverture, la vitalité, la dynamique de croissance, l'état nutritionnel, et d'autres applications liées (Baio et *al.* , 2018 ; Junior et coll., 2018). Parmi les indices largement appliqués dans le secteur agricole, l'indice de végétation par différence normalisée (NDVI) (Eckert et *al.*, 2015). Le NDVI a été proposé pour la première fois par Rouse et *al.*, 1974. Il est défini comme le rapport entre la différence et la somme des valeurs de réflectance dans le spectre proche infrarouge (NIR) et rouge (RED). Ce calcul permet de caractériser la végétation photosynthétiquement active (Tucker et *al.*, 2005). Le NDVI a été utilisé pour évaluer l'état de la couverture végétale dans des contextes de surveillance agricole et environnementale (Shanmugapriya et *al.*, 2019). A démontré une forte corrélation avec la biomasse végétale (Kumari et *al.*, 2021). Récemment, de nombreuses études ont utilisé le NDVI comme base pour prédire les rendements céréaliers, quantifier la couverture du couvert végétal et examiner les réponses écologiques aux fluctuations environnementales, notamment la sécheresse, les inondations (Daoudi et *al.*, 2023). En raison de sa simplicité et de sa large disponibilité à l'état « prêt à l'emploi » à partir de nombreux satellites et télédétection fournisseurs de données (Tenreiro et *al.*, 2021).

En plus de l'indice NDVI, il existe également la mesure de la couverture fractionnée du couvert vert (FGCC), qui est définie comme la fraction de la superficie projetée couverte par une végétation photosynthétiquement active (Wittich et Hansing, 1995). FGCC est mesurée à l'aide d'une application mobile gratuite appelée Canopeo®, utilisant l'appareil photo d'un téléphone, a été développée par "l'Oklahoma State University". Canopeo® analyse les pixels en se basant sur les rapports de couleur rouge sur vert et bleu sur vert, ainsi que sur un indice de verdure excédentaire, pour classer automatiquement les pixels comme verts ou non verts. Cette méthode

permet de mesurer la proportion de la surface couverte par une végétation active (Patrignani et Ochsner, 2015). En plus, de nos jours la majorité des agriculteurs utilisent des smartphones dotés de caméras et de capacités de traitement des images. L'évaluation basée sur les smartphones est une approche prometteuse pour estimer efficacement la couverture végétale, démontrant une forte corrélation avec les méthodes manuelles utilisées pour mesurer la biomasse aérienne, la surface foliaire et l'interception de la lumière (Xiong et *al.*, 2019).

Canopeo® a été utilisé dans différentes cultures pour mesurer la couverture du couvert végétal (Shepherd et *al.*, 2018). Casadesús et coll. (2007) ont trouvé une forte corrélation entre le FGCC, la biomasse du blé dur et le rendement en grains. Les résultats de FGCC obtenue du Canopeo® ont déjà été comparés à l'interception de la lumière (Shepherd et *al.*, 2018), à l'estimation visuelle (Büchi et *al.*, 2018) et à d'autres méthodes de classification de pixels (Büchi et *al.*, 2018 ; Patrignani&Ochsner, 2015). les différentes méthodes sont fortement corrélées et Canopeo® a été considéré comme un outil précis et objectif pour le suivi des cultures (Jáuregui et *al.*, 2019).

Pour cela les objectifs principaux pour ce travail sont :

- D'une part, évaluer le potentiel de l'utilisation des données MODIS-NDVI, FGCC et de la biomasse pour prédire les rendements céréaliers à différents stades de croissance du blé et de l'orge entre 2020 à 2022 au niveau des parcelles expérimentales de l'ITGC.
- D'autre part, évaluer le potentiel d'utilisation des valeurs MODIS-NDVI seul (1), combinaison des données NDVI avec les précipitations (2), et l'utilisation d'un ensemble de variables agro-météorologiques avec des algorithmes d'apprentissage (machine learning) (3), pour prédire le rendement des céréales (blé et orge) à différents stades de croissance entre 2002 et 2022 à trois échelles d'étude : les parcelles expérimentales de l'ITGC, la commune et la wilaya de Sétif.
- Identifier la méthode la plus efficace pour prédire les rendements céréaliers avant la récolte en créant des modèles de prévisions du rendement des céréales. Également de déterminer la date optimale pour des prévisions plus précises dans les régions semi-arides de Sétif en utilisant les données acquises à partir des images satellitaires pour le NDVI obtenues de la plateforme GLAM (Global Agricultural Monitoring) et l'application Canopeo® pour la FGCC, pour les données météorologiques et les données FLDAS

(Famine Early Warning Systems Network (FEWS NET). Land Data Assimilation System) et Climate Engine pour extraire l'humidité du sol.

Dans cette étude, la thèse est articulée autour de trois sections principales :

-La première section propose une synthèse bibliographique. Elle traite, d'une part, des aspects généraux concernant les céréales, leur répartition géographique, leur importance à l'échelle mondiale ainsi que leur place prépondérante en Algérie, tout en explorant les contraintes de production associées. D'autre part, elle examine les concepts fondamentaux de la télédétection et présente les différentes méthodes utilisées pour évaluer le rendement des cultures céréalières.

-La deuxième section détaille les matériaux et méthodes employés dans l'expérimentation, ainsi que les analyses statistiques réalisées pour évaluer les données collectées.

-La troisième section est dédiée à la présentation des résultats obtenus et à leur discussion approfondie. Elle permet d'analyser les conclusions tirées des données recueillies et de les mettre en perspective avec les objectifs de l'étude.

En conclusion, une synthèse générale des résultats ainsi que des perspectives futures seront présentées à la clôture de cette étude.

**PARTIE I :**  
**Synthèse**  
**Bibliographique**

# **Chapitre I : Aperçu générale sur les céréales (Blé et l'orge)**

## **1. Généralités sur les céréales**

### **1.1. Définition**

Le terme céréale est un dérivé du mot latin "cerealis" signifiant «grain» qui est botaniquement, un type de fruit appelé caryopse, consommable par l'homme ou par les animaux domestiques (Sarwaret *al.*, 2013). La plupart des céréales appartiennent à la famille des Poacées. Ce sont : le blé, l'orge, l'avoine, le seigle, le maïs, le riz, le millet, le sorgho. Les unes appartiennent à la sous-famille des Festucoïdées : blé, orge, avoine, seigle, les autres à la sous-famille des Panicoïdées: maïs, riz, sorgho, millet (Moule, 1971).

### **1.2. Historique**

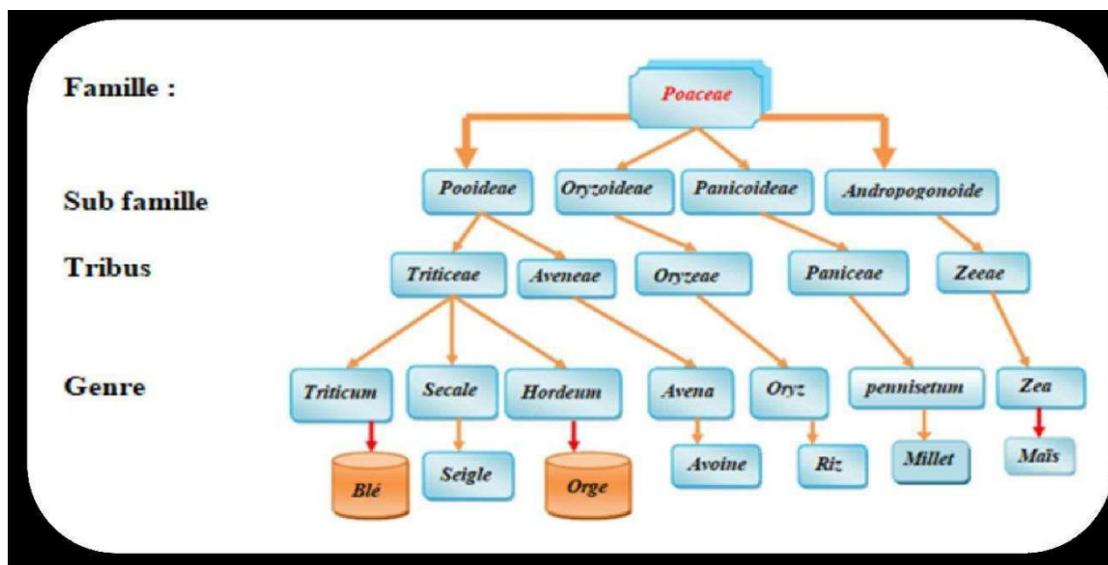
Il ya 10 000 ans, la domestication des céréales a marqué une étape cruciale dans le développement des sociétés humaines, marquant le début de l'ère néolithique qui entraînera l'adoption d'une économie de production, fondée sur l'agriculture et l'élevage (Shewry, 2009). Selon Clerget (2011), les céréales majeures se regroupent en trois catégories, représentant ensemble 75 % de la consommation mondiale de céréales. Le premier grand groupe de céréales comprend le blé, l'orge, le seigle et l'avoine, émergeant dans le triangle fertile, berceau des civilisations occidentales qui trouvent leur origine au Moyen-Orient et au Proche-Orient. Le deuxième groupe est constitué par le maïs, originaire d'Amérique centrale, il est à la base des civilisations amérindiennes. Le maïs a été introduit en Europe par les explorateurs du Nouveau Monde à la fin du XVe siècle. Enfin, un troisième groupe est centré autour du riz, plante originaire des régions chaudes et humides de l'Asie du Sud-Est. Sa domestication a eu lieu de manière synchronisée avec celle du blé, plus à l'ouest, et le riz est à la base des civilisations orientales.

### **1.3. Taxonomie**

Les céréales font partie de l'ordre des monocotylédones, où la famille des graminées étant la plus vaste. Cette famille compte plus de 600 genres et 10 000 espèces, une diversité remarquable qui s'explique, entre autres, à sa grande extension géographique (Pottier et *al.*, 2012). On la retrouve sur tous les continents, ainsi que sur la plupart des îlots océaniques. Sous toutes les latitudes où les conditions climatiques favorisent la croissance des plantes supérieures, depuis

l'équateur jusqu'aux régions proches des cercles polaires, où seuls les lichens peuvent lui succéder.

D'après (Zibouche et grimes, 2016 ) nous avons la classification ci-dessous (figure 1) :



**Figure (1):** La taxonomie de quelques céréales (Zibouche et grimes, 2016)

## 2. Cycle de développement des céréales

### 2.1. Phase végétative

#### 2.1.1. Stade de germination –Levée

La germination de la graine se manifeste par l'émergence du coléorhize, donnant lieu à la formation de racines séminales, et de la coléoptile, qui assure la protection de la première feuille fonctionnelle lors de sa sortie. La véritable levée se produit dès que les feuilles apparaissent à la surface du sol. Dans un peuplement, la levée est considérée comme atteinte lorsque la plupart des lignes de semis sont visibles. Au cours de la phase de semis-levée, l'alimentation de la plante dépend exclusivement de son système racinaire primaire et des réserves contenues dans la graine (Gate, 1995).

#### 2.1.2. Stade de tallage

Selon Soltner (1988), cette phase représente un mode de développement spécifique aux graminées, se caractérisant par la formation du plateau de tallage, l'émission des talles, et l'émergence de nouvelles racines. Ces processus sont favorisés par des températures moyennes oscillant entre 9 et 22 °C (Mekliche, 1983). Le tallage se manifeste par l'apparition d'une tige

secondaire, appelée talle, à la base de la première feuille, tandis que d'autres feuilles développent également leurs talles vertes. À l'intérieur de la tige, la pointe de croissance, qui commence à ressembler à un épi de blé, émerge initialement sous terre pour se protéger du gel. Au fur et à mesure de la reprise de la végétation, la pointe de croissance va s'élever dans la tige.

## **2.2. Phase reproductrice**

### **2.2.1. Stade de montaison**

La montaison débute au stade épi de 1 cm par l'allongement du premier entre-nœud. Pendant cette période, la croissance de la plante se caractérise par l'arrêt du tallage et la formation des premières structures d'épillets dans le tiers inférieur du futur épi. La durée de cette phase s'étend sur une période de 29 à 30 jours (Clément-Grand court et Prat, 1971). La montaison représente la phase la plus critique du développement des céréales telles que le blé dur, le blé tendre et l'orge. Elle prend fin au moment de l'émission de la dernière feuille et des signes de gonflement causés par les épis à l'intérieur de la gaine (Baldy, 1984).

### **2.2.2. Stade d'épiaison**

Ce stade englobe la période des épis, débutant avec l'émergence des premiers épis et se prolongeant jusqu'à ce que tous les épis aient complètement émergé de la gaine formée par la dernière feuille (Bourahla et Hadji, 2009).

### **2.2.3. Stade de floraison- fécondation**

La floraison débute dès l'apparition de quelques étamines visibles dans la partie médiane de l'épi, en dehors des glumelles. Lorsque les anthères se manifestent, elles affichent une couleur jaune, qui évolue vers le blanc après une exposition au soleil, signalant ainsi l'accomplissement de la fécondation. Le grain de pollen des blés est monospore, et sa dispersion est relativement limitée. Le nombre de fleurs fécondées est influencé par la nutrition azotée et une évapotranspiration modérée (Soltner, 1988).

#### **2.2.4. Stade remplissage et maturation des grains**

Les plantes poursuivent leur croissance en accumulant des réserves dans les grains. Il est nécessaire de surveiller l'apparition de maladies ou de parasites afin d'intervenir rapidement (Battinger, 2002). Cette étape se caractérise par l'augmentation de la taille du grain, l'accumulation d'amidon et la perte naturelle d'humidité des grains par déshydratation, marquant ainsi la fin du processus de maturation (Soltner, 1988). Cette phase de maturation s'étend en moyenne sur 45 jours. Les grains se remplissent progressivement et traversent différents stades :

- **Phase de la maturité laiteuse :** Cette phase se distingue par la migration des substances de réserve vers le grain et la formation des enveloppes. À ce stade, le grain présente une couleur vert clair, une texture laiteuse, et atteint ses dimensions finales (Soltner, 1988).
- **Phase maturité pâteuse :** Au cours de cette phase, les réserves se déplacent des parties vertes vers les grains. Il y a une augmentation de la teneur en amidon, tandis que le taux d'humidité diminue. Lorsque le blé atteint la maturité, la plante est sèche, et les grains des épis sont riches en réserves (Soltner, 1988).
- **Phase maturité complète :** Suite à la phase pâteuse, le grain poursuit sa maturation en se déshydratant. Il acquiert une teinte jaune, durcit et devient brillant. Ce stade est particulièrement influencé par les conditions climatiques (Soltner, 1988).

### **3. Exigences des céréales**

#### **3.1. Exigences pédoclimatique**

##### **3.1.1. Le climat**

D'après Soltner (2005), les principales exigences climatiques de la céréaliculture sont les suivantes: la température, l'ensoleillement (photopériode) et l'eau.

##### **✓ La température**

La physiologie des céréales est constamment influencée par la température. Une température supérieure à 0 °C (zéro de végétation des céréales) est nécessaire pour la germination des céréales. La température joue également un rôle crucial dans la nitrification et l'activité végétative des céréales pendant le tallage et la montaison. De plus, l'intensité de l'évaporation peut entraîner des phénomènes d'échaudage. Une baisse soudaine de la température, combinée à un assèchement intense en surface, entraîne des dommages sous forme de nécroses (Soltner, 2005).

✓ **L'ensoleillement (la photopériode)**

L'impact positif de la température sur le rendement est renforcé par le rayonnement solaire et la durée du jour lorsque la température n'est pas excessivement élevée. En revanche, ces facteurs accentuent son impact négatif dans le cas contraire. La photopériode influe également sur la durée de chacune des phases de développement mentionnées. Les variétés des trois types de céréales (blé dure, blé tendre et l'orge) présentent des différences dans leur sensibilité à la photopériode) (Kalarasse, 2018).

✓ **Les Précipitations**

L'eau joue un rôle crucial dans la croissance des plantes (Soltner, 1990), et la germination ne se produit qu'à partir d'un degré d'imbibition d'eau de 30%. En effet, les besoins en eau sont les plus importants de la phase épi 1 cm à la floraison. La période critique en termes d'approvisionnement en eau s'étend de 20 jours avant l'épiaison jusqu'à 30 à 35 jours après la floraison (Nedjah, 2015).

## **2. Le sol**

Les céréales peuvent s'adapter à des sols variés en utilisant les fertilisants et les variétés appropriées. Les qualités d'un sol propice comprennent:

- Une texture fine, de type limono-argileuse, offrant aux racines fasciculées des céréales une grande surface de contact pour une nutrition optimale.
- Une structure stable capable de résister à la dégradation causée par les précipitations.

## **3.2. Exigences en nutriments**

### **3.2.1. Les éléments majeurs**

Les céréales nécessitent trois éléments essentiels, l'azote, le phosphore et le potassium (NPK), et chacun de ces éléments joue un rôle spécifique :

- **Azote (N)** : les céréales ont un besoin crucial en azote pour favoriser leur croissance, le développement des tiges et des feuilles, ainsi que la formation des protéines. En particulier, pendant la phase de croissance végétative, l'azote joue un rôle essentiel dans le soutien de ces processus physiologiques (Lui et *al.*, 2005).
- **Phosphore (P)** : le phosphore joue un rôle crucial pour les céréales, contribuant significativement au développement racinaire, à la floraison et à la formation des graines.

Son importance est particulièrement marquée au début du cycle de croissance des céréales (Johnston, 2002).

- **Potassium (K)**: joue un rôle fondamental pour les céréales en participant à plusieurs processus physiologiques, la régulation de la turgescence cellulaire, la synthèse des protéines et des glucides, aussi renforce la résistance aux maladies. Il est essentiel pour promouvoir la santé et la productivité des céréales (Mengel *et al.*, 1987).

### 3.2.2. Les oligo-éléments

La croissance et la production des céréales dépendent de plusieurs oligo-éléments qui sont essentiels pour les plantes. Ces éléments nutritifs sont nécessaires pour le bon fonctionnement des processus biochimiques impliqués dans la croissance, le développement et la production de graines. Voici quelques-uns des oligo-éléments les plus indispensables pour la production des céréales :

**-Le magnésium (Mg)** : Le magnésium se trouve présent surtout dans les feuilles comme composant de la chlorophylle. Les prélèvements par des récoltes plus importantes et par le lessivage, font perdre chaque année au sol de 40 à 60 kg de magnésium Mg par hectare.

**-Le fer (Fe)** : Il est crucial pour la synthèse de la chlorophylle, qui est nécessaire à la photosynthèse. Le fer est également impliqué dans d'autres processus métaboliques des plantes.

**-Le zinc (Zn)** : Essentiel pour la synthèse des protéines et la régulation des hormones de croissance. Il participe également à la formation des chloroplastes.

**-Le cuivre (Cu)** : Nécessaire pour plusieurs enzymes impliquées dans la photosynthèse et le métabolisme des glucides.

**-Le manganèse (Mn)** : Joue un rôle crucial dans la photosynthèse, l'activation d'enzymes et la formation de chloroplastes.

**-Le molybdène (Mo)** : Il est impliqué dans la fixation de l'azote atmosphérique, un processus essentiel pour la synthèse des protéines.

**-Le cobalt (Co)** : Bien que nécessaire en très petites quantités, le cobalt peut jouer un rôle dans la fixation de l'azote.

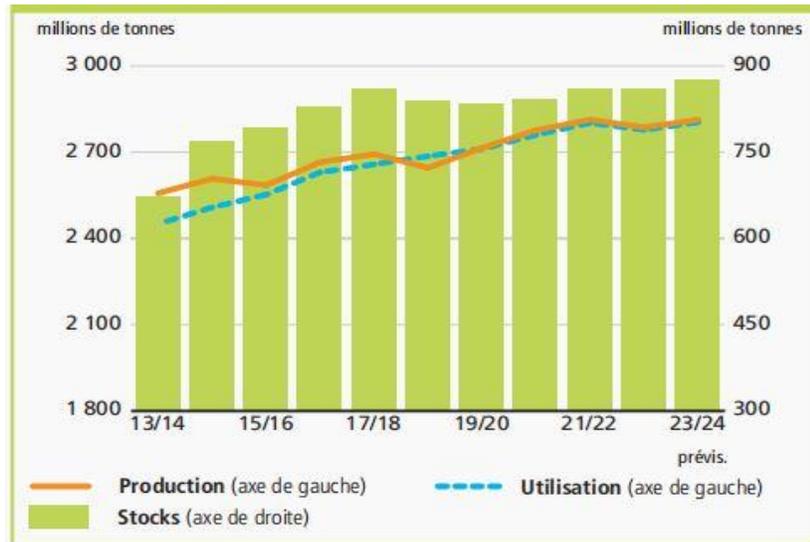
## 4. Importance des céréales et leurs distributions

### 4.1. Dans le monde

#### 4.1.1. Production

Pour l'année 2023, la production mondiale de céréales secondaires est estimée à 2823 millions de tonnes, avec une prévision de croissance de 0.9% (soit 25.9 millions de tonnes) par rapport à l'année précédente, dépassant ainsi le record établi en 2021 (FAO, 2023).

D'après les prévisions de la FAO, l'utilisation mondiale de céréales en 2023-2024 atteindra 2 813 millions de tonnes, soit 1.1 pour cent de plus qu'en 2022-2023. Une hausse étayée par les projections d'une forte croissance de l'utilisation de blé dans l'alimentation animale et de l'augmentation de la consommation par la population et d'autres utilisations (figure 2).

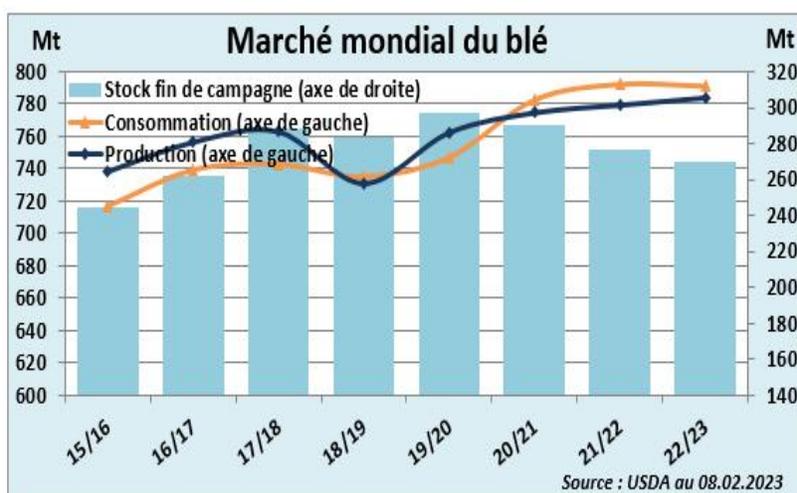


**Figure (2) :** Production, utilisation et stocks mondiale des céréales (FAO, 2023)

La guerre en Ukraine et les tensions entre la Russie et l'Occident ont perturbé la production et les exportations de céréales dans la région. Les agriculteurs ont abandonné leurs terres en raison des conflits et des troubles dans l'est de l'Ukraine, ce qui a entraîné une diminution de la production. De plus, les sanctions économiques imposées par l'Occident à la Russie ont limité la production de céréales dans les régions russes. La diminution de la production a affecté les prix des céréales sur les marchés internationaux en raison de la limitation des exportations et de la réduction des stocks mondiaux (Perthuis, 2022).

Le blé est la deuxième céréale la plus produite derrière le maïs, et le commerce mondial du blé est supérieur à toutes les autres cultures combinées. Dans le monde, la culture du blé dur n'est pas assez importante que le blé tendre mais, elle est plus lucrative. L'industrie des pâtes a augmenté la demande sur le grain de blé dur qui a un prix de 10% à 20% supérieur à celui de blé tendre (Sallet *al.*, 2019).

Selon USDA (United States Department of Agriculture), la production mondiale totale du blé en 2023 était de 787 millions de tonnes, mais elles restent 2.1% (17.1 millions de tonnes) en dessous des niveaux de 2022. Pour la troisième année consécutive, la consommation mondiale de blé devrait dépasser le volume produit. Ainsi, les stocks mondiaux vont se contracter de près de 3 %, à 269 Mt. Le ratio « stocks / consommation » s'élèverait à 34 %, soit un recul de 6 % par rapport à la moyenne des cinq dernières années (Figure 3).



**Figure (3) :** Marché mondial du blé (USDA, 2023)

Selon les statistiques de FAO (2023), la Chine, l'Inde et la Russie sont les trois plus grands producteurs individuels de blé au monde, représentant environ 41 % de la production mondiale totale du blé. Les Etats-Unis d'Amérique est le quatrième producteur individuel de blé au monde. Cependant, l'Union européenne, si elle était comptée comme un seul pays, sa production de blé dépasserait celle de n'importe quel pays à l'exception de la Chine (Figure 4).

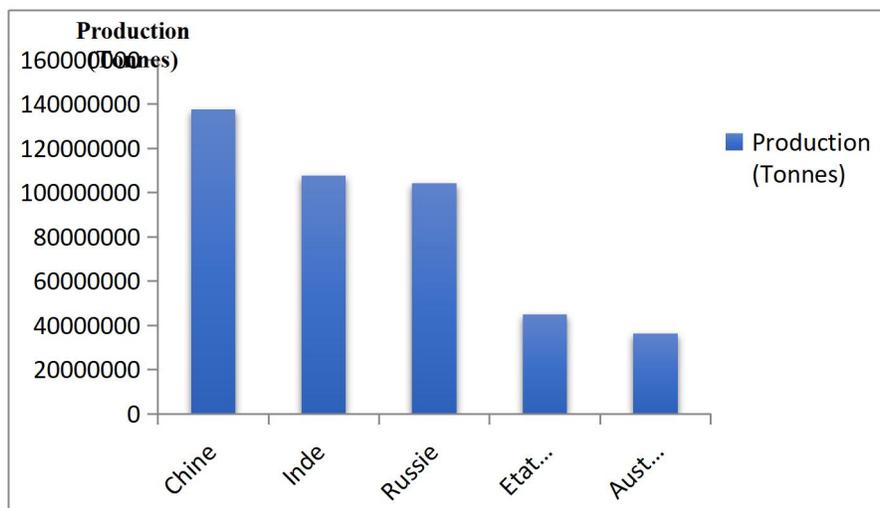


Figure (4) : Les cinq premiers pays producteurs de blé (FAO, 2023)

L'orge se classe à la 4ème place des céréales les plus cultivées dans le monde. Son utilisation se partage entre alimentation animale (70 %), de la production et malterie, et dans certains cas, pour l'alimentation humaine. D'après les données du département américain de l'Agriculture (USDA), la production mondiale d'orge pour l'année 2023 s'élève à 150 millions de tonnes, enregistrant une croissance de plus de 3 % par rapport à la récolte de l'année précédente. Il convient de noter que cette récolte demeure relativement stable par rapport aux cinq années antérieures (Figure 5).

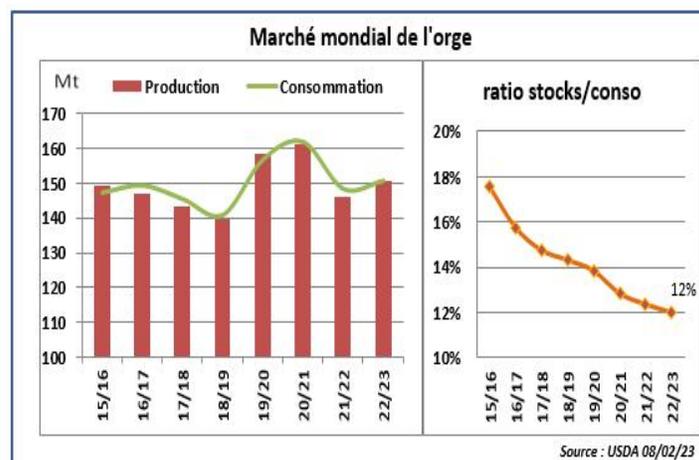
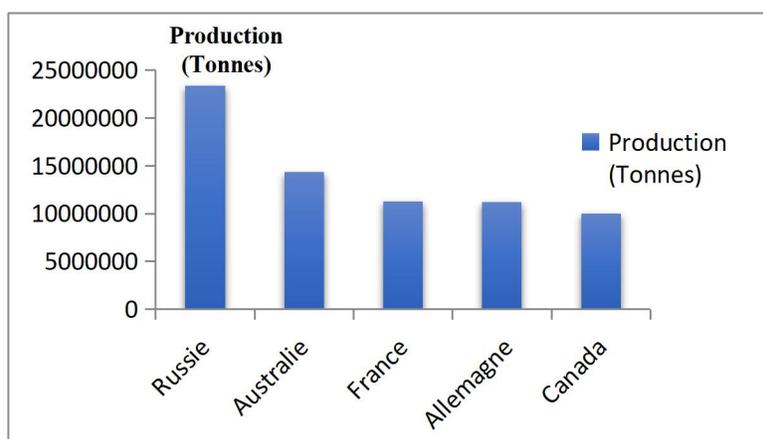


Figure (5) : Marché mondial de l'orge (USDA, 2023)

Les principaux pays producteurs mondiaux d'orge : Les pays de l'UE représentent à eux seuls 51 Mt, soit un tiers de la production mondiale. La Russie confirme sa position d'outsider

avec un niveau record de 21.5 Mt, suivie de l’Australie, dont la production connaît sa quatrième campagne consécutive sans accident climatique majeur. Suivie de la France, l’Allemagne et Canada (Figure 6).



**Figure (6):** Les cinq premiers pays producteurs de l’orge (FAO, 2023)

## 4.2. En Algérie

### 4.2.1. Répartition géographique

En Algérie, la céréaliculture est répartie en fonction de leurs espèces et son type d'adaptation dans les étages bio-climatiques suivant :

- **Zone à hautes potentialités:** on y trouve une pluviométrie moyenne supérieure à 500 mm/an, avec des rendements moyens de 20qx/ha (plaines de l’Algérois et Mitidja, bassin des Issers, vallées de la Soummam et de l’Oued El Kébir, vallée de la Seybouse...). Cette zone couvre une SAU de 400 000 (Ministre de l’agriculture, du développement rural et de pêche, 2022) ha dont moins de 20% sont consacrés à la production du blé.
- **Zone à moyennes potentialités:** caractérisée par une pluviométrie supérieure comprise entre 400 et 500 mm/an, mais sujette à des crises climatiques élevées, les rendements peuvent varier de 5 à 15qx/ha (coteaux de Tlemcen, vallées du Chélif, massif de Médéa...). La zone englobe une SAU de 1 600 000 ha dont moins de la moitié est réservée aux céréales.
- **Zone à basses potentialités:** caractérisée par un climat semi-aride et situé dans les hauts plateaux de l’Est et de l’Ouest et dans le Sud du Massif des Aurès. La moyenne des précipitations est inférieure à 350 mm par an. Ici, les rendements en grains sont le plus souvent inférieurs à 8qx/ha. La SAU de la zone atteint 4,5 millions d’ha dont près de la moitié est emblavée.

#### **4.2.2. Place et importance des céréales en Algérie**

Les céréales sont une ressource essentielle pour l'alimentation en Algérie, fournissant la majorité des calories et des protéines dans le régime alimentaire national. La production des céréales en Algérie présente une caractéristique fondamentale depuis l'indépendance travers l'extrême variabilité du volume des récoltes. Cette particularité témoigne d'une maîtrise insuffisante de cette culture et de l'indice des aléas climatiques. Cette production est conduite en extensif et à caractère essentiellement pluvial (Bencharif *et al.*, 2007). Il est donc, facile de prédire qu'elle ne pourrait satisfaire les demandes d'une population qui dépassant actuellement 44 millions (ONS, 2021) d'habitants, est traditionnellement consommatrice des céréales.

De l'époque coloniale à ce jour, les moyennes décennales de la sole céréalière occupent entre 2.4 et 3.2 millions d'hectares, soit 28 à 40% de la SAU et le reste étant réparti entre la jachère et les autres cultures. La plus grande superficie revient au blé dur qui occupe une moyenne de 1.27 millions ha, moyenne calculée sur la période de 1989 à 2019 (MADR, 2020). Le blé tendre se classe en troisième position après l'orge. Sa superficie moyenne, durant la période de 1989 à 2019, reste stable et faible. En effet, la surface moyenne du blé tendre avoisine 562 000 d'ha (Moyenne 1989/2019). Elle ne représente que 31% de la surface moyenne (1989/2019) des blés et 18% de la surface moyenne des céréales durant la même période (1989/2019).

Le blé en Algérie est la principale céréale cultivée, suivie de l'orge et de l'avoine, il occupe une place très importante dans le système agroalimentaire, qui connaît une dépendance vis-à-vis de l'étranger pour satisfaire la demande, ce qui met ainsi en péril la sécurité alimentaire du pays. En effet, la production des blés représente une fraction de 68.82% du total de la production céréalière, toutes espèces confondues et 62% du total de la superficie céréalière, toutes espèces confondues (MADR, 2020). Par ailleurs, la valeur ajoutée de la filière blé représente, selon les statistiques du Ministère de l'Agriculture (2017), environ 130.16 millions de DA, correspondant à 4.44% du total de la valeur de la production agricole.

#### **4.2.3. Production**

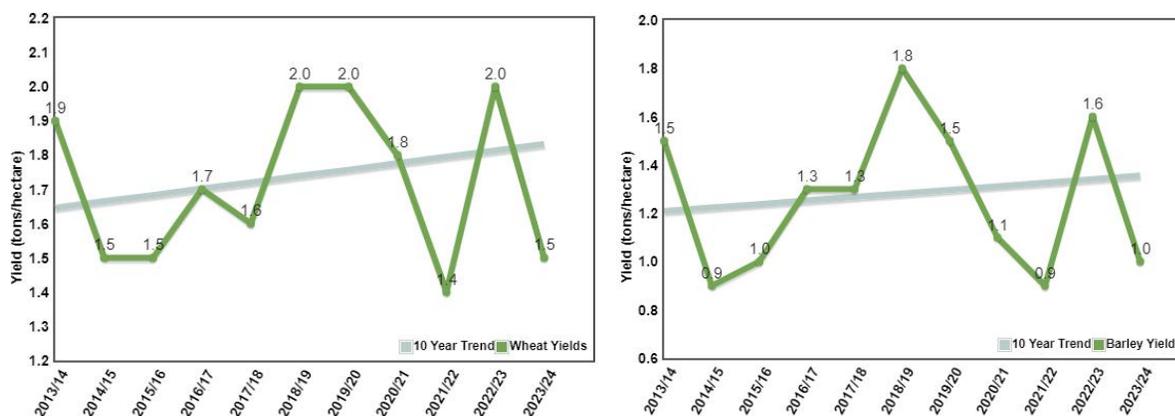
Selon les données de l'USDA (2023), au cours de la campagne agricole 2022-2023, la production de blé en Algérie a atteint 3.6 millions de tonnes. Cette performance a suivi une légère

hausse par rapport à la campagne 2021- 2022, où la production était 2.7 millions de tonnes, témoignant ainsi d'une certaine variabilité annuelle.

En ce qui concerne l'orge, la production nationale a connu une diminution significative en 2021, chutant 950 millions de quintaux, mais a enregistré une forte reprise en 2022 avec une production de 1.6 millions de tonnes. Ces fluctuations suggèrent des variations importantes des rendements céréaliers en Algérie.

En examinant l'évolution des rendements du blé en Algérie de 2013 à 2024, on observe une fluctuation notable, avec des rendements variant de 1.4 à 2 tonnes par hectare. Les campagnes 2018/2019 et 2019/2020 se caractérisent par les rendements les plus élevés, atteignant 2 t/ha. En revanche, les campagnes 2014/2015, 2015/2016 et 2021/2022 ont connu les rendements inférieurs, suggérant des défis ou des conditions défavorables pour la production de blé. Les campagnes 2016/2017, 2017/2018, 2020/2021 et 2022/2023 présentent des rendements relativement stables, bien que la campagne 2023/2024 montre une légère diminution par rapport à la campagne précédente (figure 7).

En ce qui concerne les rendements de l'orge de 2013 à 2024, une certaine instabilité est notée, avec des rendements variant de 0.9 à 1.8 t/ha. Les campagnes 2018/2019 et 2022/2023 se distinguent par les rendements les plus élevés, atteignant respectivement 1.8 et 1.6 t/ha. En revanche, les campagnes 2014/2015, 2021/2022 et 2023/2024 ont enregistré les rendements inférieurs. Les campagnes 2015/2016, 2016/2017, 2017/2018, 2019/2020 et 2020/2021 présentent des rendements relativement stables. La campagne 2023/2024 affiche une baisse significative des rendements par rapport à la campagne précédente. Ces variations annuelles des rendements des céréales peuvent être influencées par des facteurs tels que les conditions météorologiques. De plus, les épisodes de sécheresse qui ont touché l'Afrique du Nord en 2021 ont également eu un impact négatif sur la production de céréales en Algérie (figure 7).

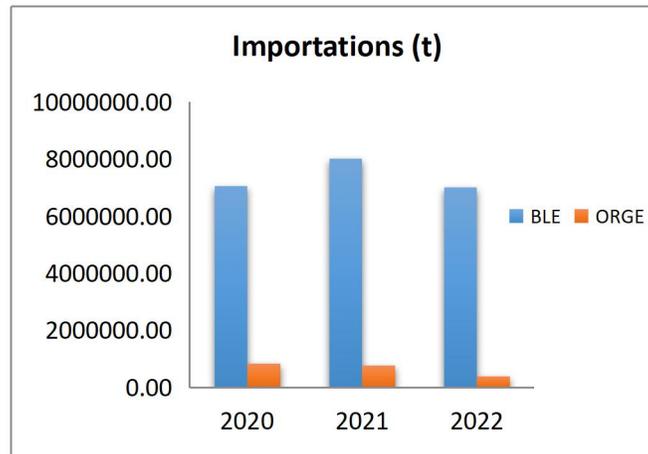


**Figure (7):** Evolution du rendement de blé (gauche) et de l'orge (à droite) en Algérie (USDA, 2023)

L'Algérie a consommé environ 11.37 millions de tonnes de blé entre juillet 2020 et juin 2021, selon un rapport conjoint du Global Agricultural Information Network et du ministère Américain de l'Agriculture. Les stocks de céréales de l'Algérie ont augmenté de 5.6 millions de tonnes en 2017 à 6.7 millions de tonnes en 2020, mais ont diminué de 6 % pour atteindre 6.3 millions de tonnes en 2021, avec une prévision de chute à 5,1 millions de tonnes en 2022, d'après la FAO (Assoko, 2022).

Malgré les efforts consentis par le gouvernement pour la filière blé et l'orge, la production de l'Algérie en blé et en orge n'est pas du tout suffisante pour satisfaire les exigences d'une population à croissance rapide. En effet, la production moyenne en blé (1989/2019) a atteint 2.2 millions de tonnes, tandis que la consommation moyenne des blés par habitant est évaluée à 256.97 kg/hab/an (moyenne 2017/2019), d'où un besoin en blé dur de 4.1 millions de tonnes/an et en blé tendre de 6.8 millions de tonnes/an pour alimenter une population avoisinant les 44 millions d'habitants (FAO, 2020).

Un déficit agricole en blé constaté annuellement. Il est assuré par les importations, ce qui classe l'Algérie parmi les premiers importateurs de blé au monde derrière l'Égypte et l'Indonésie. En effet, le pays affiche un taux de dépendance aux importations moyenne durant la période de 2015 à 2017 de 76,10% (FAO, 2020).



**Figure (8):** Evolution des importations de blé et d'orge en Algérie (FAOstat, 2023)

## 5. Les contraintes de production des céréales en Algérie

Les défis de la production céréalière en Algérie comprennent principalement les rendements limités. Plusieurs facteurs peuvent expliquer ces contraintes de la production céréalière en Algérie :

### 5.1. Contraintes pédoclimatiques

#### 5.1.1. Climat

#### 5.1.2. Le stress hydrique

Le déficit hydrique ou le manque d'eau constitue le stress abiotique le plus sévère auquel les céréales sont confrontées dans les conditions de production des zones arides et semi-arides. (Chennafi et *al.*, 2006). Les précipitations en Algérie sont peu abondantes et ne surviennent pas toujours de manière opportune et en quantités suffisantes (Bourihane et Mekkaoui, 2013). Le manque d'eau entraîne une réduction, voire un arrêt, de la croissance, un flétrissement des parties aériennes et en cas de stress intense ou prolongé, peut conduire à la mort de la plante (Despinasse, 2015). L'impact du stress varie en fonction de son intensité, de sa durée, du stade de développement de la plante, du génotype et de son interaction avec l'environnement (Foudili et Gasmi, 2017). Chez les céréales, le déficit hydrique influe sur les trois composantes majeures du rendement : le nombre d'épis, le nombre de grains par épi et le poids de 1000 grains (Debaeke et *al.*, 1996 ; Semcheddine, 2015).

### **5.1.3. Le stress thermique**

Les contraintes liées à la température peuvent exercer un impact significatif sur la production des céréales en Algérie. Ces contraintes comprennent des fluctuations de température qui peuvent être préjudiciables aux cultures, en particulier dans les régions sensibles aux variations thermiques. Selon Soltner (2005), les conditions idéales pour les céréales nécessitent une somme de températures de 150 °C pendant la période de semis-levée, de 500 °C de la phase levée à la fin du tallage. En revanche, la phase d'épiaison à la maturité physiologique nécessite 850 °C. Néanmoins, ces besoins varient en fonction des géotypes, des saisons agricoles et des conditions environnementales spécifiques (Arauset *al.*, 2003). Un excès de chaleur a un impact négatif sur les céréales, en accélérant la sénescence des feuilles, en raccourcissant la phase de remplissage du grain et en réduisant le poids de 1000 grains ainsi que le nombre de grains par mètre carré. Ces effets cumulés ont une incidence négative sur la productivité totale de la plante et la qualité du produit final (Fellahi, 2017). Les gelées tardives affectent les cultures céréalières en Algérie vers la fin d'avril et le début de mai, induisant un avortement des fleurs lors du stade de floraison (Kolai, 2008), peuvent détruire tous les épillets touchés qui ne parviennent pas à se développer et à produire des fleurs. Ces épillets restent stériles, et les dommages ne deviennent apparents qu'à partir de la phase d'épiaison, se manifestant par des épis minces et de couleur blanchâtre (Abbassenne, 1997 ; Mekhlouf *et al.*, 2006).

## **5.2. Sol**

Les sols du Nord de l'Algérie, principales zones de culture céréalière, se distinguent généralement par la présence de sels dans le sol. Ces zones céréalières sont caractérisées par une faible teneur en matière organique, un pH basique, ainsi qu'une carence en phosphore assimilable et en potasse.

### **5.2. Contraintes techniques**

La production céréalière en général est affectée par de nombreuses lacunes techniques, cependant, les principales se résument comme suit (Dali *et al.*, 2010; Hamadache, 2013 ; ITGC, 2016) :

-L'analyse du sol est souvent absente.

- Le travail du sol est standardisé avec l'utilisation d'outils inadaptés, parfois est retardé en raison de la sécheresse automnale.
- Le semis est fréquemment effectué au-delà des délais techniques requis, souvent à la volée.
- Les doses de semis sont souvent excessives, atteignant parfois 200 kg/ha.
- Les engrais azotés ne sont pas toujours disponibles en quantité et au moment opportun.
- Les semences autoproduites ne sont pas traitées contre les maladies fongiques.
- Les fertilisants sont appliqués de manière non raisonnée, le désherbage est souvent tardif, et l'absence des traitements phytosanitaires.
- Les pertes à la récolte, souvent importantes, sont principalement dues au retard de la récolte et au mauvais réglage des moissonneuses-batteuses.

### **5.3. Contraintes économiques**

Elles sont associées à des coûts de production élevés, résultant de la cherté des facteurs de production et du matériel agricole, ainsi la disponibilité insuffisante des intrants en qualité et en quantité dans les délais recommandés. La dimension limitée des exploitations céréalières, généralement entre 5 et 10 hectares, empêche souvent d'entreprendre des investissements, notamment pour l'acquisition d'équipements agricoles, et limite la capacité à bénéficier des différents soutiens financiers accordés à la céréaliculture.

### **5.4. Maladies, ennemis et accidents pathogènes et physiologiques**

#### **5.4.1. Maladies, ennemis**

Comme c'est le cas pour toutes les cultures, les céréales à paille peuvent être attaquées par divers organismes parasites, qu'ils soient macroscopiques ou microscopiques. Ces organismes peuvent être classés en plusieurs catégories, notamment les parasites animaux, les champignons, les bactéries, les virus et les mauvaises herbes.

#### **5.4.2. Pathogènes**

- **La moucheture** : La moucheture du blé dur, manifestée par l'apparition de taches brunes à noires sur les grains de blé lors de leur maturité, a un impact négatif sur l'apparence des produits dérivés tels que les pâtes alimentaires (Samson et Desclaux, 2006). Elle est générée par l'oxydation de polyphénols en présence d'air. La formation de la moucheture semble être encouragée par des niveaux d'humidité élevés entre la période de floraison et

le stade du grain laiteux (Hamadache, 2013). La gestion de la moucheture de blé implique souvent l'utilisation de variétés résistantes, la rotation des cultures, l'application de fongicides et d'autres pratiques culturales visant à réduire la présence du pathogène dans le champ.

#### **5.4.2. Accidents physiologiques**

Les accidents physiologiques des céréales se réfèrent à des perturbations dans le développement normal des plantes, souvent influencées par des conditions environnementales défavorables. Parmi ces accidents, on peut citer :

- **La verse** : Il s'agit de la déviation des tiges de céréales causée par l'action du vent ou de la pluie, ce qui résulte en une diminution du rendement (Hamadache, 2013).
- **L'échaudage** : L'échaudage du grain de blé peut être attribué à différentes causes, parmi lesquelles (Hamadache, 2013) :

-des conditions météorologiques défavorables : telles que des températures élevées, une exposition prolongée au soleil ou des périodes de sécheresse durant la phase stade laiteux-stade pâteux) au palier hydrique.

- Parasitaire : Les rouilles et la fusariose de l'épi ciblent les trois dernières feuilles, la punaise et d'autres insectes piqueurs-suceurs tels que les thrips et les pucerons attaquent les trois dernières feuilles ainsi que les grains en formation.

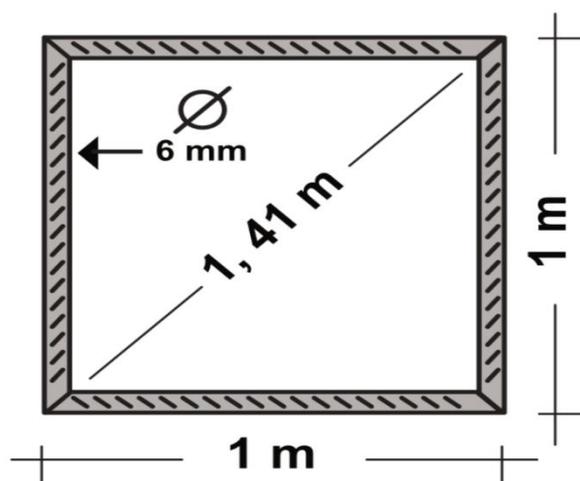
# **Chapitre II : Notions fondamentale sur la télédétection**

## 1. Méthodes d'évaluation du rendement des céréales

Le rendement agricole représente la récolte obtenue sur une superficie cultivée spécifique, généralement mesuré en quintaux métriques (1 q = 100 kg) par hectare pour les céréales, et en tonnes par hectare pour les produits à forte teneur en eau tels que les racines, les tubercules et les fruits.

### 1.1. Méthodes traditionnelle

Traditionnellement, l'évaluation du rendement de la production se fait à l'aide de carrés de rendement. Ces carrés sont positionnés aléatoirement dans une parcelle afin de compter tous les plants (Paden). Fabriqués à partir de tiges en fer rond mesurant de 1 m à 6 m de longueur et de 6 mm de diamètre, ces carrés de rendement sont assemblés par soudure (voir Figure 09). Ils sont placés dans la parcelle la journée précédant la récolte.



**Figure (9) : Carré de rendement (Paden)**

La mise en place des carrés de rendement implique trois étapes distinctes :

- Localiser la parcelle de la culture à évaluer.
- Déterminer la superficie de la parcelle, pouvant être réalisée de deux manières. La première méthode utilise un décimètre pour mesurer chaque côté de la parcelle, suivi de la détermination de la superficie. La deuxième méthode fait appel au GPS, qui permet de prendre les points limitrophes de la parcelle. En utilisant un logiciel, la superficie est ensuite déterminée, suivie de l'élaboration du schéma de la parcelle.

- Identifier les diagonales de la parcelle et les points de pose. Ensuite, il est nécessaire de placer 5 carrés par diagonale tout en évitant les bordures pour les parcelles d'un hectare. Pour les parcelles de moins d'un hectare, il faut poser 2 à 3 carrés par diagonale est recommandée, avec un minimum de 4 carrés par parcelle. Pour chaque carré :
- Compter le nombre de plants ou de pieds présents.
- Récolter les produits (tubercules, bulbes, fruits), les compter et les placer dans un contenant (sac, sachet, seau, etc.).
- Tarer le peson ou la balance avec le contenant et peser les produits récoltés.
- Reporter les données collectées sur la fiche de carré de rendement. En cas de récoltes échelonnées, additionner les données collectées de chaque récolte sur le même carré.

Pour déterminer le rendement dans le carré de rendement, diviser la quantité totale de récolte pesée en kilogrammes par la surface du carré de rendement, selon la formule :

Rendement du carré = Poids total de la récolte dans le carré / Surface du carré.

Il est généralement recommandé d'exprimer les rendements en kilogrammes par hectare (kg/ha) pour permettre une comparaison cohérente entre différentes zones et champs. Pour convertir le rendement calculé dans le carré (exprimé en kg/m<sup>2</sup>) en kg/ha, une conversion est nécessaire.

## **1.2. Méthodes innovantes**

### **1.2.1. Télédétection**

Avec les progrès scientifiques et technologiques actuels, il est possible d'observer la Terre depuis l'espace. Et exploiter ces observations pour étudier l'évolution des climats, détecter et suivre les feux de forêt, ainsi que surveiller la croissance végétale à des échelles allant de la parcelle agricole à la région, avec des fréquences temporelles pertinentes pour l'agriculture.

#### **1.2.1.1. Définition**

Le terme télédétection en anglais « remote sensing » désigne l'ensemble des techniques permettant l'étude à distance d'objets ou de phénomènes. Ce néologisme a émergé aux États-Unis dans les années soixante avec l'introduction de nouveaux capteurs complémentaires à la

photographie aérienne traditionnelle. En 1973, le terme "télédétection" a été officiellement adopté en français, avec une définition décrivant cette discipline: « Ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques d'objets par des mesures effectuées à distance, sans contact matériel avec ceux-ci » (Kergomard, 1996).

Une définition plus précise, et pour nous plus opérationnelle, de la télédétection est la suivante : « La télédétection est une technique qui permet, à l'aide d'un capteur, "d'observer" et d'enregistrer le rayonnement électromagnétique, émis ou réfléchi, par une cible quelconque sans contact direct avec celle-ci».

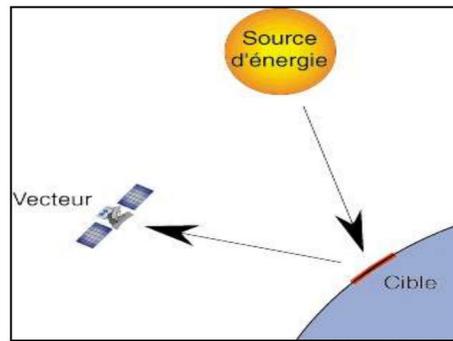
Le traitement et l'analyse des informations transportées par le rayonnement enregistré permettent d'accéder à diverses propriétés de la cible : géométriques (position, forme, dimensions), optiques (réflexion, transmission, absorption, etc.) et physico-chimiques (température, teneur en eau, chlorophylle foliaire, phyto-masse, matière organique du sol, etc.). Les plates-formes utilisées en télédétection sont variées, incluant des dispositifs au sol, des ballons, des avions et des satellites. L'acquisition de la télédétection est dite :

-Passive, lorsque la source éclairant la cible est indépendante du capteur, avec le soleil en tant que source principale. Des exemples incluent les satellites SPOT, LANDSAT et METEOSAT.

-Elle est dite active, lorsque la source éclairante est intégrée au capteur. C'est le cas des satellites LIDAR utilisant un laser et du RADAR utilisant une hyperfréquence. Contrairement à la télédétection passive, la télédétection active par radar offre la possibilité de réaliser des acquisitions en tout temps, jour ou nuit, et indépendamment des conditions météorologiques, que le ciel soit clair ou nuageux (Soudani, 2006). De ce fait, l'objectif de la télédétection est de fournir des informations sur les paysages en utilisant le rayonnement électromagnétique comme moyen de transport de ces données, sous la forme de données images (Robin, 2002).

#### **1.2.1.2. Principe de la télédétection**

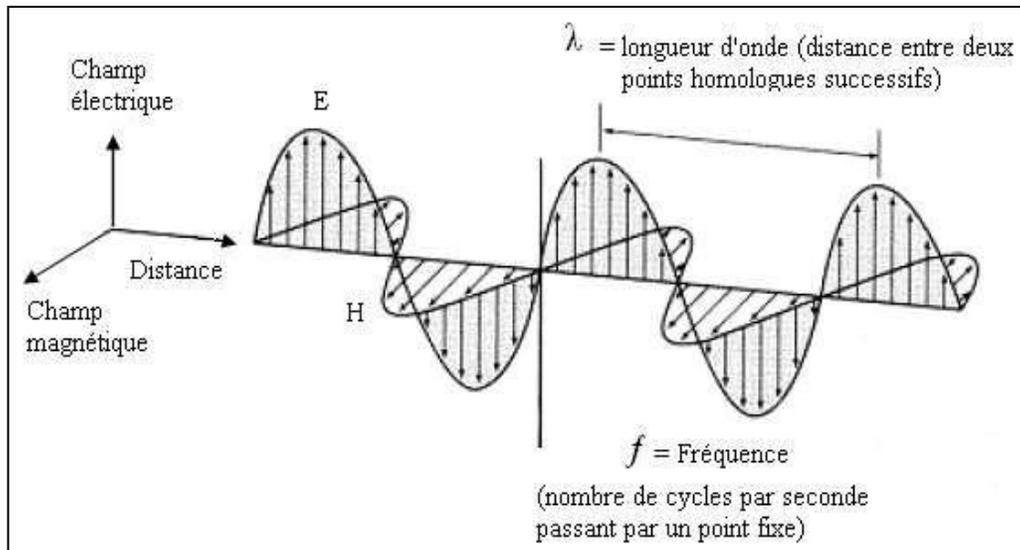
La télédétection exploite les caractéristiques du rayonnement électromagnétique afin d'analyser à distance la surface du sol, de l'océan ou de l'atmosphère. Son principe de base partage des similitudes avec la vision humaine. Ce processus découle de l'interaction entre trois éléments essentiels : une source d'énergie, une cible et un vecteur (figure 10).



**Figure (10):** principe de base de la télédétection (Martignac, 2006)

- ❖ **La cible** représente la portion de la surface terrestre observée par le satellite, pouvant s'étendre de quelques dizaines à plusieurs milliers de kilomètres carrés.
- ❖ **La source d'énergie** constitue l'élément illuminant la cible en émettant une onde électromagnétique, sous forme d'un flux de photons.
- ❖ **Le vecteur** peut être un satellite ou un avion, surplombant la cible à une altitude variant de quelques centaines de mètres à 36 000 kilomètres.
  - **Le rayonnement électromagnétique**

Le rayonnement électromagnétique représente une forme de transmission d'énergie dans la nature, la plus courante étant la lumière visible telle que perçue par l'œil humain. Il s'agit d'une énergie se propageant dans l'espace, interagissant avec la matière sous la forme d'ondes transversales indissociables (Caloz, 1992).



**Figure (11):** Représentation schématique d'une onde électromagnétique (Bonn et Rochon, 1992).

Le rayonnement électromagnétique se caractérise par :

- La période qui est le temps (T) pendant lequel l'onde effectue une oscillation complète.
- La fréquence  $\nu$  est l'inverse de la période (1/T).
- La longueur d'onde  $\lambda$  qui est la distance entre deux points homologues sur l'onde.
- La vitesse  $v$  de propagation de l'onde dans l'espace : dans le vide, et à peu de chose près, dans l'air, cette vitesse est :  $c = 300.10^3$  km/s. La relation entre vitesse, longueur d'onde et fréquence est :  $\lambda = c/\nu$ .

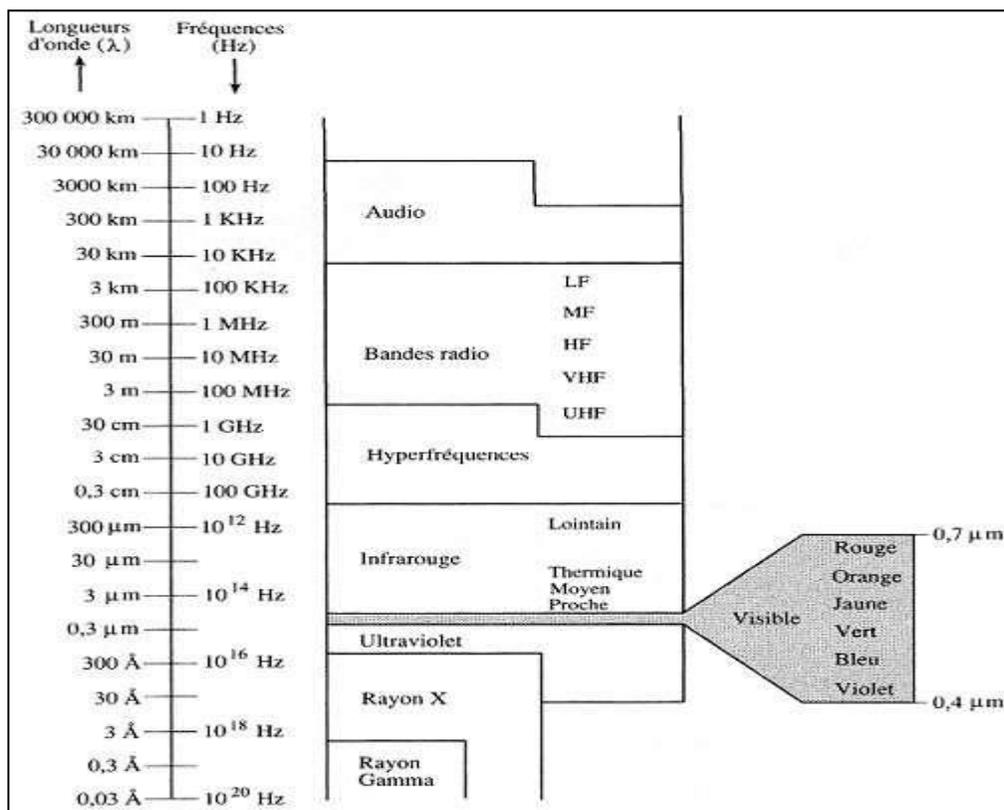
### ➤ Le spectre électromagnétique

Le spectre électromagnétique représente la répartition des ondes électromagnétiques en fonction de leur longueur d'onde, de leur fréquence ou bien encore de leur énergie. En ordre croissant de longueurs d'onde, on distingue les types de rayonnements suivants : les rayons  $\gamma$ , les rayons X, les rayonnements ultraviolets (UV), les rayonnements visibles (de 0.4 à 0.7  $\mu\text{m}$ ), l'infrarouge (IR), les micro-ondes (domaine radar), les ondes métriques (utilisées en télévision), et les ondes radioélectriques (de 10 m à quelques kilomètres) (voir figure 12). Le rayonnement se manifeste par ses interactions avec la matière. Selon l'état et la nature de cette dernière, elle peut produire un rayonnement électromagnétique (émission propre) ou interagir avec le rayonnement qui l'atteint. Ces propriétés d'absorption, d'émission et de réflexion sont fondamentales pour l'interprétation des données enregistrées par les satellites. Cependant, ces mêmes phénomènes

peuvent perturber le rayonnement enregistré par le satellite lors de sa traversée de l'atmosphère, provenant de l'objet observé (Hamimed, 2009).

Trois fenêtres spectrales sont principalement utilisées en télédétection spatiale :

- ❖ Le domaine du visible
- ❖ Le domaine des infrarouges (proche IR, IR moyen et IR thermique)
- ❖ Le domaine des micro-ondes ou hyperfréquences (pas abordé ici, même si elles ont une importance considérable en télédétection RADAR notamment).

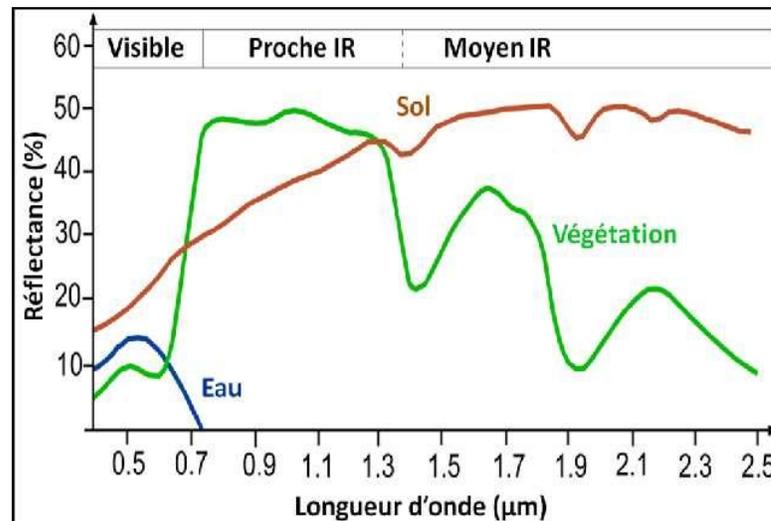


**Figure (12):** Le spectre électromagnétique (Bonn et Rochon, 1992)

➤ **Signature spectrale**

La variation de la longueur d'onde induit des réponses très diverses aux mécanismes d'absorption, de transmission et de réflexion, comme illustré dans la figure ci-dessous. Ces réponses dépendent à la fois de la composition de la cible observée et de la longueur d'onde spécifique du rayonnement. En mesurant l'énergie réfléchie ou émise par la cible à différentes longueurs d'onde, nous pouvons élaborer la signature spectrale propre à un objet. La comparaison des signatures de

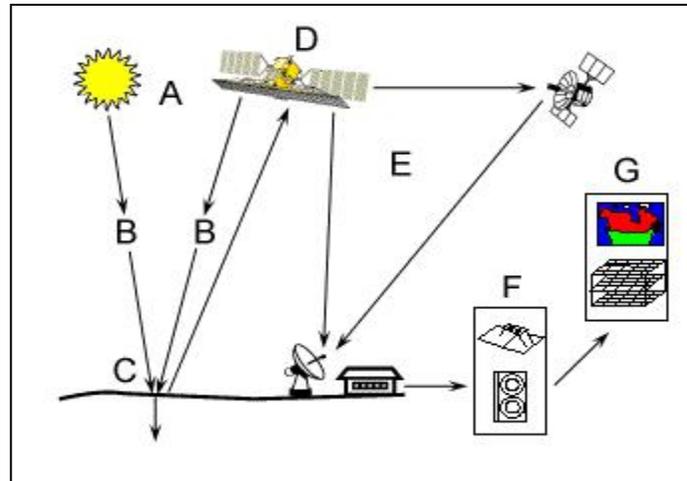
différents objets nous permet de les différencier, ce qui pourrait ne pas être possible si nous les évaluions uniquement avec une seule longueur d'onde. À titre d'exemple, bien que l'eau et la végétation puissent présenter une signature spectrale similaire dans les longueurs d'onde visibles, elles deviennent souvent distinguables dans l'infrarouge (figure 13).



**Figure (13) :** Signatures spectrales, caractéristiques des principaux types de surfaces.

### 1.2.1.3. Processus de la télédétection

Lorsque le rayonnement entre en contact avec un objet, sa trajectoire est altérée en raison des propriétés triples de la matière, qui peut absorber, réfléchir ou transmettre la lumière. Ces trois caractéristiques de la lumière au moment de son interaction avec la matière sont mesurées en relation avec la quantité de lumière incidente frappant l'objet (lumière incidente) : l'absorptance ( $\alpha$ ), la réflectance ( $\rho$ ), et la transmittance ( $\tau$ ), où  $(\alpha) + (\tau) + (\rho) = 1$ . En télédétection spatiale, des instruments appelés capteurs sont employés pour enregistrer le rayonnement électromagnétique réfléchi ou émis par les objets situés à la surface terrestre ou dans l'atmosphère (voir figure 14).



**Figure (14) : Le processus de télédétection**

- ✓ **Source d'énergie (A)** : une source d'énergie ou d'illumination pour illuminer la cible.
- ✓ **Interaction rayonnement et atmosphère (B)**: Pendant son trajet entre la source d'énergie et la cible, le rayonnement entre en interaction avec l'atmosphère.
- ✓ **Interaction avec la cible (C)** : Une fois parvenu à la cible, l'énergie interagit avec sa surface.
- ✓ **Enregistrement de l'énergie par le capteur (D)** : L'énergie émise ou diffusée par la cible est capturée puis enregistrée.
- ✓ **Transmission, réception et traitement (E)** L'énergie capturée par le capteur est transmise à une station de réception, où elle est convertie en images numériques ou photographiques.
- ✓ **Interprétation et analyse (F)** : Une interprétation visuelle ou numérique de l'image traitée est essentielle pour extraire les informations souhaitées concernant la cible.
- ✓ **Application (G)** : Il s'agit d'utiliser les informations extraites de l'image afin de mieux comprendre la cible, de révéler de nouveaux aspects ou de contribuer à résoudre un problème spécifique (Kluser, 2000).

#### 1.2.1.4. Vecteur et capteur

Les vecteurs sont des dispositifs artificiels positionnés en orbite autour de la Terre pour répondre aux exigences de la télédétection, des communications et de la télémétrie. Grâce à leur trajectoire orbitale, les plates-formes spatiales permettent une couverture répétitive et continue de

la surface terrestre. En télédétection spatiale, un capteur représente un instrument qui collecte l'énergie radiative provenant de la scène observée et produit un signal électrique correspondant et mesurable.

## **2. Technologie d'imagerie numérique**

Une image satellite est une représentation visuelle, vue du dessus, d'une zone étendue de la Terre. Ce type d'image est capturé par un satellite en orbite autour de la planète. Contrairement aux images obtenues à l'aide d'un appareil photo ou dessinées sur du papier, une image satellite est une forme numérique, traitée par des outils informatiques à partir des signaux transmis par le satellite. Elle se présente donc comme une mosaïque de pixels ou unités de surface, référencés en lignes et en colonnes, formant une grille régulière couvrant toute la surface balayée par le capteur. Chaque pixel contient une compilation d'informations codées par les valeurs des comptes radiométriques et les coordonnées en pixels (Canny, 1986).

### **2.1. Résolutions d'image numérique**

En télédétection, la résolution des images est souvent définie par la dimension de la zone couverte par un pixel. Cela représente la capacité d'un système optique à distinguer des signaux qui sont proches spatialement ou spectralement. La capacité à recueillir des informations par télédétection nécessite la prise en compte minutieuse de quatre types de résolution :

#### **2.1.1. Résolution spatiale**

La résolution spatiale d'un capteur est aussi définie par la distance minimum entre deux sources ponctuelles permettant de les distinguer. Elle correspond à la dimension du plus petit élément (Pixel) discernable à la surface terrestre (Erus, 2008). Cette résolution est exprimée en mètres ou en kilomètres et mesure la dimension d'un pixel. Par exemple, pour une image avec une résolution de 5 mètres, chaque pixel représente une superficie de 5 x 5 mètres au sol. Ainsi, une résolution spatiale élevée permet la distinction d'objets de petite taille tels que des bâtiments ou de petites parcelles agricoles, tandis qu'une résolution grossière (de plusieurs centaines de mètres à plusieurs kilomètres) ne permettra pas de discerner des objets de petite taille. Plus la résolution augmente, plus la taille des images croît également, ce qui ralentit les processus et nécessite un matériel plus performant, entraînant une hausse du coût de l'image (SAHEB,

2007). Il est donc crucial d'adapter les choix d'images à la problématique de recherche plutôt que l'inverse. On distingue en fonction des caractéristiques techniques des capteurs

- ✓ Haute résolution spatiale: 0.5 - 4 m
- ✓ Moyenne résolution spatiale 4 - 30 m
- ✓ Faible résolution spatiale 30 - 1000 m

### **2.1.2. Résolution spectrale**

La résolution spectrale décrit la capacité d'un capteur à utiliser de petites fenêtres de longueur d'onde, elle caractérise aussi la précision de la mesure radiométrique. L'eau ou la végétation par exemple ne nécessitent pas une résolution spectrale fine, à l'inverse des roches et minéraux qui réclament une résolution beaucoup plus grande. Si la résolution spectrale est trop grossière, il ne sera alors plus possible de bien différencier les différents minéraux. Les images panchromatiques sont thématiquement moins intéressantes que les images multi-spectrales. Effectivement, il est préférable d'avoir trois bandes spectrales : vert, bleu et rouge, plutôt qu'une seule bande spectrale visible. Il existe également des capteurs très perfectionnés capables de détecter des centaines de bandes spectrales, on les appelle capteurs hyper-spectraux (Gadal, 2003).

### **2.1.3. Résolution temporelle**

La résolution temporelle du système de télédétection est définie par la période de passage au nadir d'un satellite, correspondant au temps nécessaire à un satellite pour accomplir un cycle orbital complet. Cette période s'étend généralement sur quelques jours. Ainsi, plusieurs jours sont nécessaires pour qu'un satellite puisse observer à nouveau précisément la même scène depuis le même point dans l'espace (Bennour, 2015), on distingue :

- ✓ Haute résolution temporelle : < 1 - 3 jours
- ✓ Moyenne résolution temporelle : 4 - 16 jours
- ✓ Faible résolution temporelle : > 16 jours

#### **2.1.4. Résolution radiométrique**

Il s'agit de la sensibilité d'un détecteur à percevoir les variations d'intensité de l'énergie électromagnétique émise, réfléchi ou diffusée, déterminée par les divers niveaux d'intensité que le capteur peut discriminer au sein d'une bande. Une résolution radiométrique plus fine indique une plus grande sensibilité du capteur aux petites variations d'intensité de l'énergie reçue (Gadal, 2003).

### **2.2. Types de capteurs d'image**

#### **2.2.1. Photographie aérienne**

Les photographies aériennes représentent le système de télédétection le plus élémentaire et ancien employé pour l'observation de la surface terrestre. Prises à différentes altitudes depuis des ballons, avions ou satellites, ces photographies reproduisent de manière fidèle et exhaustive la surface du sol à un moment précis (Chevallier, 1971). Elles sont utilisées depuis longtemps dans divers domaines d'étude tels que la géologie, la géomorphologie, la météorologie, l'hydrologie, la biogéographie, l'agronomie, la géographie, l'archéologie, l'urbanisme, le génie civil, l'aménagement du territoire, la cartographie, la foresterie, etc.

#### **2.2.2. Drones**

Les drones sont inclus dans la catégorie des plateformes de « télédétection basse altitude » (TBA). Ces plateformes survolent le terrain à des altitudes par rapport au sol allant de quelques dizaines de mètres à plusieurs centaines de mètres, la limite de plafond de vol étant principalement déterminée par des réglementations. À des altitudes inférieures à 2 mètres, nous parlons plutôt de la télédétection proximale, qui implique l'utilisation de véhicules terrestres comme plateformes pour les capteurs de télédétection. Les drones contribuent quotidiennement à accroître la productivité des agriculteurs en fournissant des analyses de terrain en temps réel basées sur des vues détaillées du terrain, ce qui constitue une valeur ajoutée significative. Dotés de caméras et de divers capteurs, ces drones captent des images aériennes qui peuvent constituer des sources d'informations précieuses. Les images de haute qualité en temps réel permettent une reconnaissance et une classification efficaces des cultures, facilitant ainsi la surveillance de leur

croissance (Li et *al.*, 2018). Prévenir les maladies et les ravageurs qui peuvent les endommager (Rango et *al.*, 2017). En assurant une surveillance continue des cultures et en identifiant les zones nécessitant des interventions, les agriculteurs peuvent diminuer l'utilisation d'intrants tels que les pesticides et l'eau. Cette approche leur offre la possibilité de réaliser des économies tout en améliorant les rendements, en se concentrant sur les zones spécifiques nécessitant des traitements et en gérant de manière précise les quantités utilisées (Alsalam et *al.*, 2017).



**Figure (15):** Drones à rotors multiples

### 2.2.3. Satellites

Un satellite d'observation terrestre est un dispositif artificiel positionné en orbite autour de la planète. L'orbite, qui est la trajectoire tracée par un corps céleste autour d'un autre plus massif, adopte, généralement une forme quasi circulaire. D'après les informations du Centre National d'Études Spatiales d'Amérique (CNES), près de 2500 satellites orbitent autour de la Terre, générant continuellement des milliers d'images diverses, souvent spécialisées dans l'observation de la météo, des paysages, des catastrophes naturelles et de la végétation.

Les satellites d'observation de la Terre et leurs capteurs à bord présentent des caractéristiques spécifiques qui influent sur l'aptitude plus ou moins adaptée de leurs images à la cartographie des différentes composantes. Ces caractéristiques incluent les propriétés spectrales liées au domaine et au nombre de bandes spectrales (richesse spectrale), ainsi qu'à leur largeur (résolution spectrale). De plus, la fréquence d'acquisition des images (ou répétition des passages du satellite) détermine la résolution temporelle, à la résolution spatiale et à la fauchée (surface « observée »

par le capteur) (Girard et Girard, 2010). Le tableau (1) montre les caractéristiques des principales missions satellitaires pour le suivi de la végétation qui sont actuellement opérationnelles.

**Tableau (1) :** Caractéristiques des principaux satellites opérationnels en 2017, offrant des images dans le domaine optique avec des résolutions spatiales élevées, moyennes et basses, notamment pour le suivi de la végétation (Bégué et *al.*, 2018).

Satellite/Constellation	Début d'acquisition	Capteur	Bandes spectrales VIS et IR	Résolution temporelle (jour)	Résolution spatiale (m)	Fauchée (km)	Organisation(s)	Politique de données
Basse résolution spatiale (kilométrique)								
NOAA 18 ; 19	2005 ; 2009	AVHRR	R, PIR, MIR, IRT	Journalière	1100	2399	NOAA	Libre accès
Metop-A ; B	2006 ; 2012					2893	EUMESAT	
Moyenne résolution spatiale (hectométrique)								
PROBA-V	2013	VGT-P	B, R, PIR, MIR	Journalière	100 ; 300 ; 1000	2285	ESA	Libre accès
Terra ; Aqua	2000 ; 2002	MODIS	B, V, R, PIR, MIR, IRT	Journalière	250 ; 500 ; 1000	2330	NASA	Libre accès
Haute résolution spatiale (décamétrique et métrique)								
Landsat 8	2013	OLI	B, V, R, PIR, MIR	16 jours	100	185	USGS/NASA	Libre accès
		TIRS	IRT					
Landsat 7	1999	ETM+	B, V, R, PIR, MIR, IRT		15* ; 30 ; 60			
Sentinel-2A ; 2B	2015 ; 2017	MSI	B, V, R, PIR, MIR	5 jours	10 ; 20	290	ESA	Libre accès
Venus	2017	VSSC	B, V, R, PIR	2 jours	5	27,5	CNES/ISA	Projets scientifiques sélectionnés
ResourceSat-2A	2016	LISS-3	V, R, PIR, MIR	5 jours	23,5	140	ISRO	Commerciale
		LISS-4	V, R, PIR					
		AWiFS	V, R, PIR, MIR					
SPOT 6 ; 7	2012 ; 2014	NAOMI	B, V, R, PIR	Journalière	1,5* ; 6	60	CNES/Airbus	Commerciale

\* Correspond aux résolutions spatiales des bandes panchromatiques. Les autres valeurs de résolution spatiale correspondent aux bandes multi-spectrales dans les domaines visible (VIS) et Infra-Rouge (IR) : Bleu (B), Vert (V), Rouge (R), Proche Infra-Rouge (PIR), Moyen Infra-Rouge (MIR) et Infra-Rouge Thermique (IRT).

Les images provenant des satellites présentés dans le tableau (1), réunissent les critères essentiels permettant la cartographie et la surveillance des surfaces cultivées, des pratiques agricoles et de la production agricole. Ces critères incluent, d'une part, des canaux spectraux ou bandes spectrales dans les domaines visible et infrarouge, indispensables pour le calcul des indices de végétation, et d'autre part, une résolution temporelle élevée. En outre, la durée des missions satellitaires pluriannuelles facilite le suivi des évolutions dans l'utilisation des terres agricoles et la projection des tendances de production sur plusieurs années successives. De plus, grâce à la continuité des missions satellitaires garantit une disponibilité accrue d'images d'archives, permettant ainsi la réalisation d'études rétrospectives et un suivi à long terme de ces dynamiques.

### **3. Application de l'imagerie numérique dans évaluation du rendement des cultures**

#### **3.1. Le spectroradiomètre imageur à résolution moyenne (MODIS)**

A l'échelle globale, la littérature montre que le suivi de la croissance végétale se fait à partir de capteurs tels que MODIS (Terra) ou AVHRR (NOAA), grâce à la haute fréquence temporelle de ces images à basse ou moyenne résolution spatiale (Zhang, 2020; Atzberger et *al.*, 2016; Li et Liu, 2020). À l'échelle régionale, l'utilisation d'images à résolution spatiale moyenne, telles que LANDSAT (30 m), permet également de surveiller des zones cultivées en exploitant des séries temporelles tout au long de la saison de croissance des plantes (Wang et al, 2021 ; Hénonin et *al.*, 2021 ; Xu et *al.*, 2020). Enfin, avec des images à haute résolution spatiale (décamétrique), obtenues par des satellites tels que SPOT5 (Vatin et *al.*, 2021).

MODIS (The ModerateResolution Imaging Spectroradiometer), est développé pour le suivi de la Terre, de l'océan et de l'atmosphère. Cette technologie d'imagerie combine les capacités du radiomètre à ultra haute résolution avancé (AVHRR) et du LandsatThematic Mapper en intégrant des bandes spectrales dans le long et moyen infrarouge. Il propose une résolution spatiale de 250 m, 500 m et 1 km. Des canaux spectraux avec une meilleure caractérisation de l'atmosphère et des nuages sont spécialement conçus pour éliminer l'effet de l'atmosphère sur les observations de surface et fournir des mesures atmosphériques précises (Dimiceli et *al.*, 2020). Le capteur MODIS est fréquemment exploité comme source d'imagerie en télédétection pour une surveillance à l'échelle globale. Sa gratuité, vaste gamme de bandes spectrales et sa fréquence

régulière de révision (quotidienne) en font un choix populaire. Les images MODIS comprennent 36 bandes spectrales, comprenant deux bandes avec une résolution spatiale de 250 m (bandes 1 et 2), cinq bandes de 500 m (bandes 3-7) et 29 bandes de 1 km. Les sept premières bandes sont largement exploitées pour la surveillance de la couverture terrestre et de l'utilisation des sols (Wang *et al.*, 2015).

Le capteur MODIS est installé sur les satellites Terra et Aqua de la National Aeronautical and Space Administration (NASA), qui ont été lancés en 1999 et 2002 respectivement. Le satellite Terra suit une orbite polaire héliosynchrone, avec des passages à 10h30 (en descente) et 22h30 (en montée). De même, le satellite Aqua constitue une paire d'observations supplémentaire à 01h30 (en descente) et à 13h30 (en montée) heure solaire (Giglio *et al.*, 2006).

### **3.2. Les indices de végétation**

Un indice se définit comme une variable « synthétique, numérique qui caractérise l'intensité ou l'extension d'un phénomène trop complexe pour être décomposé en un nombre maîtrisable de paramètres » (Caloz *et al.*, 1993). Les indices de végétation sont des mesures quantitatives qui indiquent la vigueur de la végétation (Campbell, 1987). Ce sont des transformations spectrales qui utilisent deux bandes ou plus dans le but d'optimiser l'influence des caractéristiques de la végétation, et permettre des comparaisons spatiales et temporelles fiables de l'activité photosynthétique terrestre et des changements structuraux au niveau de la canopée (Huete *et al.*, 2000). L'intérêt de ces indices réside dans les applications telles que la prévision du rendement (Lykhovyd, 2020 ; Vozhehova *et al.*, 2020 ; Nagy *et al.*, 2021), l'évaluation des besoins nutritifs des plantes (en détectant la concentration d'azote dans les plantes) (Reyniers *et al.*, 2004), la détection des maladies et des infections, l'évaluation des besoins de la plante en eau et la lutte contre les mauvaises herbes (Ranjitha *et al.*, 2014). Les plus utilisés sont le PVI (Perpendicular Vegetation Index) de Richardson et Wiegand (1977). Le TSAVI (Transformed Soil Adjusted Vegetation Index) de Baret et Guyot (1991). Le NDVI (Normalized Difference Vegetation Index) de Rouse *et al.*, (1974).

### **3.2.1. L'indice de végétation par différence normalisée (NDVI)**

L'indice de végétation par différence normalisée, le plus connu et le plus populaire des indices de végétation, fut développé dans les années 1970, à partir des images issues du capteur MSS des missions Landsat (Rouse *et al.*, 1974). Le plus couramment désigné sous le nom de NDVI (Normalized Difference Vegetation Index), selon son acronyme anglais, est calculé à partir de deux bandes spectrales, le rouge (R) et l'infrarouge (IR), selon la formule suivante :

$$\text{NDVI} = (\text{IR} - \text{R})/(\text{IR} + \text{R}). \quad (1)$$

Il varie de 0, indiquant une surface désertique, à 1 correspondant aux forêts denses. Cet indice est particulièrement accessible à mettre en œuvre, et son utilisation est soutenue par une littérature abondante. Il est déjà largement utilisé avec les images satellitaires, notamment pour évaluer la condition, le stade de développement et la biomasse des cultures. Il est utilisé également pour la prévision du rendement (Caturegli *et al.*, 2016 ; Tattaris *et al.*, 2016). Son utilité pour décrire le couvert végétal repose sur le fait que celui-ci, d'une part, absorbe de manière sélective (bien que non exclusive) l'énergie lumineuse dans les longueurs d'onde du rouge pour la photosynthèse, et d'autre part, réfléchit intensément cette énergie dans le proche infrarouge, en fonction de la structure inter-cellulaire du matériel végétal photo-synthétisant (Gausman, 1985).

### **3.3. Méthodes d'évaluation du rendement des cultures à partir des images satellitaires**

La télédétection spatiale offre la possibilité d'obtenir des informations sur la surface terrestre, y compris les couverts végétaux, à différentes échelles spatiales et temporelles. Ces informations peuvent couvrir à la fois des parcelles spécifiques et des régions entières, avec une fréquence adaptée pour surveiller la dynamique des cultures. Dans le monde entier, plusieurs systèmes opérationnels de suivi des cultures et de prévision des rendements agricoles utilisent divers degrés d'intégration des données satellites. La précision des prévisions de rendement agricole dépend largement du niveau de détail des informations relatives à l'utilisation des terres agricoles et à l'état des cultures (Commission Européenne, 2010). L'utilisation de la télédétection pour le suivi de la croissance des cultures a débuté aux Etats Unis d'Amérique, entre 1974 et 1977, notamment avec le programme LACIE (Large Area CropInventory Experiment) développé conjointement par l'USDA et la NASA (National Aeronautics and Space Administration). Ce

programme fut l'un des premiers à démontrer les capacités opérationnelles de la télédétection pour la prévision de la production du blé (MacDonald and Hall, 1980).

### **3.3.1. Prévision du rendement des cultures avec des méthodes statistiques**

Les premiers modèles utilisés pour la prévision du rendement à grande échelle sont généralement des modèles statistiques (Thompson, 1969). Les modèles statistiques destinés à la prévision du rendement établissent des relations entre les rendements antérieurs et des indicateurs des conditions agro-environnementales, tels que la température, les précipitations, et les indices de végétation. L'application de ces modèles à la prévision du rendement des cultures repose sur l'hypothèse que les relations entre les conditions agro-environnementales de l'année de prévision sont comparables aux liens historiques utilisés dans l'élaboration du modèle. De plus, il suppose que le rendement anticipé dans les conditions de l'année de prévision est expliqué par le même ensemble de variables indépendantes que celui utilisé lors de l'élaboration du modèle à partir des données des années précédentes (Basso *et al.*, 2013).

D'autre part, les données de télédétection dans les différentes longueurs d'ondes fournissent des informations sur la végétation et sur l'état du sol à grande échelle spatiale. La régression entre le NDVI à partir de début de développement et le rendement final des cultures est le modèle statistique le plus largement utilisé à partir de données de télédétection (Basso et Liu, 2019). De nombreux chercheurs ont calculé le NDVI à partir de données MODIS pendant une période de 2 à 3 mois avant la récolte, puis ont mis au point des modèles de régression linéaire pour prévoir le rendement du maïs, du blé et du riz (Johnson, 2014; Skakunet *al.*, 2016; Wang *et al.*, 2014 ; Adeniyi *et al.*, 2020 ; Panek et Gozdowski, 2021 ; Nagy *et al.*, 2021).

Les modèles statistiques, bien qu'ils sont efficaces dans leur simplicité d'utilisation, présentent des limitations en termes d'informations fournies, notamment lorsqu'ils sont appliqués à des conditions qui diffèrent de celles pour lesquelles ils ont été initialement développés, particulièrement dans des années aux conditions extrêmes (Basso et Liu, 2019).

En termes de méthodologie de régression, on peut distinguer deux approches, l'approche linéaire et l'approche non-linéaire, qui sont détaillées ci-dessous.

#### **3.3.1.1. Les approches linéaires**

L'approche linéaire pour la prévision du rendement des cultures repose sur l'utilisation de la régression linéaire multiple entre les variables indépendantes (telles que les données agrométéorologiques, de télédétection, etc.) et la variable dépendante, à savoir le rendement final en grain. Plusieurs études ont adopté cette méthode pour la prévision des rendements de cultures dans diverses régions du monde. L'équation associée à cette approche est représentée par l'équation (Eq.2) suivante :

$$y_t = \sum_{i=1}^n a_i x_i + \epsilon \quad (2)$$

Avec

- $y_t$  représente le rendement prédit.
- $x_i$  représente le prédicteur ou la variable indépendante (température, précipitation NDVI, ...) pendant l'année  $t$ ,
- $(a_i)$  sont les coefficients de régression associés à chaque variable indépendante
- $(\epsilon)$  est le terme d'erreur ou de résidu (il capture la variation non expliquée par les prédicteurs).

### **3.3.1.2. Les approches non-linéaires**

Le machine learning est un champ d'études de l'intelligence artificielle qui a émergé avec les technologies de « big data » et la haute performance de calcul (Liakos *et al.*, 2018). Entre autres définitions, le machine learning est définie comme le domaine scientifique qui donne aux machines la capacité d'apprendre sans être strictement programmées (Samuel, 1959). Une machine intelligente assimile et dérive des connaissances des données, élaborant un modèle pour effectuer des prédictions. Par conséquent, le processus d'apprentissage automatique est subdivisé en trois étapes principales. Ces étapes comprennent le prétraitement des données, la construction du modèle, et enfin la généralisation : un processus permettant de prédire la sortie en fonction des entrées sur lesquelles l'algorithme n'a pas été précédemment formé (Sharma *et al.*, 2021).

L'un des bénéfices majeurs du machine learning réside dans sa capacité à résoudre de façon autonome des problèmes non linéaires en tirant parti de données provenant de diverses sources (Chlingaryan *et al.*, 2018). Plus récemment, l'application de techniques d'apprentissage automatique s'est étendue à la prédiction du rendement des cultures, englobant des méthodes

telles que la régression multi variée, les arbres de décision, l'extraction de règles d'association et les réseaux de neurones artificiels. Une caractéristique fondamentale de ces modèles réside dans leur capacité à traiter la production agricole (rendement des cultures) en tant que fonction implicite des variables d'entrée (gènes et composantes environnementales). Cette fonction peut être hautement non linéaire et complexe. Au cours des récentes années, diverses techniques d'apprentissage automatique ont été appliquées à la prédiction des rendements de différentes cultures, démontrant ainsi des résultats performances. Liu et coll. (2001), ont utilisé un réseau neuronal avec un réseau caché de couche pour prédire le rendement du maïs à l'aide de données d'entrée sur le sol, la météo, et le management. Marko et coll. (2016), ont proposé des histogrammes pondérés régression pour prédire le rendement de différentes variétés de soja, qui a démontré des performances supérieures aux classiques algorithmes de régression.

# **PARTIE II :**

# **Matériel et Méthodes**

# **Chapitre I :**

# **Présentation des zones**

# **d'étude**

### **Introduction**

Dans le cadre de cette recherche, nous avons adopté une approche multidimensionnelle en analysant les rendements à trois échelles spatiales distinctes : la wilaya de Sétif, la commune de Sétif et des parcelles expérimentales spécifiques de l'ITGC-Sétif. Cette division en échelles géographiques permet une exploration approfondie des interactions entre les facteurs climatiques, les caractéristiques du sol et d'autres variables locales qui impactent la productivité agricole.

Chaque échelle offre un niveau de résolution unique, allant de l'analyse régionale à une échelle plus fine au niveau des parcelles expérimentales. Cette approche permet de capturer les variations spatiales des conditions environnementales et d'identifier les facteurs clés qui influent sur le rendement des cultures céréalières dans les régions d'étude.

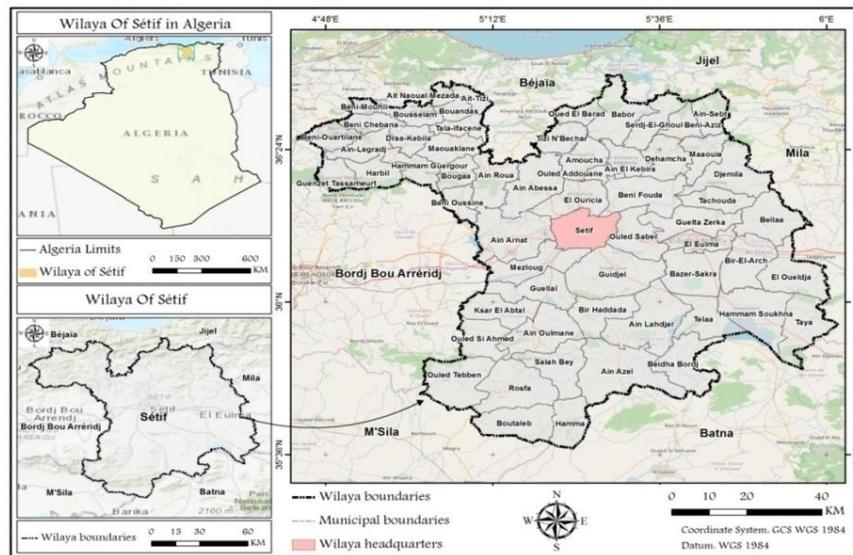
### **1. La wilaya de Sétif**

#### **1.1. Cadre géographique**

La wilaya de Sétif se situe à 300 km au Sud-Est de la capitale Algérienne, couvrant une superficie de 6504 km<sup>2</sup>, se trouve dans la région des hauts plateaux à une altitude qui varie entre 900 et 1300 m, avec des coordonnées géographiques de 36°.15' en latitude et 5.43' en longitude.

Les frontières de la wilaya sont délimitées par:

- Au Nord : par les wilayas de Bejaïa et Jijel.
- A l'Est : par la wilaya de Mila.
- Au Sud : par les wilayas de Batna et M'sila.
- A l'Ouest : par la wilaya de Bordj-Bou-Argeridj.



**Figure (16) :** Situation géographique de la wilaya de Sétif.

## 1.2. Caractéristiques agro-pédologiques

Elle se divise en trois zones principales sur le plan topographique:

- **Zone montagneuse:** se compose de trois zones montagneuses, la première est la zone montagneuse de la région nord (Babor), qui s'étend sur plus de 100 kilomètres et atteint une altitude maximale de 2004 mètres. Ensuite, on trouve les montagnes de Bibans, dont l'extrémité orientale couvre la partie nord-ouest de la wilaya. Enfin, les montagnes du Hodna, s'étendent vers le Sud et le Sud-Ouest, où l'altitude atteint à Djebel Boutaleb 1890 m. Cette zone occupe plus de 40% de la superficie de la Wilaya.
- **Zone des hautes plaines :** C'est une vaste zone qui couvre 50% de la superficie totale de la Wilaya, qui est relativement plate, avec des altitudes allant de 900 à 1200 mètres. Malgré leur apparence plate, les mamelons de montagne (Djebel Youcef et Braou) ont été découverts dans cette région.
- **Zone de dépression Sud et Sud Est :** Située dans les régions sud et sud-est de la Wilaya, l'altitude est rarement supérieure à 900 m. Cette zone, qui est en réalité plate, couvre 10% de la zone de Wilaya et se distingue par la présence de "chots" ou dépressions salines.

### **1.3. Les ressources hydrauliques**

La wilaya compte 55 oueds dont le plus important est Oued Boussellam qui rejoint la Soummam, Oued Agrioun, Oued El Kebir, Oued Bouslah et Oued Rhumel. Ces réseaux hydrographiques alimentent les barrages d'Ain Zada et les retenues collinaires de la région. Le nombre des sources d'eau s'élève à 21 104 sources dont 124 sources d'eau potable (ANIREF, 2023).

### **1.4. Les potentialités agricoles**

La Wilaya de Sétif possède un potentiel en sol assez important, une superficie agricole de 360.968,07 hectares et dont 23.994,64 hectares de terres irriguées. Cette agriculture repose essentiellement sur la céréaliculture (Ministère de l'intérieur, 2016).

Selon le Service de la Planification et des Statistiques budgétaires de la wilaya de Sétif (DPSB, 2021), la répartition des terres agricoles se présente comme suit :

- La Superficie Agricole Totale (SAT) s'élève à 459 848 hectares.
- La Superficie Agricole Utile (SAU) représente 365 191 hectares, équivalant à 79% de la SAT.
- La Superficie Irriguée atteint 35 000 hectares, soit 0.10% de la SAU.
- La Superficie Forestière couvre 102 313 hectares.
- Les Maquis s'étendent sur une superficie de 21 130 hectares.

De plus, la Steppe occupe environ 8,2% de la superficie totale de la wilaya.

### **1.5. Production végétale**

Les principales productions végétales se concentrent essentiellement sur la céréaliculture, fourragères, bien que dans une moindre mesure (Tableau 2). L'arboriculture, notamment la culture de l'olivier et du figuier, représente une richesse notable de la zone montagneuse.

**Tableau 2** : Principaux production végétales (DPSB, 2021)

<b>Produits</b>	<b>Quantités produites</b>
Blé dur	758 868 Qx
Blé tendre	35 000 Qx
Orge	69 352 Qx
Avoine	9 989 Qx
Fourrages	1 052 439 Qx
Légumes secs	5 869 Qx
Tabac	26 355 Qx
Maraîchage	1 627 915 Qx
Viticulture	291 Qx
Fruits à noyaux et à pépins	277 049 Qx
Olives	185 267 Qx

Les projets visant à transférer les eaux des barrages d'Erraguen et d'IghilEmda vers ceux de Draa Dis et de Maouane, ainsi que la création de vastes périmètres d'irrigation, devraient significativement accroître la superficie agricole utile irriguée. Cela est envisagé dans le cadre de la politique nationale d'atteindre l'autosuffisance agro-alimentaire, en favorisant une production céréalière optimale.

### **1.6. Production animale**

Dans le domaine de l'élevage, les données de l'année en question révèlent les effectifs suivants :

- Le nombre d'ovins s'élève à 414 569 têtes.
- La population bovine compte 143 710 individus.
- Les caprins sont au nombre de 55 929.
- En ce qui concerne l'aviculture, on dénombre 1 456 batteries.

**Tableau 3:** Principaux production animales (DPSB, 2021)

<b>Produits</b>	<b>Quantités produites</b>
Viande ovine	68 327 QX
Viande bovine	102 000 QX
Viande caprine	12 119 QX
Poulet de chair	231 874 QX
Dinde	30 445 QX
Lait	239 824 103 litres
Œufs	831 092 103 unités
Miel	2 484 QX
Laine	5 797 QX
Peaux et cuir	27 767 QX

### **1. 7. Caractéristiques climatiques**

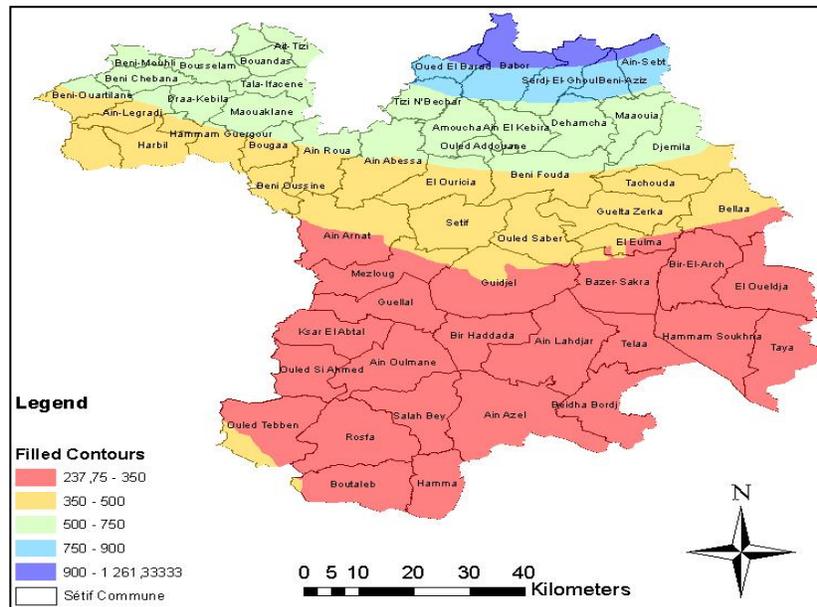
Les conditions climatiques exercent une influence significative sur le développement des plantes. L'émission des feuilles et l'évolution des divers stades dépendent fortement de la quantité de pluie et la somme des températures journalières depuis le semis jusqu'à la fin du cycle de la plante.

Le climat de la wilaya de Sétif est continental et semi-aride, avec des hivers rigoureux et des étés chauds. En hiver, la température est généralement faible et il y a beaucoup de pluie, ce qui provoque une érosion évidente, notamment dans les régions du Nord.

#### **1.7.1. Pluviométrie**

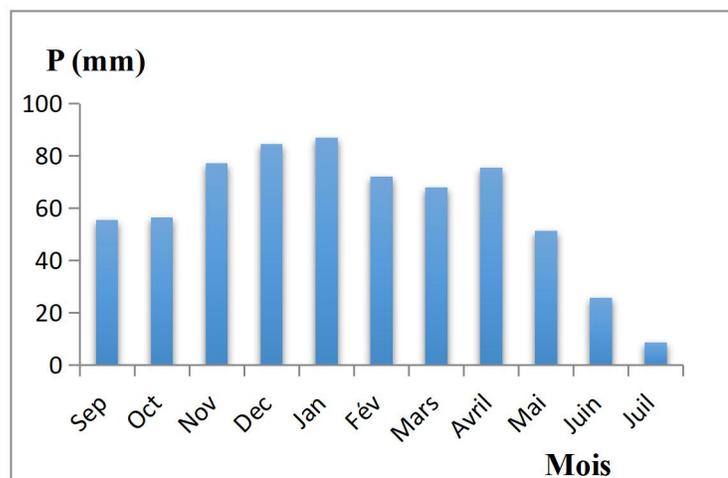
La principale activité agricole dans les régions arides et semi-arides, telles que les hautes plaines Sétifiennes, est la culture des céréales sous régime pluvial. Cependant, la production des céréales reste limitée par les conditions climatiques, en particulier les précipitations qui se caractérisent par une variabilité spatio-temporelle importante avec des moyennes annuelles de 700mm, 400mm et 200mm/an, respectivement au Nord, centre et Sud (figure 17 ).

En été, les précipitations sont quasiment nulles, alors qu'en hiver, la wilaya enregistre de fortes précipitations, avec parfois des chutes de neige, bénéfiques pour le sol jusqu'en Mars et parfois même jusqu'en Avril.



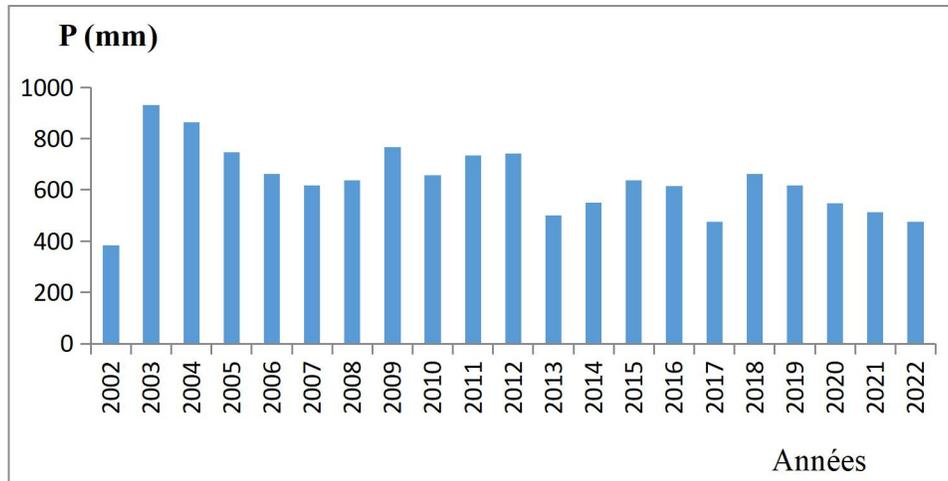
**Figure (17):** Carte de répartition des pluies de la wilaya de Sétif (Rouabhi, 2017)

La figure (18) présente la moyenne de la répartition mensuelle des précipitations dans la wilaya de Sétif entre 2002 et 2022, soulignant une variation irrégulière d'un mois à l'autre. Le cumul annuel des précipitations pour la période de Septembre à Juin, entre 2002 et 2022, atteint 663 mm, et la pluviométrie la plus élevée a été enregistrée pendant le mois de Janvier atteignant 87.07 mm, et la quantité la plus faible a été enregistrée en Juillet avec 8.7 mm.



**Figure (18) :** Précipitations moyennes mensuelles au niveau de la wilaya de Sétif entre 2002 et 2022 (Glam, 2022).

La figure (19) présente les précipitations annuelles dans la wilaya de Sétif entre 2002 et 2022, soulignant une variation significative d'une année à l'autre. L'année 2003 représente la plus humide avec 931.38 mm, par contre l'année 2002 la plus sèche avec 384.27 mm, cette fluctuation dans les précipitations d'une année à l'autre peut avoir un impact sur la croissance et le développement des cultures céréalières.

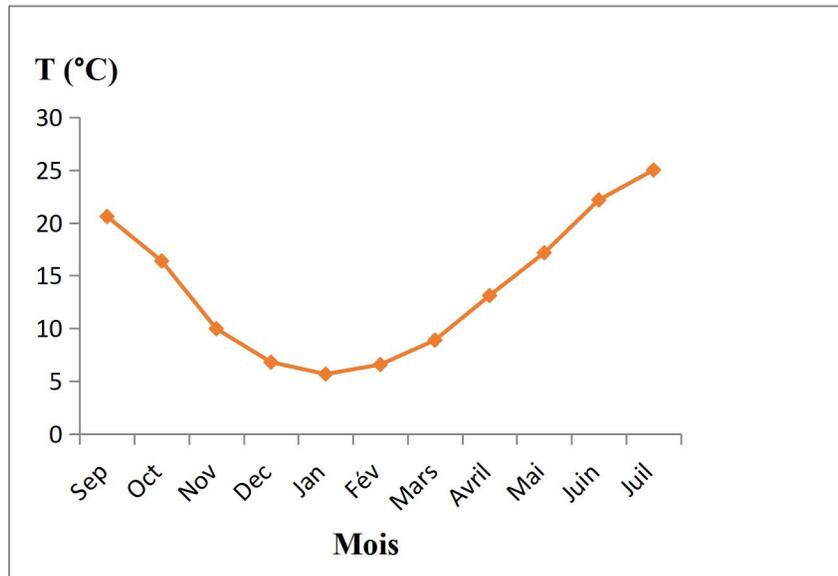


**Figure (19) :** Précipitations annuelle au niveau de la wilaya de Sétif entre 2002 et 2022 (Glam, 2022).

### 1.7.2. Température

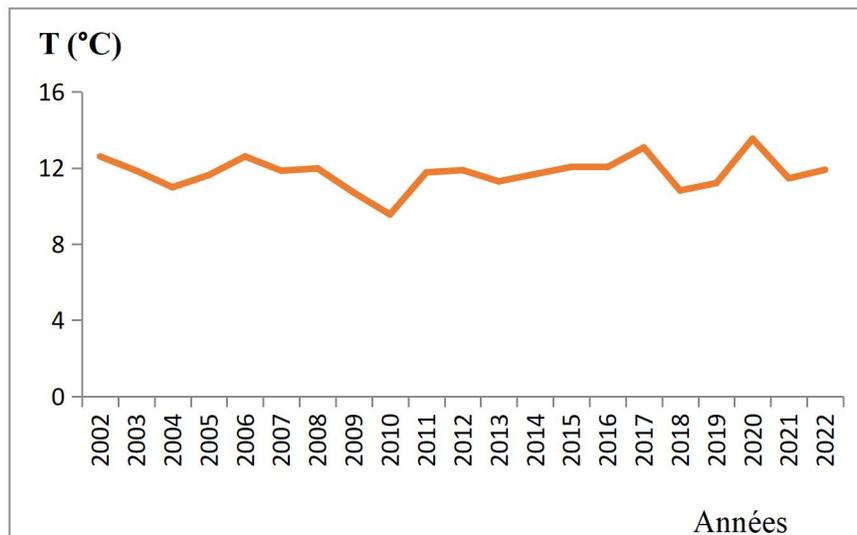
Parmi les divers stress environnementaux, les basses températures printanières, notamment les gelées tardives, ainsi que les températures élevées en fin de saison, qui sont responsables des dommages les plus importants dans la céréaliculture. Contrairement à l'accumulation des précipitations, la température moyenne mensuelle présente une distribution bimodale, basse au cours des stades végétatifs tels que le tallage, la montaison et le gonflement et élevée pendant la floraison, remplissage du grain et la phase d'installation de la culture (semis - début tallage)

Les températures moyennes des campagnes 2002 jusqu'à 2022 varient de 5.66 °C pour le mois de Janvier qui correspond au mois le plus froid à 25.8°C en Juillet correspondant au mois le plus chaud (figure 20).



**Figure (20):** Températures moyennes mensuelles au niveau de la wilaya de Sétif entre 2002 et 2022 (Glam,2022)

D’après la figure (21), on remarque qu’il ya des variations significatives des températures d’une année à l’autre. En particulier, l’année 2020 ressort comme la plus chaude avec une température moyenne de 13.53 °C, tandis que l’année 2010 fraîche, avec une moyenne de 9.55 °C.

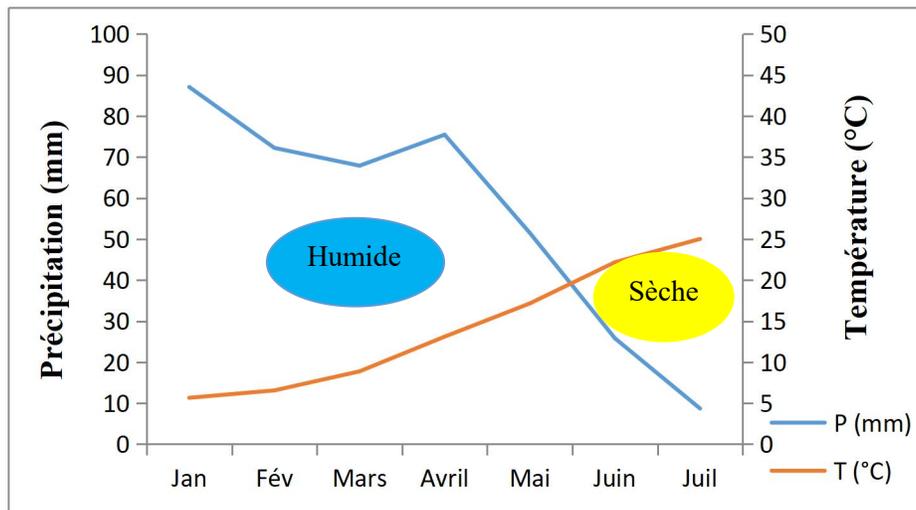


**Figure (21):** Températures moyennes annuelles au niveau de la wilaya de Sétif entre 2002 et 2022 (Glam, 2022)

### 1.7.3. Diagramme ombrothermique

Les températures sont reportées à l'échelle double des précipitations selon la définition de Gaussen (1958). Une période sèche est une période humide pendant laquelle les précipitations totales du mois sont inférieures ou égales au double de la température ( $P \leq 2T$ ). Cette relation permet d'établir un diagramme Ombro-thermique représenté dans les figures

D'après le diagramme ombrothermique (figure 22), réalisé à partir des moyennes des températures et des précipitations des vingt campagnes (de 2002 à 2022) de la wilaya de Sétif, on observe que la période humide s'étale du mois de Janvier jusqu'au début du mois de Mai, tandis que la période sèche commence au début du mois de Mai.



**Figure (22) :** Diagramme ombrothermique de la wilaya de Sétif entre 2002 et 2022 (Glam, 2022)

## 2. L'échelle communale (la commune de Sétif)

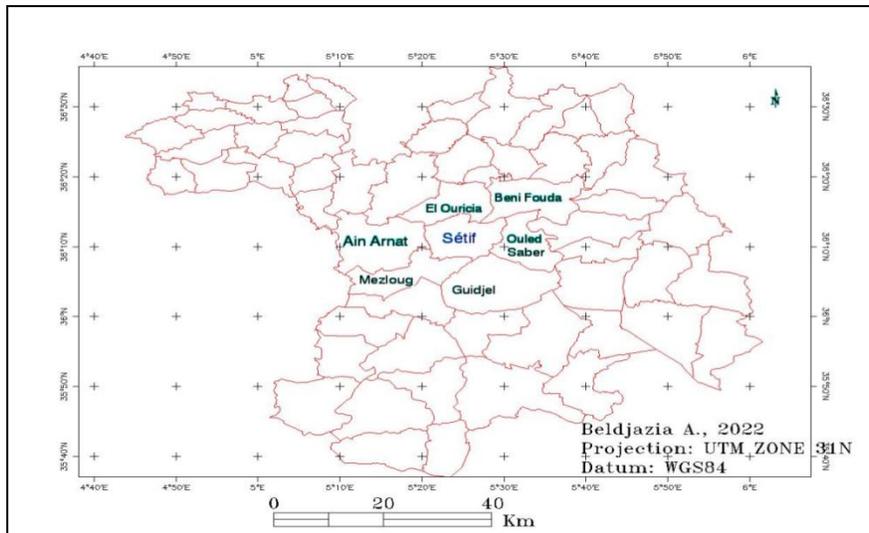
### 2.1. Situation géographique

La commune de Sétif est située dans l'Est Algérien dans la wilaya de Sétif, dans la région des hauts plateaux, à une distance de 300 Km de la capitale Alger, ayant pour coordonnées géographiques : 36°11'N pour la latitude et 5°24' E pour la longitude, avec une altitude de 1100 m, elle couvre une superficie estimée à 127.3 Km<sup>2</sup>. La commune de Sétif se trouve au cœur de la partie centrale de la Wilaya entourée par:

- Au nord : par la commune de Ouricia.

- A l'est : par la commune de Ouled saber.
- A l'ouest : par les communes de Mezloug et Ain Arnet.
- Au sud : par la commune de Guedjel.

Elle occupe une position prédominante parmi les villes des hauts plateaux de l'Est, ce qui fait d'elle leur capitale.

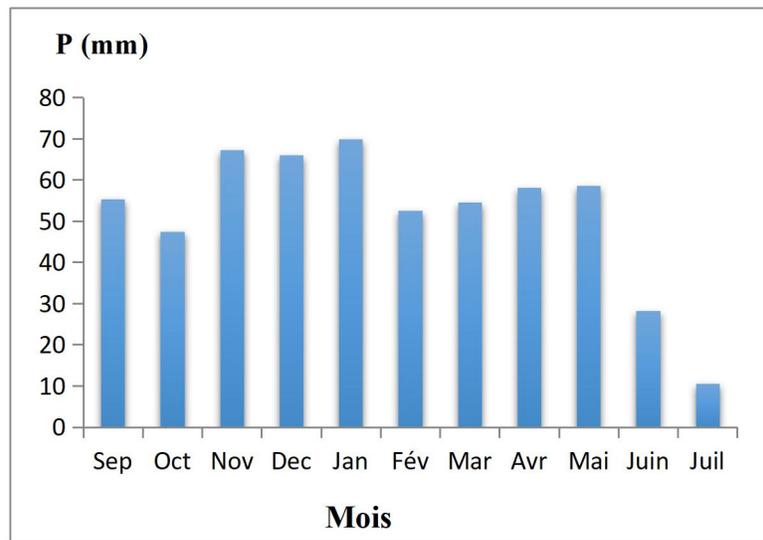


**Figure (23):** Localisation géographique de la commune de Sétif (Beldjazia, 2022)

## 2.2. Caractéristiques climatiques

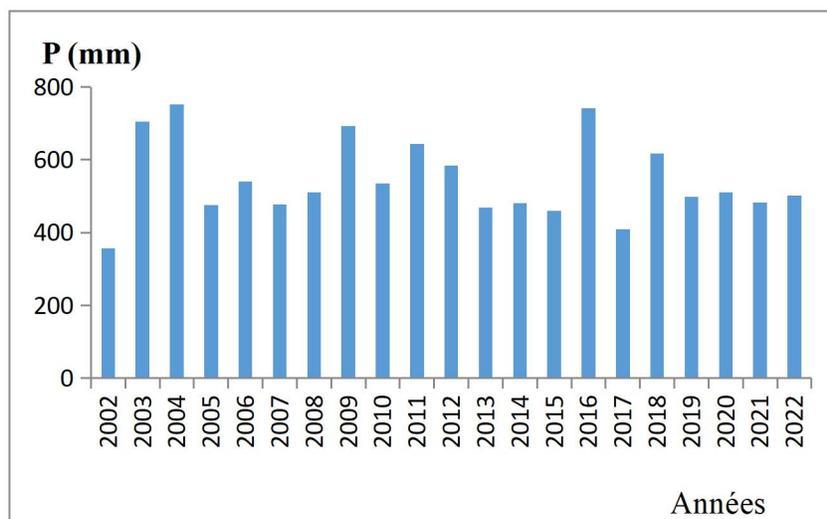
### 2.2.1. Pluviométrie

La répartition mensuelle moyenne des précipitations dans la commune de Sétif observée entre 2002 et 2022, présente une variabilité significative d'un mois à l'autre. Le cumul annuel des précipitations pour la période de Septembre à Juin au cours de ces années, atteint 568.48 mm, Notamment, le mois de Janvier a enregistré la plus forte pluviométrie, atteignant 69.92 mm, tandis que la quantité la plus faible a été enregistrée en Juillet avec 10.6mm (figure 24).



**Figure (24) :** Précipitations moyennes mensuelles enregistrées au niveau de la commune de Sétif de 2002 à 2022 (Glam, 2022)

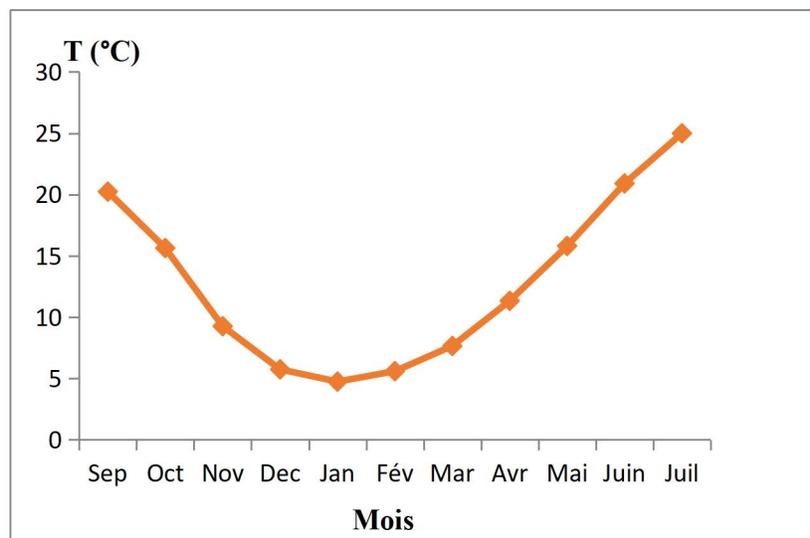
La figure (25) présente les précipitations annuelles dans la commune de Sétif entre 2002 et 2022, soulignant une variation significative d'une année à l'autre. L'année 2004 représente la plus humide avec 751.13 mm, tandis que l'année 2002 la plus sèche avec 355.9 mm, cette fluctuation dans les précipitations d'une année à l'autre peut avoir un effet sur la croissance et le développement des cultures céréalières dans cette région.



**Figure (25):** Précipitations annuelle au niveau de la commune de Sétif entre 2002 et 2022 (Glam, 2022).

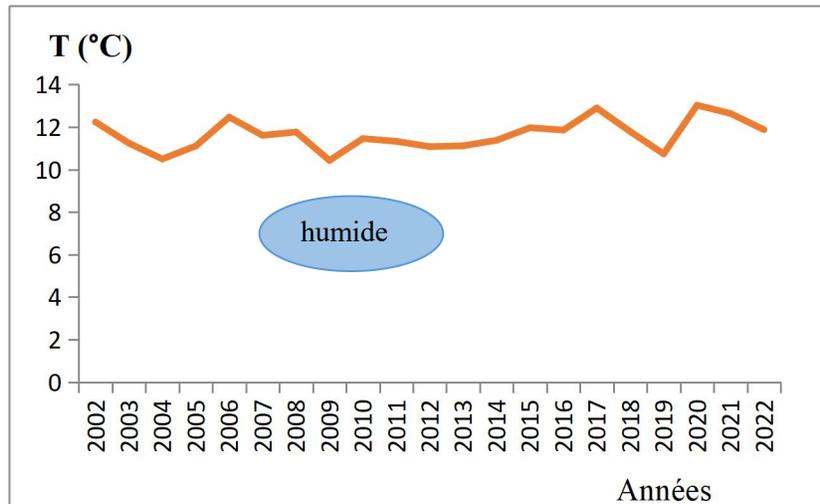
### 2.2.2. Température

L'analyse des températures moyennes mensuelles au niveau de la commune de Sétif de 2002 à 2022 (figure 26) indique que : les périodes les plus froides, allant de décembre à mars, se caractérisent par les températures les plus basses, tandis que les autres mois affichent des températures généralement plus élevées. La température moyenne minimale est enregistrée au mois de Janvier atteignant 4.73°C, tandis que la température moyenne maximale est enregistrée au mois de Juillet s'élevant à 24.99 °C.



**Figure (26):** Températures moyennes mensuelles au niveau de la commune de Sétif  
Entre 2002 à 2022 (Glam, 2022)

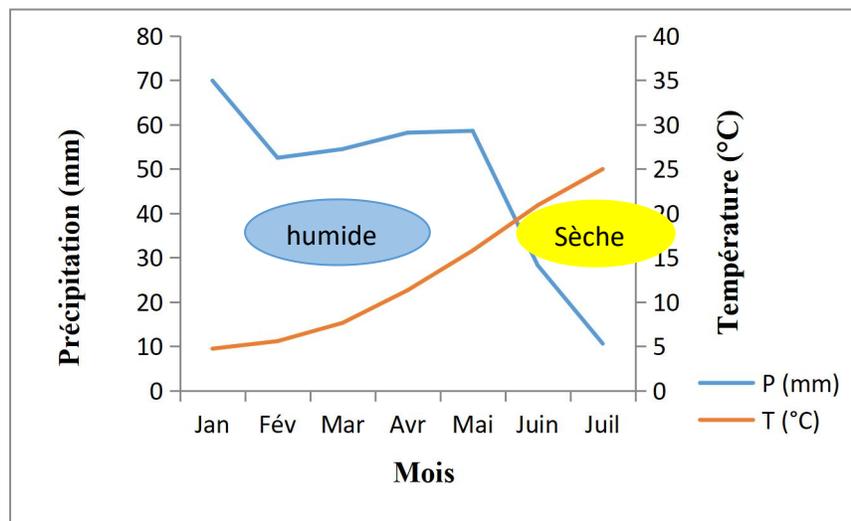
D'après la figure (27), on remarque qu'il ya des variations significatives des températures d'une année à l'autre. En particulier, l'année 2020 ressort comme la plus chaude avec une température moyenne de 13.03 °C, tandis que l'année 2009 fraîche, avec une moyenne de 10.43 °C.



**Figure (27):** Températures moyennes annuelles au niveau de la commune de Sétif  
Entre 2002 à 2022 (Glam, 2022)

### 2.2.3. Diagramme ombrothermique

D’après le diagramme ombrothermique (figure 28), réalisé à partir des moyennes des températures et des précipitations des vingt campagnes (de 2002 à 2022) de la commune de Sétif, on observe que la période humide s’étale du mois de Janvier jusqu’au mois de mi-Mai et la période de sécheresse commence au mi-mai.



**Figure (28) :** Diagramme ombrothermique de la commune de Sétif entre 2002 et 2022 (Glam, 2022).

### 3. L'échelle parcellaire

#### 3.1. Situation géographique

L'expérimentation a été conduite durant les campagnes agricoles 2019/2020, 2020/2021 et 2021/2022, dans les parcelles expérimentales de l'Institut Technique de Grandes Cultures (ITGC) de Sétif (figures 29, 30), le site située à 5 km au sud -ouest de la ville de Sétif. Situé à une latitude de 36°10'17" Nord, une longitude de 5°21'55" Est et à une altitude de 1080 m (Google Earth Pro, 2023). Ce site expérimental est localisé dans la zone centrale des hautes plaines, propice à la culture céréalière. Le climat du site est de type méditerranéen, continental, appartenant à l'étage bioclimatique semi-aride, caractérisé par deux saisons distinctes, la prédominance des pluies en saison froide, un été chaud et sec (Baldy, 1986; Chennafi et al., 2006).



**Figure (30) :** Localisation géographique des parcelles expérimentales de l'ITGC.

**3.2. Caractéristiques physico-chimiques du sol**

Afin de déterminer les principales propriétés qui caractérisent le sol des parcelles expérimentales, une série d'analyses a été réalisée dans le laboratoire de l'université de Ferhat Abbas -Sétif 1- sur un échantillon issu d'un mélange de cinq prélèvements du sol à l'aide d'une tarière suivant les diagonales de la parcelle à une profondeur de 30 cm. Les résultats sont rapportés dans le tableau 4.

**Tableau4** : Résultats d'analyses physico-chimiques du sol des parcelles expérimentales de l'ITGC-Sétif-

<b>Types d'analyse</b>	<b>Paramètres</b>	<b>Résultats d'analyse</b>	<b>Méthodes d'analyse</b>
<b>Analyse granulométrique</b>	Argile (%)	40.12	Pipette de ROBINSON
	Limon (%)	30.68	
	Sable (%)	29.2	
<b>Analyse chimique</b>	Calcaire total (%)	21.25	Calcimètre de BERNARD
	pH eau	8.35	pH mètre
	pH KCL	7.59	pH mètre
	CE (ms/cm)	0.152	Conductimètre
	CaCO <sub>3</sub> (%)	37	JORET-HEBERT
<b>Analyse biochimique</b>	Carbone (%)	1.23	Méthode Anne
	Matière organique (%)	2.13	MO= C %*1.72

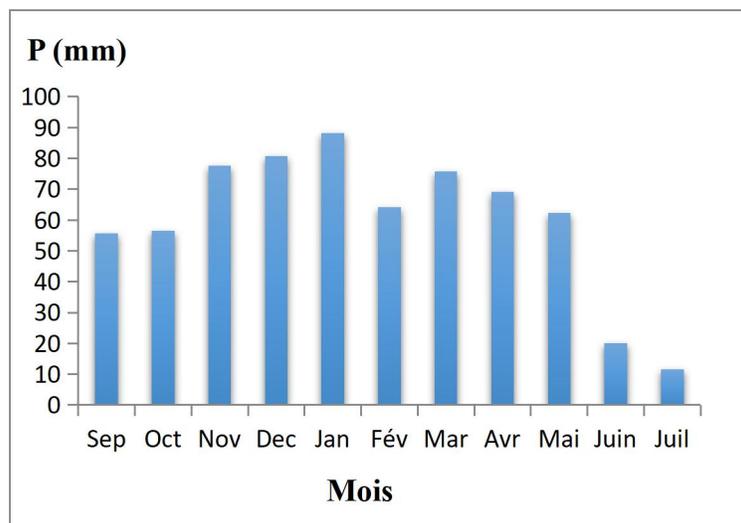
D'après le triangle de Hénin (1969), le sol présente une texture argilo-limono-sableuse. Son pH est classé comme basique selon les normes MDLC (1993), Cette caractéristique basique est typique des sols calcaires, comme le confirme un taux de calcaire mesuré à 21.25% (entre 15% et 30%) (Baize, 1988). Selon les normes établies par différents auteurs, le sol est non salé car la

conductivité électrique est de 0.152 (<2) (USSSL, 1954). De plus, sa teneur en matière organique est modérément faible, égale à 2.13% (entre 1.15% et 3%) (Clavet et viellemin, 1986).

### 3.3. Caractéristiques climatiques

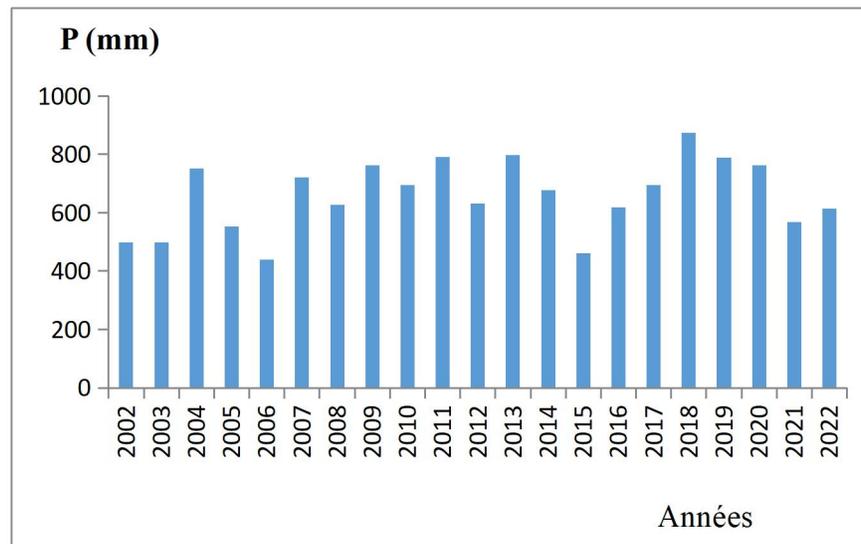
#### 3.3.1. Pluviométrie

D'après la figure (31), on remarque que la moyenne de la répartition mensuelle des précipitations dans les parcelles expérimentales de l'ITGC de Sétif entre 2002 et 2022, soulignant une variation irrégulière d'un mois à l'autre. Le cumul annuel des précipitations pour la période de septembre à Juin, entre 2002 et 2022, atteint 662.19 mm, et la pluviométrie la plus élevée a été enregistrée pendant le mois de Janvier atteignant 88.27 mm.



**Figure (31) :** Précipitations moyennes mensuelles enregistrées au niveau d'ITGC (Sétif) entre 2002 et 2022(Glam, 2022)

La figure (32) présente les précipitations annuelles dans les parcelles expérimentales de l'ITGC de Sétif entre 2002 et 2022, soulignant une fluctuation significative d'une année à l'autre. L'année 2018 représente la plus humide avec 872.55 mm, tandis que l'année 2006 la plus sèche avec 439.24 mm, cette fluctuation dans les précipitations d'une année à l'autre peut avoir un effet sur le rendement des céréales dans cette région.

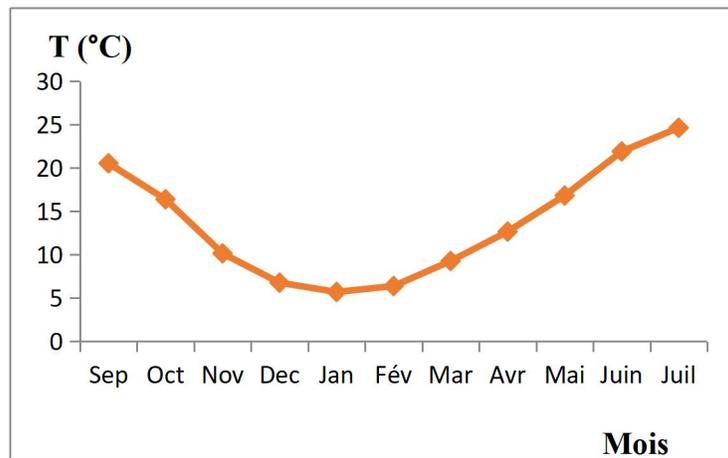


**Figure (32) :** Précipitations annuelle enregistrées au niveau d’ITGC (Sétif) entre 2002 et 2022 (Glam, 2022).

### 3.3.2. Température

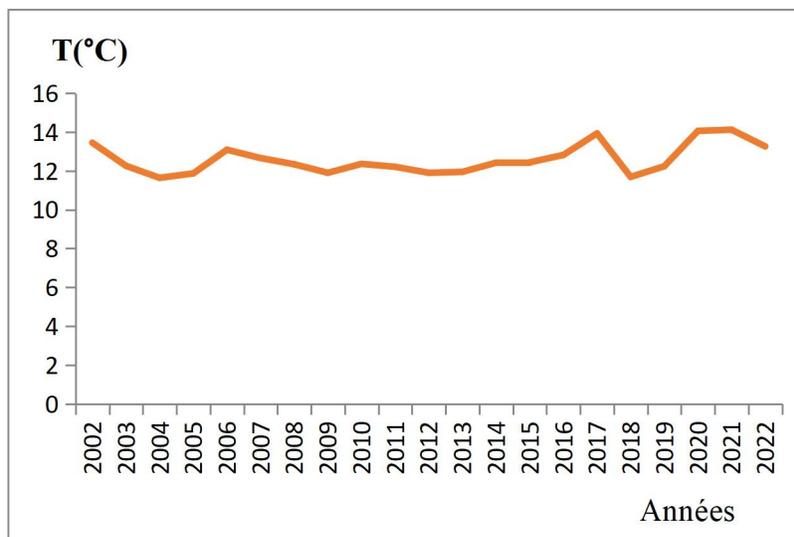
En plus les températures, l’humidité joue un rôle essentiel pendant la période de croissance des céréales. Des conditions excessivement sèches ou humides peuvent entraîner des conséquences négatives sur leur développement.

D’après la figure 33, les températures moyennes des campagnes 2002 jusqu’à 2022, enregistre une diminution à partir du mois d’Octobre (environ 16.37°C) jusqu’au mois de Février (environ 6.36°C), puis elle remonte progressivement, vers le mois de Mars jusqu’au mois de Juillet (environ 24.62°C). Cette augmentation de la température coïncide avec la phase de remplissage des grains.



**Figure(33) :** Températures moyennes mensuelles au niveau d’ITGC (Sétif) entre 2002 et 2022 (Glam, 2022).

D’après la figure (34), on remarque qu’il ya des variations significatives des températures d’une année à l’autre au niveau des parcelles de l’ITGC. En particulier, l’année 2021 ressort comme la plus chaude avec une température moyenne de 14.12 °C, tandis que l’année 2004 ressort comme la plus fraîche, avec une moyenne de 11.65 °C.



**Figure (34):** Températures moyennes annuelles au niveau d’ITGC (Sétif) entre 2002 et 2022 (Glam, 2022).

### 3.3.3. Diagramme ombrothermique

D'après le diagramme ombrothermique (figure 35), réalisé à partir des moyennes des températures et des précipitations des vingt campagnes (de 2002 à 2022) au sein des parcelles expérimentales de ITGC (Sétif), on observe que la période humide, caractérisée par des précipitations plus importantes, s'étale du mois de Novembre jusqu'au mois de Mai, et la période sèche commence au mi Mai.

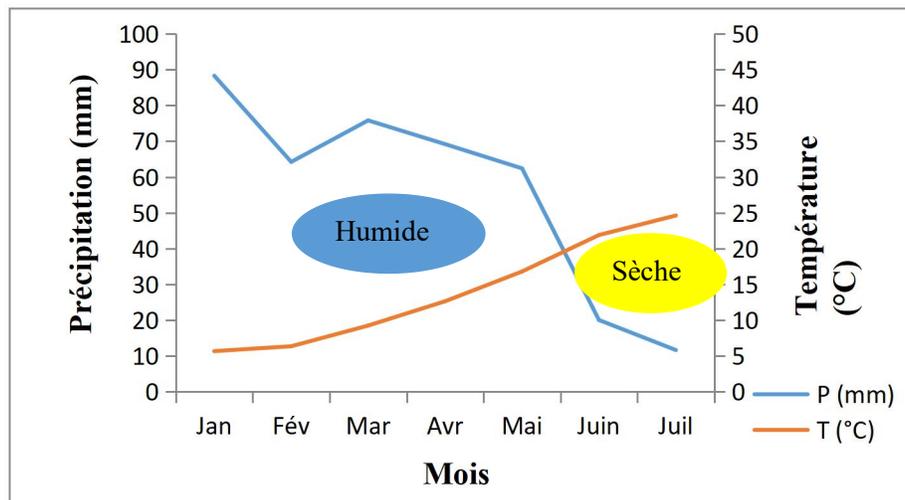


Figure (35): Diagramme ombrothermique au niveau de l'ITGC-Sétif- entre (2002-2022).

#### 4. Le matériel végétal

Le matériel végétal utilisé dans les parcelles expérimentales est le suivant :

- **Blé dur: Variété Boussellam**

Sélectionné localement, à partir du matériel génétique fourni par CIMMYT/ICARDA, qui vient de connaître un début d'adoption dans les régions de Tiaret et Sétif, c'est une variété qui se caractérise par une haute paille, avec des épis blancs, barbe noire-grise, demi-lâche, une longueur et une robustesse considérables. Sa hauteur moyenne varie de 90 à 100 cm, c'est une variété résistante aux maladies cryptogamiques, mais, il est recommandé de traiter les semences avec des fongicides pour renforcer cette résistance au froid, à la verse, à la sécheresse (Baghem, 2012).

- **Orge: Variété Fouara**

Fouara, est une orge à 6 rangs, elle est le résultat d'une sélection effectuée au sein de la station de l'Institut Technique des Grandes Cultures (ITGC) de Sétif, cette sélection a été réalisée

à partir du matériel en ségrégation provenant de l'ICARDA (International Center of Agricultural Research in Dry Areas) en 1991/92 (Benmahammed et *al.*, 2005). C'est une variété très productive et de bonne stabilité, Fouara a été inscrite au catalogue national depuis 1997 (Benmahammed et *al.*, 2005).

### **5. Itinéraire technique et conduite de la culture**

L'itinéraire technique est une combinaison logique et ordonnée de techniques appliquées à une culture visant à optimiser son rendement.

#### **5.1. La préparation du sol**

Le travail du sol en agriculture désigne l'ensemble des opérations mécaniques effectuées sur la couche superficielle du sol, c'est une composante principale de l'itinéraire technique des grandes cultures, jouant un rôle crucial en ameublissant le sol en profondeur, favorisant ainsi le développement des plantes cultivée en:

- Facilitant la germination des graines.
- Améliorer l'état physique (aération, infiltration), ce qui favorise la circulation des nutriments dans le sol.
- Améliorer la rétention d'eau du sol.
- Eliminer les mauvaises herbes.
- Préparer le lit de semences.

Un labour profond a été effectué au niveau des stations expérimentales d'ITGC au mois de Février, représentant l'étape principale du travail du sol pour préparer le terrain pour l'installation de la culture. Cette opération, à été réalisée à l'aide d'une charrue à disques, à une profondeur de 30 cm. Suivi, le premier recroisement, comprenant deux passages croisés au cover-crop, ont été effectués au mois de mai pour briser les grosses mottes et préparer le lit de semence. Le second recroisement, suivi d'un hersage en octobre pour éliminer les mauvaises herbes. Cette pratique agricole vise à incorporer tous les résidus végétaux de la culture précédente dans le sol et à préparer le lit de semence. Le roulage à été effectué après semis est effectué à l'aide d'un rouleau croskill ou lisse pour assurer un bon contact de la graine avec le sol. Toutefois, cette opération en conditions assez humides.

### **5.2. Fumure de fond**

La fumure de fond est une pratique agricole consistant à apporter des engrais au sol avant le semis ou la plantation des cultures. Un engrais de fond, le triple super phosphate TSP à 46 % a été épandu au mois d'Octobre avant le semis à raison, de 100 kg / ha. Il vise à fournir aux plantes les éléments nutritifs nécessaires à leur croissance tout au long de leur cycle de vie.

### **5.3. Semis**

Le semis en ligne a été effectué à une profondeur de 3cm au mois 03 Novembre concernant la campagne (2019/2020), 04 Novembre pour la campagne (2020/2021), et le 07 Novembre pendant la campagne 2021/2022, à l'aide d'un semoir. La dose de semis a été calculée en fonction du poids de mille grains et de la faculté germinative des semences utilisées.

### **5.4. Fumure azotée**

L'alimentation azotée détermine non seulement le niveau de production de la culture, mais aussi la qualité de la récolte en particulier à travers le mitadinage. L'engrais de couverture a été apporté sous forme d'urée à 46 % 75 kg/ha. Cette application a été divisée en deux périodes : 1/3 au semis (au stade de la levée) et 2/3 à la fin du tallage (au stade épi 1 cm).

### **5.5. Désherbage chimique**

Le désherbage est une opération agricole cruciale visant à maîtriser les principales mauvaises herbes des cultures céréalières. Le désherbage chimique a eu lieu dès le stade 2-3 feuilles, se prolongeant jusqu'au stade de mi-tallage, en utilisant un herbicide combinant un anti-graminées et un anti dicotylédones (Zoom : Triasulfuron-Dicamba) a été utilisé à raison de 150 g/ha, complété par le Brimby (Aryloxyphénoxy-Propionates) à raison de 0.75 L/ha.

### **5.6. Récolte**

La récolte a été effectuée par une moissonneuse batteuse expérimentale. Elle a eu lieu fin juin jusqu'à début juillet.

# **Chapitre II : Données et Outils utilisées**

### **1. Objectif**

Le présent chapitre montre les différentes données et outils utilisées à l'échelle wilaya, commune et parcelles, Afin de développer des modèles de prévision des rendements des céréales, notamment le blé et l'orge, et d'acquérir une compréhension approfondie des facteurs impactant de la production céréalière, sur une période de 20 ans, de 2002 à 2022: des données satellitaires (constituées des séries temporelles de NDVI issues du capteur MODIS/TERRA), le pourcentage de couverture verte fractionnée de la canopée (FGCC) extrait de l'application Canopeo®, la biomasse verte de trois ans (2020, 2021, 2021), les données agro-météorologiques ainsi que les données du rendement de blé et d'orge. Les caractéristiques spécifiques de ces données, ainsi que les outils méthodologiques employés, sont détaillés dans les sous-sections suivantes.

### **2. Les produits satellitaires MODIS**

A l'échelle mondiale, les satellites de type « grand champ » jouent un rôle essentiel dans la surveillance des variations saisonnières et annuelles des couverts végétaux. Ils sont des outils indispensables et fournissent une source d'informations cruciale pour mieux comprendre le fonctionnement de la biosphère.

Dans le cadre de cette étude, nous avons exploité les images du spectroradiomètre imageur à moyenne résolution (MODIS) à bord des satellites TERRA de la NASA. Ce dispositif a été spécialement conçu pour l'étude et la surveillance de l'environnement dans son ensemble, englobant la biosphère, l'hydrosphère et l'atmosphère. MODIS offre une résolution spatiale de 250 m (Pour le rouge et l'infrarouge de 0.6 à 0.9  $\mu\text{m}$ ), et une bonne résolution radiométrique (12 bits) et acquière les images de la surface du globe dans 36 bandes spectrales allant du visible (405 nm) à l'infrarouge thermique (14.385  $\mu\text{m}$ ).

#### **2. 1. Choix de la source de données de télédétection**

Le capteur MODIS a été choisi, en raison de ses caractéristiques particulièrement adaptées :

- 1- La disponibilité et gratuité des données, combinée à la mise à disposition en temps quasi-réel de données ortho-rectifiées et en réflectance de surface, simplifie leur utilisation directe et minimisent les étapes de prétraitement.

2- La capacité de couverture globale avec une fréquence temporelle élevée (journalière) permet la surveillance continue du développement de la végétation à intervalles réguliers sur n'importe quelle région du monde.

3- La fourniture de produits synthétiques d'indices de végétation offre un gain de temps appréciable pour les utilisateurs, évitant ainsi les étapes de calcul de néocanaux.

4- La disponibilité de plus de 30 produits dérivés permettant de caractériser les couverts végétaux, en plus, la profondeur temporelle des archives disponibles depuis février 2000.

### **2.2. L'indice de végétation par différence normalisée ou indice de Tucker (NDVI) :**

L'Indice de Végétation par Différence Normalisée (NDVI), également connu sous le nom de «Normalized Difference Vegetation Index» en anglais, est largement reconnu comme l'un des indices végétaux les plus utilisés. Il joue un rôle crucial dans la prédiction des rendements agricoles et a été introduit par Rouse en 1974.

Le NDVI permet de caractériser l'activité photosynthétique de la surface végétale, offrant ainsi la possibilité de discriminer facilement les zones recouvertes de végétation, NDVI est construit à partir des canaux rouges (R) et proches infra rouge (NIR). La formule de calcul du NDVI est la suivante :

$$\text{NDVI} = (\text{NIR} - \text{R}) / (\text{NIR} + \text{R}) \quad (2)$$

où NIR est la réflectance de la cible dans le canal Proche infrarouge (PIR) et R sa réflectance dans le Rouge (R).

Cet indice est sensible à la vigueur et à la quantité de la végétation. Ces valeurs varient entre -1 et +1, les valeurs négatives correspondant aux nuages, l'eau et à d'autres surfaces non végétales. Les roches et le sol nu présentent également des indices de réflexion proches de zéro. Les fortes valeurs (>0.3) indiquent systématiquement une surface végétale, plus ces valeurs sont élevées, plus le couvert végétal est dense. À l'inverse les valeurs les plus faibles (<0.1) caractérisent les types de couverts plus clairsemés, ces caractéristiques font que le NDVI est couramment utilisé en tant qu'indicateur indirect de la biomasse (Banari et *al*, 1995).

Le point fort de cet indice réside dans sa capacité de distinguer même entre les espèces végétales, car il est directement lié à l'activité photosynthétique des plantes et par conséquent, à

leur capacité d'absorption énergétique. Il agit comme un indicateur de la biomasse chlorophyllienne des plantes.

### **2.3. Données MODIS-NDVI**

Les séries temporelles du NDVI pour les trois zones d'étude (les parcelles d'ITGC, commune et wilaya de Sétif), sont obtenues de la base de données du Global Agricultural Monitoring (GLAM) système (<https://glam1.gsfc.nasa.gov/>), hébergé par l'USDA et la NASA (National Aeronautics and Space Administration). Le système GLAM a été développé dans le cadre d'un projet de surveillance mondiale de l'agriculture. Cette initiative vise à évaluer de manière régulière les prévisions mondiales de l'économie agricole, production et les conditions affectant la sécurité alimentaire mondiale d'une manière impartiale et en temps opportun.

Ces jeux de données NDVI proviennent du produit MOD09Q1 version 6, du capteur MODIS du satellite Terra, qui fournit une estimation de la réflectance spectrale de surface des bandes 1 et 2 du spectroradiomètre imageur à résolution moyenne, corrigée pour les conditions atmosphériques telles que les gaz, les aérosols et la diffusion de Rayleigh. Les bandes de réflectance de surface de 250 mètres (m) sont produites avec deux couches de qualité. Pour chaque pixel, une valeur est sélectionnée parmi toutes les acquisitions au cours de la période composite de 8 jours des indices de végétation NDVI.

Pour obtenir les valeurs NDVI au niveau des parcelles de l'ITGC, nous avons utilisé le masque de culture (GFSAD30-2015) développé par le Projet "NASA Global Food Security-Support Analysis Data", qui a une résolution spatiale de 30 m (<https://croplands.org>) (USGS, 2021). Afin de garantir des données de haute qualité, les informations recueillies ont subi des corrections radiologiques, atmosphériques et géométriques. Ces mesures ont été prises dans le but d'améliorer la précision et la fiabilité des données.

## **3. Calcul de la fraction de couverture verte de la canopée (FGCC)**

### **3.1. L'application "Canopeo"**

Canopeo® est une application mobile d'analyse d'images développée en 2015 à l'Oklahoma State University, aux États-Unis, en utilisant le langage de programmation MATLAB de Mathworks. Cette application innovante exploite la caméra du téléphone pour analyser et classer tous les pixels d'une image ou d'une vidéo en fonction des valeurs de couleur dans le

système rouge-vert-bleu (RVB). Le résultat de cette analyse se présente sous la forme d'une image binaire, où les pixels blancs représentent ceux qui satisfont aux critères de sélection (représentant la canopée verte), tandis que les pixels noirs correspondent à ceux qui ne répondent pas à ces critères (absence de canopée verte). La fraction de couverture végétale varie de 0 (absence totale de couverture végétale verte) à 1 (couverture végétale entièrement verte à 100%) (Patrignani et Ochsner, 2015).

En utilisant des photographies de la culture, l'application Canopeo® calcule la fraction de couverture verte de la canopée (FGCC) dans une image, cette analyse est applicable à toutes les cultures. Il convient de noter que cette application est la plus téléchargée sur le Google Play Store (figure 36), parmi les différentes options proposées pour l'analyse de la couverture végétale verte.



**Figure (36) :** Canopeo® application

### **3.2. Fonctionnement de l'application**

Il suffit de :

- Prendre une photo à une distance d'environ 50 cm au-dessus de la végétation
- L'application offre trois options pour l'entrée d'image, offrant à l'utilisateur la possibilité d'analyser la couverture du sol à partir d'une photo nouvellement prise, d'une image provenant de la galerie, ou encore à partir d'une vidéo.
- Elle effectue le calcul du pourcentage de couverture verte (FGCC), puis compare l'image couleur initiale avec l'image traitée afin de vérifier si les paramètres d'analyse par défaut captent correctement les pixels verts. Sur l'image traitée, les pixels blancs correspondent aux pixels verts présents dans la photographie d'origine.

- L'application offre la possibilité d'ajuster le degré de précision de l'analyse, permettant ainsi la possibilité de modifier la quantité de pixels considérés comme étant verts.

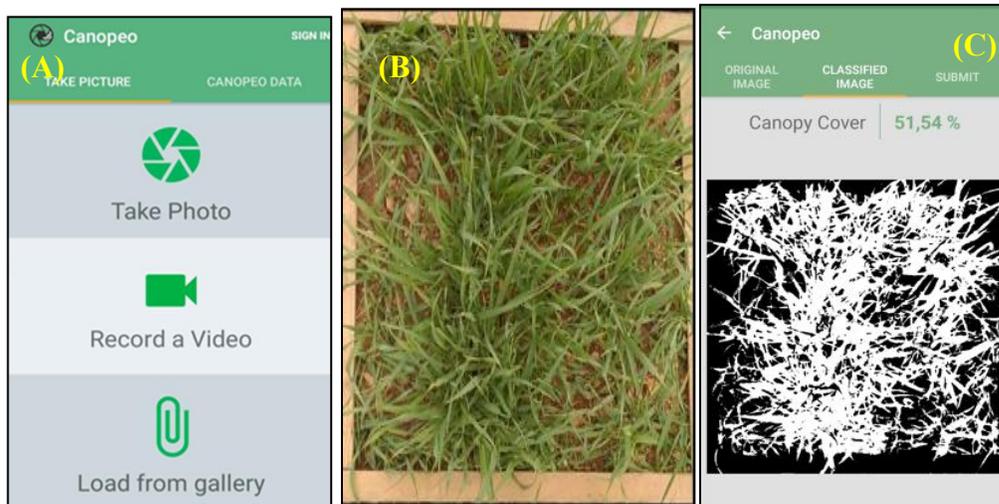
En outre, l'application offre diverses fonctionnalités supplémentaires, notamment:

- La géolocalisation de la photo.
- La saisie des paramètres de la culture tels que le type de culture, la date d'implantation et la hauteur de végétation.
- L'enregistrement des traitements effectués dans le module "Canopeo Data".
- La visualisation des traitements sur une carte.
- Enfin, pour les utilisateurs disposant d'un compte et connectés à l'application, la possibilité de sauvegarder les résultats pour un accès ultérieur.

### **3.3. Calcul du FGCC au niveau des parcelles expérimentales**

Sur le terrain, nous avons utilisé un cadre quadrant et un monopode pour garantir le même emplacement et le même recadrage des photos. En se positionnant au-dessus du point d'échantillonnage et en tenant la caméra à bout de bras à une hauteur de 1.2 m, une image perpendiculaire au sol a été capturée, en suivant les directions principales (0°, 90°, 180° et 270°).

Il est important de noter que la caméra Samsung S20 a été utilisée pour prendre les photos de végétation en mode automatique et avec une résolution de 64 MP. Pour garantir plus de précision, nous avons calculé la couverture fractionnée du couvert vert (FGCC) à partir de quatre échantillons pour chaque culture (blé et orge), puis nous avons effectué une moyenne des résultats obtenus à chaque point d'échantillonnage. Cette approche de moyenne a été adoptée pour atténuer les erreurs résultant de légères variations de position entre les photos et également pour tenir compte d'éventuelles influences de l'appareil photo du téléphone sur les résultats de l'estimation du FGCC.



**Figure (37):** Application Canopeo® (A), Image originale (B), Image de FGCC après traitement par Canopeo® (C).

#### 4. Estimation de la biomasse

L'évaluation de la biomasse verte, a été effectuée en utilisant des cadres de dimensions  $0.5 \times 0.5$  m pour collecter les plantes à proximité des zones d'échantillonnage à l'aide de faucilles. Quatre répétitions de chaque culture ont été réalisées. Ensuite, les plantes ont été placées dans des sacs identifiés, puis soumises à un séchage dans un four à circulation d'air pulsé pendant 48 heures, jusqu'à ce qu'elles atteignent une masse constante à une température de  $105\text{ }^{\circ}\text{C}$ , leurs poids secs finaux ont été enregistrés et nous avons fait la moyenne des résultats des quatre répétitions pour chaque culture (blé et orge) Voir la figure (38).



**Figure (38) :** collection de la biomasse (A), échantillons placés dans des sacs identifiés (B), échantillons soumises à un séchage (C), mesure du poids secs des échantillons (D).

### Remarque

Six collections ont été réalisées à différents stades de développement de blé et d'orge, afin de couvrir une gamme plus large de conditions végétales, notamment le début tallage, plein tallage, la montaison, l'épiaison, la floraison et le développement du grain (stade laiteux). Chaque culture était échantillonnée à quatre points différents. Les mesures de la biomasse, du FGCC et du NVDI ont été prises simultanément les mêmes jours pour garantir l'uniformité autant que possible.

### 5. Données agro-météorologiques

#### 5.1. Données météorologiques

Les données historiques des températures moyennes, minimales et maximales et les précipitations ont été obtenues à partir de la plateforme du Global Agricultural Monitoring (système de surveillance agricole mondial) (GLAM) système (<https://glam1.gsfc.nasa.gov/>), hébergé par l'USDA et la NASA à une résolution spatiale de 0.1°. Ces données obtenues à partir d'une interpolation combinant divers sources:

- La durée des nuages à sommets froids (CCD – Cold Cloud Duration) provenant de Météosat.
- Les données satellite micro-onde provenant du SpecialSensorMicrowave/Imager (SSM/I) des satellites du Programme de satellites météorologiques pour la Défense (DMSP). Ces données telles que décrites par Jobard, sont des mesures directes de l'absorption et de l'émission ou de la diffusion du rayonnement par les gouttes d'eau et les cristaux de glace à l'intérieur des nuages.
- Des données issues du radiomètre Advanced Microwave Sounding Unit (AMSU) embarqué à bord du satellite météorologique polaire NOAA-15. Ce radiomètre fournit des données sur la température et l'humidité de l'atmosphère.
- Les données de précipitations et de températures provenant du Système Global de télécommunications (GTS).

#### 5.2. L'Évapotranspiration potentielle

L'évapotranspiration mensuelle « ETP » (mm/mois) a été calculé par le logiciel "Drought Indices Calculator" (DrinC) (Tigkas et *al.* 2015) avec la méthode Hargreaves (Hargreaves et Samani 1985), en utilisant les températures minimales et maximales comme des inputs, avec l'intégration des coordonnées géographiques des régions d'études pour estimer le rayonnement (Rh). Ces valeurs ont été calculées à partir du mois Février à Juin.

$$ET_0 = 0.0023 (R_a) (T_{\text{mean}} + 17.8) (T_{\text{max}} - T_{\text{min}})^{0.5} \quad (3)$$

Avec :  $ET_0$  : évapotranspiration de référence mensuelle (mm).

$R_a$  : rayonnement solaire extra terrestre moyenne ( $W/m^2$ ).

$T_{\text{max}}$  : Température maximale mensuelle ( $^{\circ}C$ ).

T<sub>min</sub> : Température minimale mensuelle (°C).

### **5.3. L'indice de sécheresse de reconnaissance efficace (Effective Reconnaissance Drought Index) « eRDI ».**

Est un indice physique, universel pour évaluer les effets de la sécheresse sur le rendement des cultures, en particulier dans l'agriculture pluviale (Tsakiris et Vangelis, 2005). Il combine deux paramètres fondamentaux, les précipitations cumulées (P) et l'Évapotranspiration Potentielle (ETP) sur des périodes de référence définies, fournissant ainsi une mesure du bilan hydrique. Le logiciel "Drought Indices Calculator" (DrinC) (Tigkas et al. 2015) à été utilisé également pour le calcul de l'eRDI des 3 mois (Février- Mars- Avril)

$$\mathbf{eRDI} = \frac{P-PET}{P} \times 100\% \quad (4)$$

P : représente les précipitations pendant une certaine période de temps.

PET : représente l'évapotranspiration potentielle pendant la même période de temps.

### **5.4. Indice de précipitation agricole standardisé (aSPI)**

Crée par McKee et al. (1993) pour la caractérisation et le suivi de la sécheresse. Il s'agit d'un indice largement utilisé, flexible et facile à calculer, qui ne nécessite que des données sur les précipitations. Le logiciel "DrinC" (Tigkas et al. 2015) à été utilisé pour le calcul de l'aSPI. des 3 mois (Février- Mars- Avril).

$$\mathbf{aSPI} = \frac{P_i - \bar{P}}{\sigma_P} \quad (5)$$

P<sub>i</sub> : représente la précipitation observée pendant une période spécifique.

$\bar{P}$  : est la moyenne de la précipitation observée sur une période de référence.

$\sigma_P$  : est l'écart type de la précipitation observée sur la même période de référence.

### **5.5. L'humidité mensuelle du sol (%)**

Elle a été obtenue à partir de la plateforme FEWS.NET, en utilisant " l'EWX Next generation viewer" au niveau communal et wilaya. Au niveau parcellaire on a utilisé l'application "Climate engine" pour obtenir l'humidité du sol à partir les données FLDAS (Famine Early Warning

Systems Network (FEWS NET) Land Data Assimilation System) avec une résolution de 9.6 km. Ces valeurs ont été obtenues pour des profondeurs 10 cm et 100 cm de Février à Juin.

Il convient de noter, que toutes les données utilisées couvrent la période de 2002 à 2022.

### **6. Données du rendement**

Les données statistiques sur les rendements de blé et d'orge, ont été recueillies auprès de la Direction des Services Agricoles (DSA) de Sétif pour la wilaya et la commune de Sétif. Elles couvrent la période allant de 2002 à 2022, cette plage temporelle étendue capture les variations interannuelles et les tendances à long terme des rendements de blé et de l'orge, offrant ainsi une base solide pour la modélisation, ainsi pour une analyse plus robuste et précise des prévisions de rendement des cultures céréalières.

Par ailleurs, les données statistiques sur les rendements de blé et d'orge utilisées à l'échelle parcellaire ont été obtenues au auprès de l'Institut Technique des Grandes Cultures (ITGC) de Sétif.

# **Chapitre (III) : Approche de prédiction des rendements céréaliers**

### **1. Analyse et outils statistique des résultats**

#### **1.1. Analyse par régression linéaire**

L'analyse par régression linéaire représente une méthode statistique puissante utilisée pour évaluer la relation entre une variable dépendante et une ou plusieurs variables indépendantes. Nous avons adopté cette approche pour comprendre et quantifier la relation entre les rendements agricoles de blé et d'orge d'une part, et divers indicateurs environnementaux d'autre part. Elle permet d'établir une équation linéaire qui décrit au mieux la relation entre ces variables, facilitant ainsi la prédiction des rendements en fonction des paramètres étudiés, Pour cela nous avons utilisé :

##### **1.1.1. Régression linéaire simple**

Tout d'abord, nous avons utilisé la régression linéaire simple, pour examiner et quantifier la relation empirique entre les rendements de blé et d'orge et l'indice de végétation (NDVI), cette approche nous a permis d'évaluer de manière directe la corrélation entre les performances des cultures céréalières (blé et orge) au fil des années et les variations du NDVI entre 2002 et 2022. Nous pouvons représenter cette relation à l'aide de l'équation suivante :

$$Y = \beta_0 + \beta_1 X \quad (6)$$

Où  $\beta_1$  représente le coefficient de régression. Le paramètre  $\beta_0$  ne peut généralement être interprété mathématiquement que si la variable  $X$  est définie sur 0, alors  $\beta_0$  est l'estimation donnée à 0 dans  $X$ .

##### **1.1.2. Régression linéaire multiple**

On a employé la régression linéaire multiple pour examiner de manière plus approfondie les facteurs qui influent sur les variations des rendements de blé et d'orge pour une analyse plus précise. Cette méthode a impliqué la corrélation des rendements de blé et d'orge avec une combinaison de NDVI, ainsi que l'inclusion de la pluviométrie, pour mieux renforcer la précision des résultats.

**Remarque :** L'analyse statistique de prédiction du rendement à l'aide du FGCC a été employée exclusivement au niveau parcellaire, pendant une période expérimentale de trois ans, couvrant les campagnes agricoles 2019-2020, 2020-2021 et 2021-2022.

### **1.2. Analyse par régression non linéaire**

Enfin, pour affiner davantage notre analyse, nous avons adopté une approche de régression non linéaire plus complexes en utilisant le logiciel Orange3, avec les techniques de data mining suivantes : (SVM , Neural Network, KNN, Random Forest, Tree, Gradient Boosting, AdaBoost, Stochastic Gradient Descent. Cette méthode permettra de déterminer le modèle le plus précis et prédictif des rendements de blé et orge. En exploitant les capacités avancées de ce logiciel, nous avons pu identifier les relations non linéaires subtiles entre les rendements (blé et orge) et l'indices de végétation (NDVI) et les données agro-météorologiques citer précédemment dans le chapitre II.

#### **1.2.1. Le logiciel Orange3**

##### **1.2.1.1. Présentation et principe de fonctionnement du logiciel Orange3**

Orange3 est un logiciel open-source conçu pour le traitement des données et la création de modèles, développée pour l'analyse et la visualisation de données, il permet aux utilisateurs, de manipuler des données, construire des workflows d'analyse et développer des modèles de machine learning.

Pour construire des modèles de prédiction des rendements céréaliers avec Orange3, nous avons procédé les étapes suivantes:

**1. Importation de données :** Orange3 prend en charge divers formats de données, tels que CSV ou Excel, nous avons importé les fichiers qui contiennent les données des rendements céréaliers (variable cible), ainsi que les variables explicatives (données satellitaires NDVI, données agro-météorologiques) dans Orange3 pour les explorer et les analyser.

**2. Prétraitement des données :** pour nettoyer et préparer les ensembles de données, en utilisant les widgets process pour normaliser les données (entre -1 et 1) et effectuer d'autres transformations nécessaires, tels que la différenciation des variables pour éliminer les auto-corrélations. Ainsi, les données ont été divisées en deux parties : 70% pour l'entraînement des modèles (training) et 30% pour le test.

**3. Exploration des données :** une fois les données prétraitées, nous avons exploré les relations entre les variables en utilisant des widgets de visualisation tels que des diagrammes de corrélation, de dispersion, des histogrammes,... etc.

**4. Construction de modèles:** Ensuite, nous avons procédé à la construction de plusieurs modèles de prédiction, en utilisant 9 algorithmes de machine learning disponibles dans Orange3 : Linear regression, SVM , Neural Network, KNN, Random Forest, Tree, Gradient Boosting, AdaBoost, Stochastic Gradient Descent.

**5. Évaluation des modèles (Test and score):** Pour évaluer la performance des modèles on a utilisé la technique de la validation croisée et les métriques d'évaluation appropriées, telles que : R,  $R^2$ , RMSE, MSE.

- ✓ Le coefficient de corrélation (R), un paramètre clé dans l'analyse de régression. Il mesure la force et la direction de la relation linéaire entre une variable dépendante et une ou plusieurs variables indépendante. Dans une régression linéaire simple, le coefficient de régression représente la pente de la ligne de régression, indiquant combien la variable dépendante change en moyenne pour un changement unitaire dans la variable indépendante. Et dans une régression linéaire multiple, chaque variable indépendante a son propre coefficient de régression, mesurant la contribution spécifique de cette variable aux variations de la variable dépendante, tout en maintenant les autres variables constantes.
- ✓ Le coefficient de détermination ( $R^2$ ), a été utilisé pour évaluer la proportion de la variance de la variable dépendante expliquée par le modèle de régression et aussi pour mesurer le degré de relation entre les rendements céréaliers observés et prédits, il varie de 0 à 1, où 1 indique une adéquation parfaite du modèle aux données.
- ✓ L'erreur quadratique moyenne (MSE), ou Mean Squared Error en anglais a été utilisée pour évaluer la précision du modèle de prévision. Il calcule la moyenne des carrés des écarts entre les valeurs prédites par le modèle et les valeurs réelles.
- ✓ L'erreur quadratique moyenne (RMSE), ou root mean squared error en anglais, est une mesure de la racine de l'erreur quadratique moyenne (MSE), a été utilisée pour mesurer l'écart entre les valeurs prédites et les valeurs réelles.
- ✓ coefficient de détermination ajusté ( $R^2_{\text{ajust}}$ ): est une mesure statistique utilisée en régression linéaire pour évaluer la proportion de la variance d'une variable dépendante qui est expliquée par les variables indépendantes, tout en tenant compte du nombre de

prédicteurs dans le modèle. Contrairement au  $R^2$  simple, il pénalise l'ajout de variables qui n'améliorent pas significativement le modèle.

Les  $R^2$ , RMSE, MSE et  $R^2$ -ajust ont été calculés à l'aide des équations 7, 8, 9 et 10.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (8)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

$$R^2 - ajust = 1 - \left( \frac{(1 - R^2)(n - 1)}{n - k - 1} \right) \quad (10)$$

**6. Sélection de la date et les meilleures variables explicatives :** Après avoir choisi le modèle, on procède à la sélection de la bonne date de prédiction, ainsi qu'au choix des meilleures variables explicatives pour déterminer celles qui contribuent effectivement à la prédiction des rendements à l'échelle de la wilaya, communale et parcellaire.

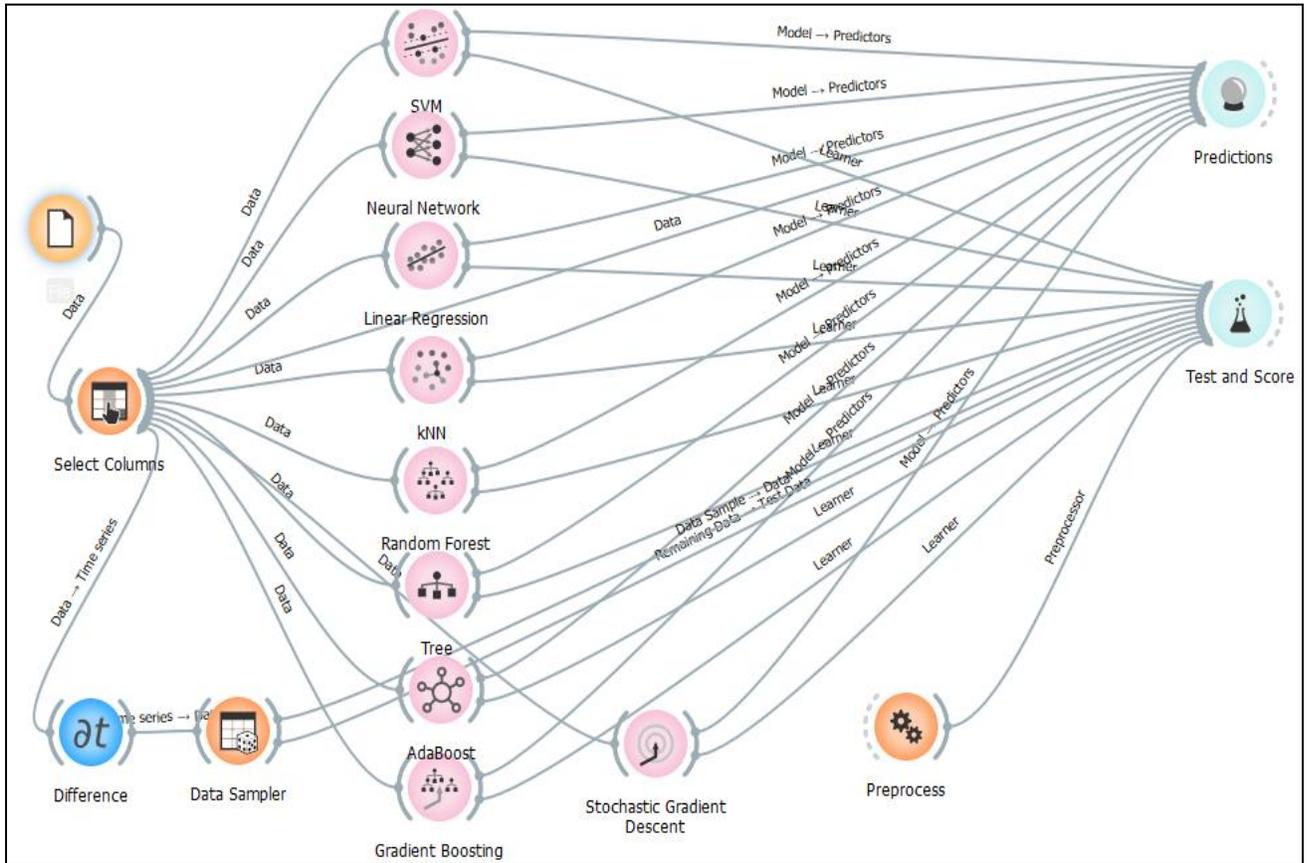


Figure (39) : Méthodologie de prévision des rendements céréaliers avec le logiciel Orange3.

# **PARTIE III : Résultats et Discussion**

# **Chapitre (I) : Analyses des paramètres agro-météorologiques**

### 1. Analyses des rendements céréaliers

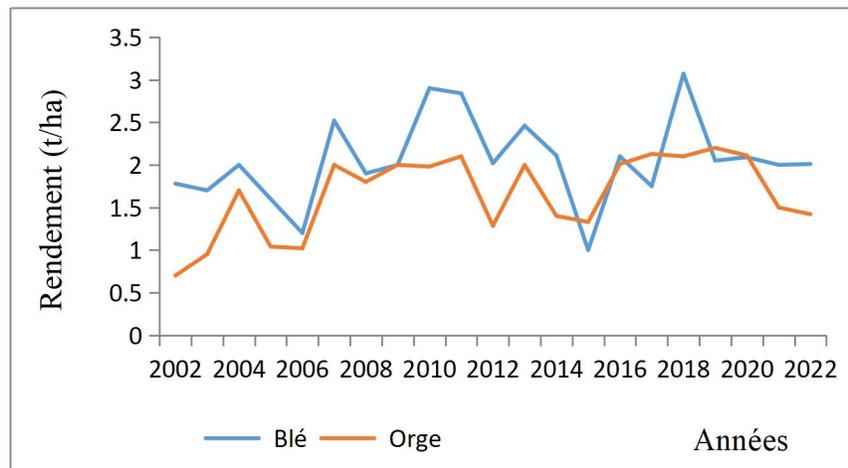
#### 1.1. Au niveau parcellaire : les parcelles expérimentales de l'ITGC

La variabilité temporelle des rendements moyens en grains de blé et d'orge dans les parcelles expérimentales de l'ITGC entre 2002 et 2022 est illustrée dans la figure (40). En 2018, le blé a affiché son rendement moyen le plus élevé avec 3.07 t/ha, tandis que l'orge a atteint son rendement moyen le plus élevé en 2019 avec 2.2 t/ha. À l'inverse, les rendements les plus faibles ont été observés respectivement en 2015 pour le blé, avec 1 t/ha, et en 2002 pour l'orge, avec 0.7 t/ha.

La moyenne des rendements pour le blé et l'orge sont 2.05 t/ha, 1.65 t/ha respectivement avec un écart type de 0.5, 0.46 pour le blé et l'orge, ce qui indique que les rendements annuels du blé et de l'orge sont relativement stables et cohérents d'une année à l'autre, avec peu de fluctuations significatives (voir tableau 5).

**Tableau (5):** tableau des statistiques descriptif des rendements de blé et d'orge au niveau des parcelles de l'ITGC entre 2002 et 2022.

	Blé (t/ha)	Orge (t/ha)
Moyenne	2.05	1.65
Minimum	1	0.7
Maximum	3.07	2.2
Ecart type	0.5	0.46



**Figure (40):** Variabilité temporelle des rendements (blé et orge) de 2002 à 2022 au niveau des parcellaire.

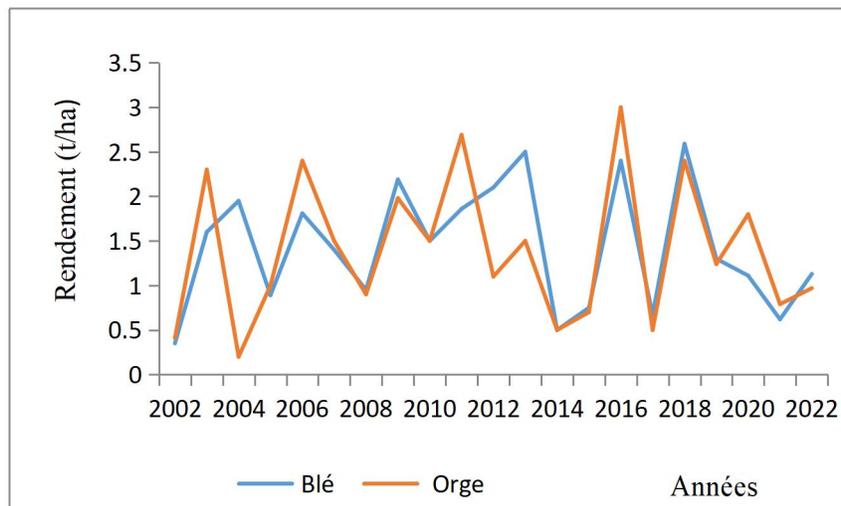
### 1.2. Au niveau de la commune de Sétif

La figure (41) montre la variation temporelle des rendements moyens en grains de blé et d'orge dans la commune de Sétif entre 2002 et 2022. En 2018, le blé a enregistré son rendement moyen le plus élevé, atteignant 2.59 t/ha, tandis que l'orge atteint son rendement moyen le plus élevé en 2016, avec 3 t/ha. En revanche, les rendements les plus bas ont été observés respectivement en 2021 pour le blé avec 0.62 t/ha, et en 2004 pour l'orge avec 0.2 t/ha.

La moyenne des rendements pour le blé et l'orge sont 1.43 t/ha, 1.39 t/ha respectivement avec un écart type de 0.69, 0.81 pour le blé et l'orge, ce qui signifie qu'il y a généralement une plus grande variation dans les rendements annuels de l'orge par rapport à sa moyenne, en comparaison à la variation des rendements annuels du blé, (voir tableau 6).

**Tableau (6):** tableau des statistiques descriptif des rendements de blé et d'orge au niveau des de la commune de Sétif entre 2002 et 2022.

	Blé (t/ha)	Orge (t/ha)
Moyenne	1.43	1.39
Minimum	0.62	0.2
Maximum	2.59	3
Ecart type	0.69	0.81



**Figure (41):** Variabilité temporelle des rendements (blé et orge) de 2002 à 2022 au niveau de la commune de Sétif

### 1.3. Au niveau de la wilaya de Sétif

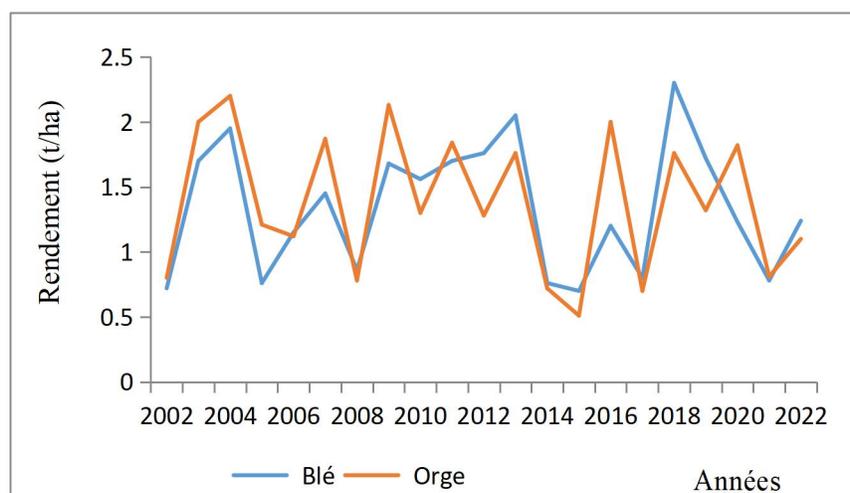
La figure (42) montre la variation temporelle des rendements moyens en grains de blé et d'orge dans la wilaya de Sétif entre 2002 et 2022. En 2018, le blé a enregistré son rendement moyen le plus élevé, atteignant 2.3 t/ha, tandis que l'orge atteint son rendement moyen le plus élevé en 2004, avec 2.2 t/ha. En revanche, les rendements les plus bas ont été observés respectivement en 2015 pour le blé avec 0.7 t/ha, et en 2015 pour l'orge avec 0.51 t/ha.

Ces fluctuations annuelles dans les rendements moyens des deux céréales sont principalement attribuables aux conditions météorologiques, telles que la variabilité des précipitations, les températures hivernales très basses et les sécheresses au printemps et au début de l'été.

La moyenne des rendements pour le blé et l'orge sont 1.34 t/ha, 1.38 t/ha respectivement avec un écart type de 0.49, 0.54 pour le blé et l'orge, indique que rendements annuels de blé et d'orge sont les relativement stables et cohérents d'une année à l'autre, avec peu de fluctuations significatives (voir tableau 7).

**Tableau (7):** tableau des statistiques descriptif des rendements de blé et d'orge au niveau des parcelles de l'ITGC entre 2002 et 2022.

	Blé (t/ha)	Orge (t/ha)
Moyenne	1.34	1.38
Minimum	0.7	0.51
Maximum	2.3	2.2
Ecart type	0.49	0.54



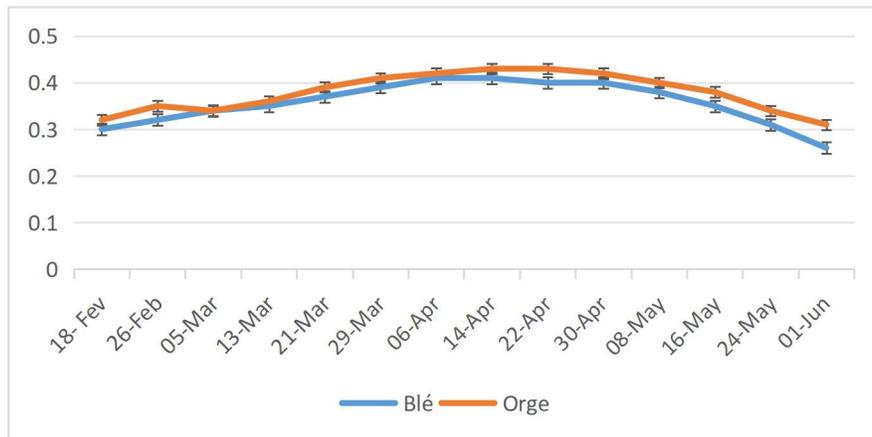
**Figure (42):** Variabilité temporelle des rendements (blé et orge) de 2002 à 2022 au niveau de la wilaya de Sétif

### **2. Analyses de la variation saisonnière et interannuelle du NDVI**

L'analyse de la variabilité saisonnière et interannuelle du NDVI tout au long du cycle de croissance des cultures de blé et d'orge dans les trois zones d'étude, révèle les résultats suivants : pendant la phase "levée- début tallage" de Février au début de Mars, ou les valeurs du NDVI sont les plus bas, correspondant à un stade précoce de développement végétatif. Ensuite, ces valeurs augmentent progressivement au fur à mesure que les parties végétatives des cultures se développent, atteignant leur apogée à la fin de la phase végétative "plein tallage" entre début et mi-Mars, cette hausse reflète la croissance et le développement des cultures durant cette période. Les valeurs du NDVI demeurant élevées jusqu'à la phase de floraison, qui se produisant entre mi-Mars et Avril, signifiant ainsi une biomasse végétale importante et un bon état de santé des cultures, c'est une phase cruciale dans le cycle de croissance, elle est déterminante pour la formation des grains. Par la suite, pendant la phase "post-floraison ou phase de maturation", une diminution graduelle des valeurs du NDVI est observée à mesure que les cultures mûrissent. Cette réduction reflète la sénescence des plantes et la redistribution des ressources vers les parties nécessaires à la reproduction et à la formation des graines. Les valeurs du NDVI atteignent leur niveau le plus bas pendant la phase de "maturité" en juin, indiquant ainsi la fin du cycle de croissance (figures 43, 44, 45).

#### **2.1. Au niveau parcellaire**

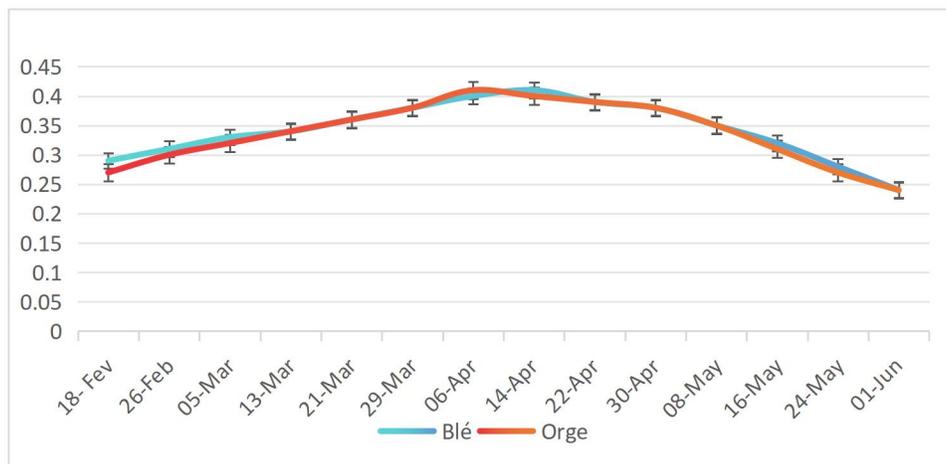
Les valeurs de NDVI variaient entre 0.212, enregistré le 1<sup>er</sup> Juin 2004 et 0.539 enregistré le 30 Avril 2007 pour toutes les années d'étude pour l'orge variaient entre 0.197 le 1<sup>er</sup> Juin 2002 et 0.537 le 30 Avril 2007 pour le blé (figure 43).



**Figure (43):**Variabilité saisonnière du NDVI de blé et d'orge au niveau parcellaire entre 2002 et 2022

## 2.2. Au niveau communal

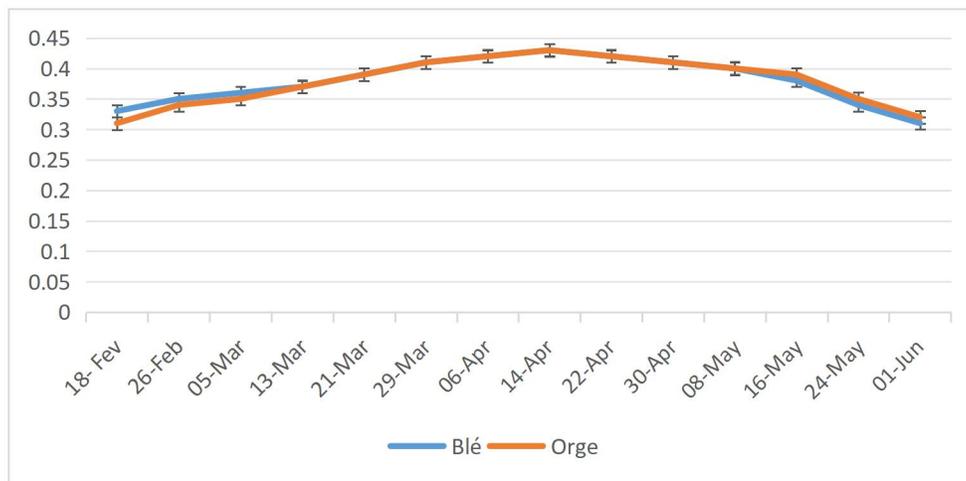
Les valeurs du NDVI de l'orge, ont fluctué entre 0.175 (1<sup>er</sup> Juin, 2017) et 0.51 (30 Avril, 2007) durant toutes les années. Pour le blé, ces valeurs ont varié de 0.175 (1<sup>er</sup> Juin, 2002) entre 0.545, (30 Avril 2022) (voir figure 44).



**Figure (44):**Variabilité saisonnière du NDVI de blé et d'orge au niveau de la commune de Sétif entre 2002 et 2022

## 2.3. Au niveau wilaya

Les valeurs du NDVI ont fluctué entre 0.214 (1<sup>er</sup> juin 2002) et 0.518 (6 avril en 2019), pour l'orge. Concernant le blé, ces valeurs ont varié de 0.204 (1<sup>er</sup>Jun 2017) à 0.536 (30 Avril, 2022) (figure 45).

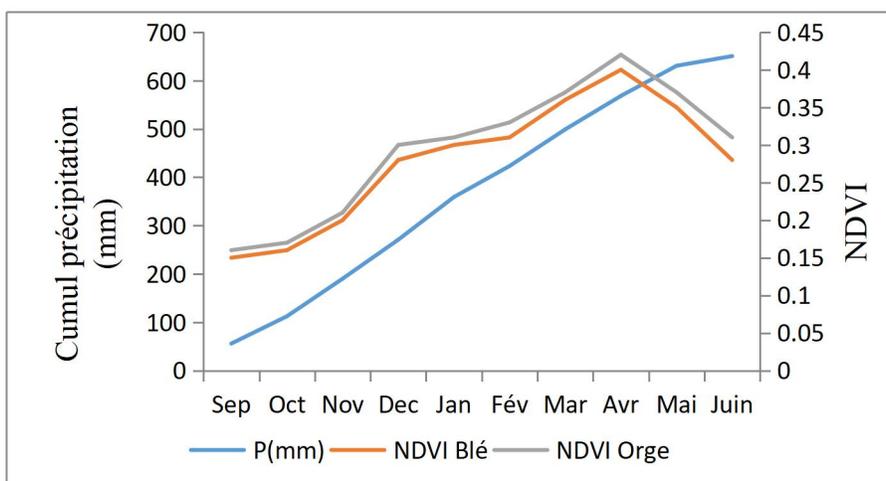


**Figure (45):** Variabilité saisonnière du NDVI de blé et d’orge au niveau de la wilaya de Sétif entre 2002 et 2022

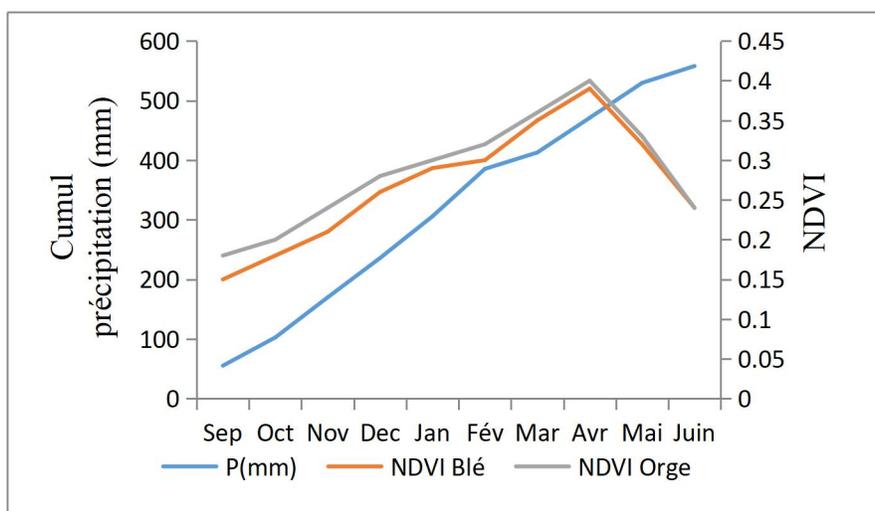
### 1.5. Relation entre NDVI et les précipitations

D'après les figures (46, 47, 48) ci dessous on observe une forte corrélation entre le NDVI moyen mensuel du blé et de l'orge ainsi que les précipitations moyennes mensuelles à l'échelle parcellaire, communale et wilaya, cette corrélation augmente progressivement jusqu'aux seuils spécifiques de précipitations dans chaque échelle (568 mm pour les parcelles, 471 mm pour la commune et 576.96 mm pour la wilaya). Cela est dû au fait que les plantes ont besoin d'eau pour la photosynthèse et une augmentation des précipitations peut favoriser leurs croissance, se traduisant par une augmentation du NDVI.

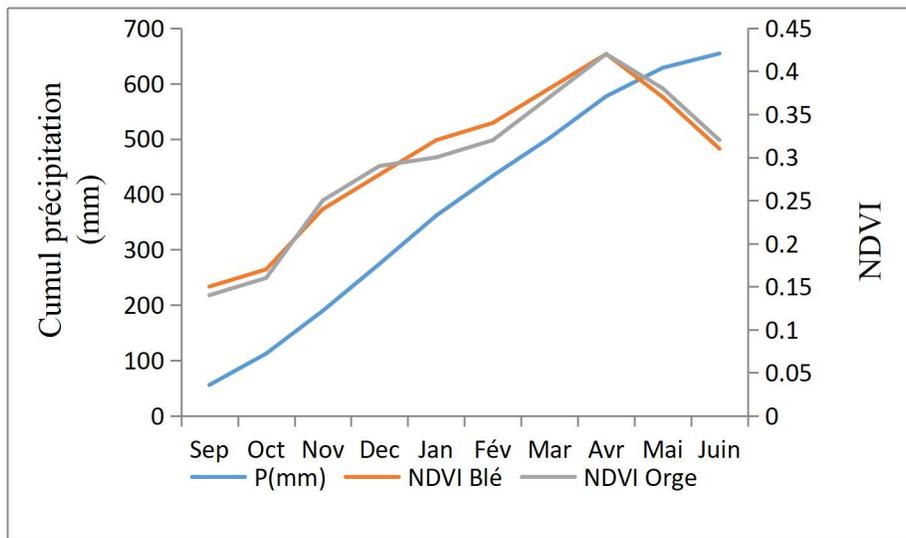
Cependant, une fois que les précipitations dépassent un certain seuil, la relation entre les deux variables se relâche, indiquant qu'une augmentation supplémentaire des précipitations peuvent ne pas avoir le même impact sur le NDVI moyen du blé et de l'orge, c'est ce qu'on appelle la saturation du NDVI.



**Figure (46):** Evolution du NDVI moyen (de blé et d’orge) et la pluviométrie cumulée (mm) au niveau des parcelles expérimentales de l’ITGC (2002 à 2022)



**Figure (47):** Evolution du NDVI moyen (de blé et d’orge) et la pluviométrie cumulée (mm) au niveau de la commune de Sétif (2002 à 2022)



**Figure (48):** Evolution du NDVI moyen (de blé et d'orge) et la pluviométrie cumulée (mm) au niveau de la wilaya de Sétif (2002 à 2022)

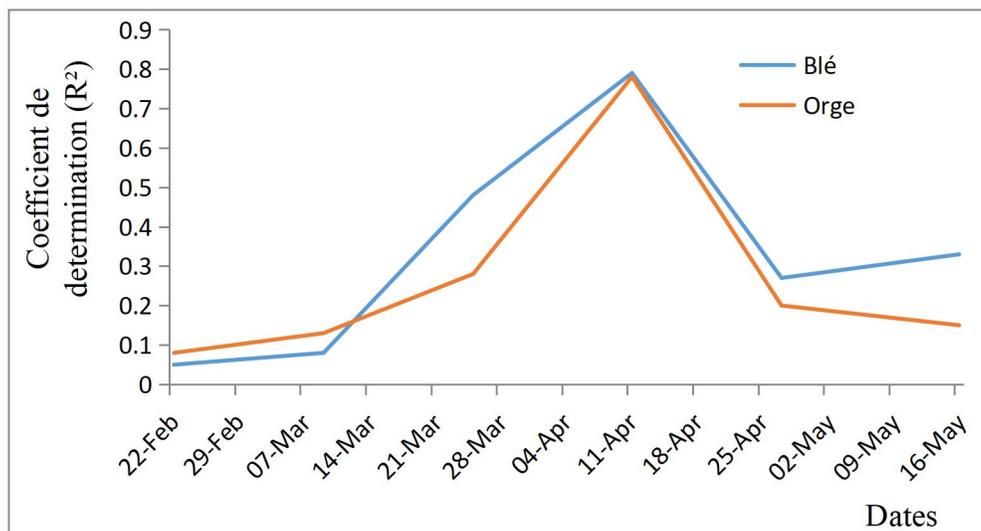
# **Chapitre (II) : Prédiction des rendements**

### I) Prédiction du rendement des céréales avec des données expérimentales de court terme (2020-2022) en utilisant le FGCC, le NDVI et la biomasse végétale au niveau parcellaire

#### 1. Prédiction des rendements par le FGCC (Fractional Green Canopy Cover)

Le pourcentage de couverture fractionnée de canopée verte (FGCC) est un outil efficace dans l'évaluation des rendements des céréales, car il reflète la santé et la vigueur des plantes. Des niveaux élevés de chlorophylle dans les feuilles sont généralement associés à une activité photosynthétique accrue, ce qui favorise une croissance et un développement optimaux des plantes. Ainsi, un FGCC élevé indique souvent des plantes en meilleure santé, capables de produire plus de biomasse et potentiellement de meilleurs rendements.

Les résultats de cette étude démontrent une relation linéaire positive entre le FGCC et le rendement en grains pour les deux céréales d'hiver (blé et orge) à l'échelle parcellaire. L'évolution des coefficients de déterminations sont présentés graphiquement sur la figure (49).



**Figure (49) :** Evolution du coefficient de détermination entre le rendement en grains et le FGCC pour les deux céréales (orge, blé) de 2020 à 2022 au niveau des parcelles de l'ITGC.

La corrélation entre FGCC et le rendement en grains pour les deux céréales (orge et blé) devient de plus en plus significative au fur et à mesure de l'avancement de la campagne agricole, à partir du 09 Mars jusqu'à atteindre un pic le 11 Avril, avec des valeurs de  $R^2$  variaient de 0.08 à 0.78 pour l'orge,  $R^2$  variaient 0.05 à 0.79 pour le blé. Ce pic de corrélation coïncide avec le

sommet moyen des valeurs de FGCC observées au cours des trois années, ou l'activité photosynthétique est optimale. Par la suite, la relation entre ces variables diminue progressivement à des dates ultérieures, ce déclin étant probablement dû au vieillissement des feuilles et à la perte d'activité photosynthétique (figure 49).

Sur la base de ces résultats, le moment optimal pour prédire avec précision le rendement en grains du blé et d'orge en utilisant FGCC au niveau des parcelles expérimentales est le début du mi-printemps, plus précisément le 11 Avril (environ 150 jours après le semis) qui correspond au stade "plein tallage- épiaison ".

Une analyse de régression linéaire a été effectuée afin de développer des modèles de prédiction des rendements céréaliers, on utilisant les valeurs de FGCC du 11 Avril (de 2020 à 2022) comme variables indépendantes, tandis que l'historique des rendements en grains de blé et d'orge ont été pris comme variables dépendantes, indiquant une forte relation entre le FGCC et le rendement en grains pour les deux types de céréales, avec des coefficients de détermination  $R^2=0.78$  pour l'orge et  $0.79$  pour le blé, ce qui signifie que 78 % et 79% de la variabilité du rendement d'orge et de blé peuvent être expliquée par le FGCC. Cela démontre que le FGCC a une forte influence sur le rendement de blé et d'orge. Les coefficients de régression pour le blé et l'orge étaient respectivement de 4.221 et 14.744, ce qui implique qu'une augmentation de 0.1 t/ha du FGCC est associée à une augmentation moyenne de 0.4221 t/ha et 1.477 t/ha du rendement en grains pour le blé et l'orge. Cette relation est soutenue également par des coefficients de corrélation (R) élevées,  $R=0.89$  et  $0.88$  pour le blé et l'orge respectivement, ce qui exprime un lien fort et une relation positive entre les deux variables, car la valeur du R est aussi positive au niveau de signification de 0.05 ( $\alpha=5\%$ ).

Selon cette régression linéaire simple, l'estimation des rendements de blé suit l'équation suivante :

$$Y = 4.221x \text{ FGCC}_{(11 \text{ avril})} - 0.617$$

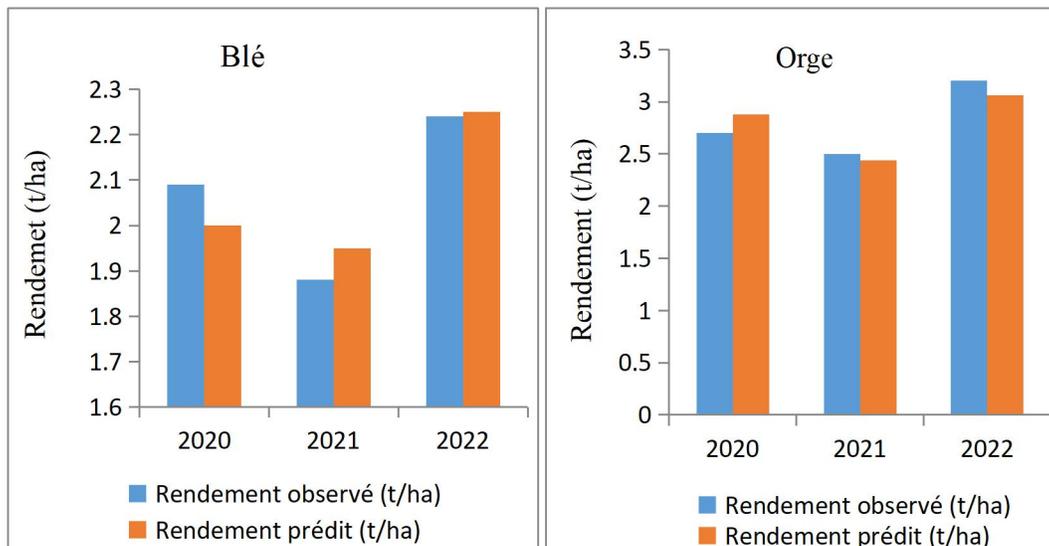
Pour de l'orge, l'estimation des rendements suit l'équation suivante :

$$Y = 14.744x \text{ FGCC}_{(11 \text{ avril})} - 8.317$$

### **1.1. Vérification de la performance des modèles**

Quatre paramètres de performance ont été utilisés pour évaluer la performance des modèles de prévision: la racine carrée de l'erreur quadratique moyenne (RMSE), l'erreur quadratique moyenne (MSE), le coefficient de corrélation (R) et le coefficient de détermination ( $R^2$ ).

Les résultats de la prédiction des rendements à l'aide du FGCC ont révélé une forte corrélation entre les rendements réels et prédits, les coefficients de corrélation (R) de 0.89 pour le blé et 0.88 pour l'orge indiquent une forte corrélation linéaire entre les rendements observés et les rendements prédits au niveau de signification de 0.05 ( $\alpha = 5\%$ ). De plus les coefficients de détermination ( $R^2$ ) de 0.80 et 0.87 pour le blé et l'orge respectivement, cela signifie que 80% et 78% de la variation dans les rendements observés de blé et d'orge sont expliquées par le modèle de régression. Les faibles valeurs de RMSE, allant de 0.51 à 0.107 t/ha, indiquent que les prédictions du modèle sont très proches des valeurs réelles. De même les valeurs de l'erreur quadratique moyenne (MSE) sont basses allant de 0.004 à 0.018, suggèrent que les prédictions du modèle sont précises (figure 50). Les résultats de performance sont présentés dans le tableau (08).



**Figure (50) :** Comparaison des rendements observés et prédit selon le modèle créé pour le blé et l'orge en utilisant FGCC au niveau parcellaire.

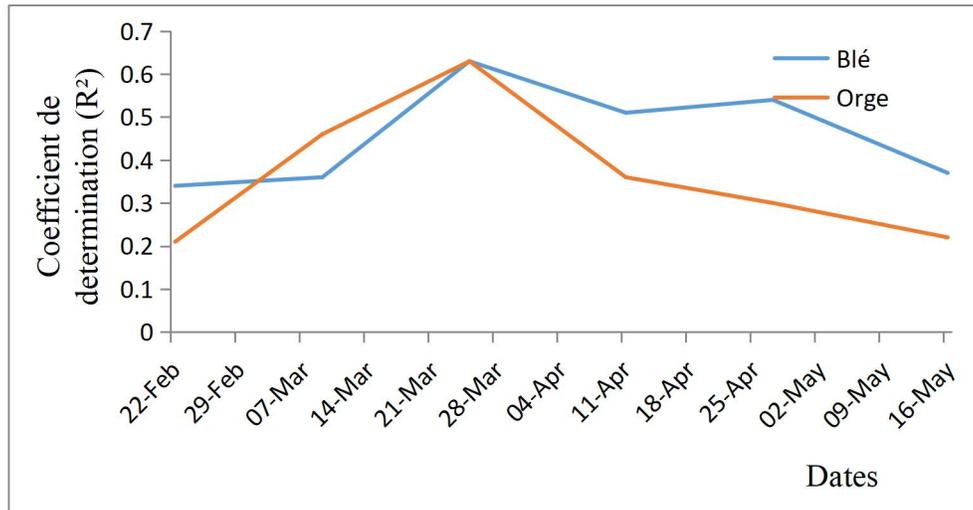
**Tableau (08) :** Résultats de performances des modèles au niveau parcellaire en utilisant le FGCC.

Cultures	$R^2$	R	RMSE	MSE	Equation
Ogre	0.78	0.88	0.134	0.018	$Y=1.003 x + 0.001$
Blé	0.8	0.89	0.063	0.004	$Y=1.006 x + 0.010$

## 2. Prédiction des rendements par le NDVI

Le NDVI est un outil efficace dans l'évaluation des rendements des céréales, Il fournit des informations précieuses sur les effets combinés de divers facteurs édapho-climatiques tels que:

les conditions météorologiques, les variétés de cultures, les types de sols, et d'autres facteurs sur la croissance et le rendement des céréales. Les résultats démontrent une relation linéaire positive entre MODIS-NDVI et le rendement en grains pour les deux céréales d'hiver (blé et orge) à l'échelle parcellaire. Les coefficients de déterminations sont présentés graphiquement dans la figure (51).



**Figure (51) :** Evolution du coefficient de détermination entre le rendement en grains et le NDVI pour les deux céréales (orge, blé) de 2020 à 2022 au niveau parcellaire.

La corrélation entre le NDVI et le rendement en grains pour les deux céréales (orge et blé) devient de plus en plus forte au fur et à mesure de l'avancement de la campagne agricole, à partir du 09 Mars et atteint un pic le 25 Mars, avec des valeurs de  $R^2$  variaient de 0.21 à 0.64 à pour l'orge,  $R^2$  variaient de 0.34 à 0.63 pour le blé. Ce pic de corrélation coïncide avec le sommet moyen des valeurs de NDVI observées au cours des 3 ans d'expérimentation. Par la suite, la relation entre ces variables diminue progressivement à des dates ultérieures, ce déclin pouvant être attribué à la saturation du NDVI lors des derniers stades de croissance des céréales (figure 48).

Sur la base de ces résultats, on peut déduire que le moment optimal de prédiction avec précision le rendement en grains du blé et d'orge en utilisant MODIS-NDVI au niveau des parcelles expérimentales de l'ITGC- Sétif- est le début du printemps, plus précisément le 25 Mars qui correspondant au stade "plein tallage".

Une analyse de régression linéaire a été effectuée afin de développer des modèles de prédiction des rendements céréaliers, en utilisant les valeurs du NDVI du 25 Mars (de 2020 à 2022) comme variables indépendantes, tandis que l'historique des rendements en grains de blé et

d'orge ont été pris comme variables dépendantes, indiquant une forte corrélation entre le NDVI et le rendement en grains pour les deux types de céréales, avec des coefficients de détermination  $R^2 = 0.64$  pour l'orge et  $0.63$  pour le blé. Cela démontre que les valeurs de NDVI ont une forte influence sur la détermination du rendement de blé et d'orge. A partir des coefficients de régression des modèles, on peut estimer qu'une augmentation de  $0.1$  du NDVI est associée à une augmentation moyenne de  $1.251$  t/ha et  $4.072$  t/ha du rendement en grains pour le blé et l'orge. Cette relation est soutenue également par des coefficients de corrélation ( $R$ ) élevés,  $R = 0.82$  et  $0.80$  pour le blé et l'orge respectivement, ce qui exprime un lien fort et une relation positive entre les deux variables, car la valeur du  $R$  est aussi positive au niveau de signification de  $0.05$  ( $\alpha = 5\%$ ).

L'estimation des rendements de blé suit l'équation suivante :

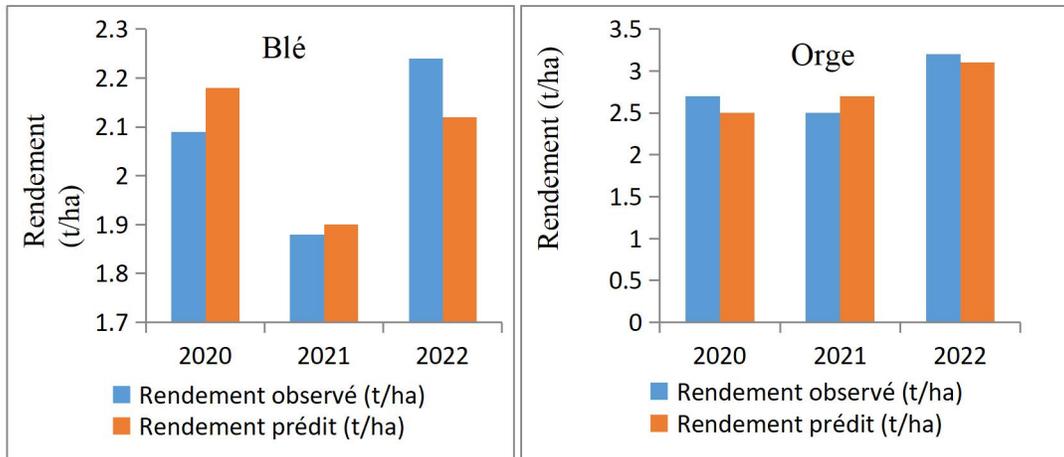
$$\text{Rendements blé} = 12.519 \times \text{NDVI}_{(25\text{mars})} - 3.1$$

Pour de l'orge, la prévision des rendements suit l'équation suivante :

$$\text{Rendements orge} = 40.728 \times \text{NDVI}_{(25\text{mars})} - 14.279$$

### **2.1. Vérification de la performance des modèles**

Les résultats de la prédiction des rendements à l'aide du NDVI au niveau des parcelles, ont révélé une forte corrélation entre les rendements réels et prédits, les coefficients de corrélation ( $R$ ) de  $0.80$  pour le blé et  $0.82$  pour l'orge indiquent une forte corrélation linéaire entre les rendements observés et les rendements prédits au niveau de signification de  $0.05$  ( $\alpha = 5\%$ ). De plus les coefficients de détermination ( $R^2$ ) de  $0.64$  et  $0.68$  pour le blé et l'orge respectivement, cela signifie que  $64\%$  et  $68\%$  de la variation dans les rendements observés de blé et d'orge sont expliquées par le modèle de régression. Les faibles valeurs de RMSE, allant de  $0.086$  à  $0.185$  t/ha, indiquent que les prédictions du modèle sont très proches des valeurs réelles. De même les valeurs de MSE sont basses allant de  $0.007$  à  $0.031$  t/ha (figure 52). Les résultats de performance des modèles sont présentés dans le tableau (09).



**Figure (52) :** Comparaison des rendements observés et prédit selon le modèle créé pour le blé et l’orge en utilisant le NDVI au niveau parcellaire

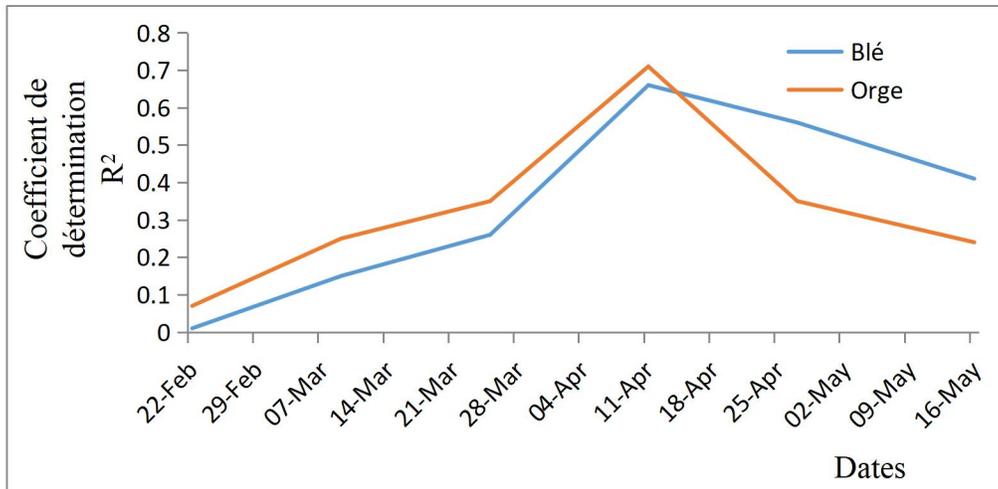
**Tableau (09) :** Résultats de performances des modèles au niveau parcellaire en utilisant le NDVI durant la période 2020-2022.

Cultures	R <sup>2</sup>	R	RMSE	MSE	Equation
Ogre	0.68	0.82	0.083	0.007	Y=1.022 x-0.046
Blé	0.64	0.8	0.176	0.031	Y=1.018 x-0.054

### 3. Prédiction des rendements en utilisant la biomasse sèche

La biomasse sèche, représente la quantité de matière végétale présente dans les cultures, est souvent étroitement liée à la productivité des céréales, une bonne biomasse sèche indique généralement une croissance vigoureuse et des rendements potentiels. En établissant des relations entre la biomasse sèche mesurée et les rendements observés, il est possible de prédire les rendements futurs des céréales.

D’après la figure (53), on observe une relation linéaire positive entre la biomasse sèche et le rendement en grains pour les deux céréales d'hiver (blé et orge) à l’échelle parcellaire.



**Figure (53) :** Evolution du coefficient de détermination entre le rendement en grains et la biomasse sèche pour les deux céréales (orge, blé) de 2020 à 2022 au niveau des parcelles de l'ITGC.

La corrélation entre la biomasse sèche et le rendement en grains des céréales (orge et blé) se renforce au fur et à mesure que progresse la campagne agricole, atteignant son point culminant le 11 avril, avec des coefficients de détermination ( $R^2$ ) variant de 0.07 à 0.71 pour l'orge et de 0.01 à 0.66 pour le blé. Par la suite, cette corrélation diminue progressivement aux dates ultérieures.

Ces résultats suggèrent que le moment optimal pour prédire le rendement du blé et de l'orge en se basant sur la biomasse sèche des parcelles expérimentales est le 11 avril, correspondant au stade "plein tallage-épiaison".

En utilisant les valeurs de biomasse sèche du 11 avril (de 2020 à 2022) comme variables indépendantes et les rendements moyens du blé et de l'orge comme variables dépendantes, la régression linéaire révèle une forte corrélation entre la biomasse sèche et le rendement en grains pour les deux types de céréales, avec des coefficients de détermination ( $R^2$ ) de 0.71 pour l'orge et 0.66 pour le blé. Les coefficients de régression pour le blé et l'orge étaient respectivement de 0.31 et 0.202, ce qui suggère qu'une augmentation de 0.1 t/ha de la biomasse sèche est associée à une augmentation moyenne de 0.031 t/ha et 0.020 t/ha du rendement du blé et de l'orge, respectivement.

Selon le modèle de prévision, l'estimation des rendements de blé suit l'équation suivante :

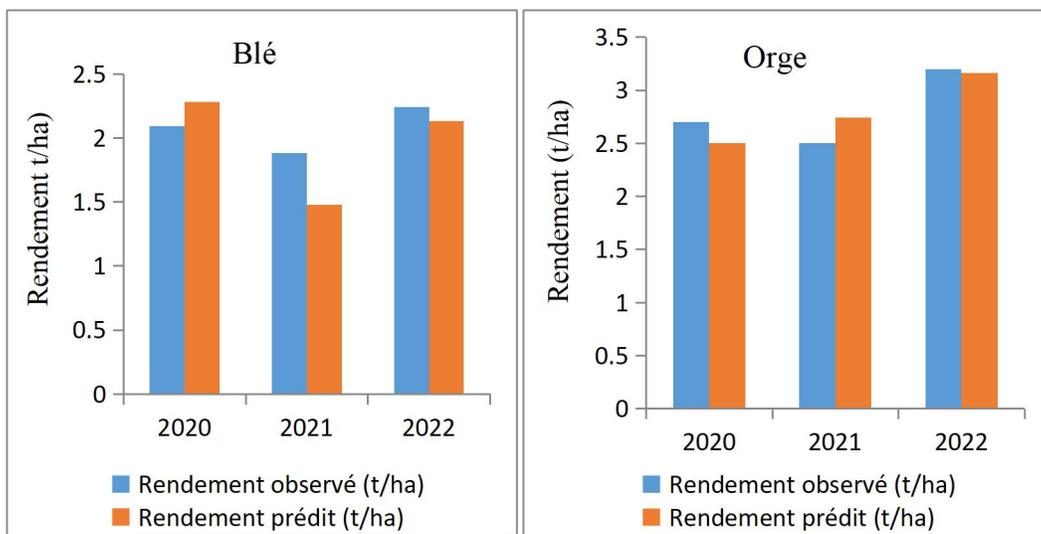
$$\text{Rendement blé} = 0.031 \times \text{biomasse}_{(11\text{avril})} - 30.06$$

Alors que, l'estimation des rendements de l'orge suit la relation ci-dessous :

$$\text{Rendement orge} = 0.202 \times \text{biomasse}_{(11\text{avril})} - 22.176$$

### 3.1. Vérification de la performance des modèles

Les résultats de la prédiction des rendements à l'aide de la biomasse au niveau des parcelles de l'ITGC ont révélé une forte corrélation entre les rendements réels et prédits, les coefficients de corrélation (R) de 0.82 pour le blé et 0.84 pour l'orge au niveau de signification de 0.05 ( $\alpha= 5\%$ ). De plus les coefficients de détermination ( $R^2$ ) de 0.64 et 0.71 pour le blé et l'orge respectivement, cela signifie que 64% et 71 % de la variabilité des rendements observés de blé et d'orge sont expliquées par le modèle de régression. Les valeurs de RMSE (0.092 à 0.172 t/ha) et du MSE (0.007 à 0.025t/ha), sont basses, indiquant que les prédictions des modèles sont très proches des valeurs réelles (figure 54, tableau 10).



**Figure (54) :** Corrélation entre les rendements observés et prédits selon le modèle créé pour le blé et l'orge en utilisant la biomasse sèche au niveau parcellaire

**Tableau (10) :** Résultats de performances des modèles au niveau parcellaire en utilisant la biomasse sèche.

Cultures	R <sup>2</sup>	R	RMSE	MSE	Equation
Ogre	0.71	0.84	0.083	0.007	Y=1.005 x-0.003
Blé	0.64	0.82	0.158	0.025	Y=1.037 x-0.060

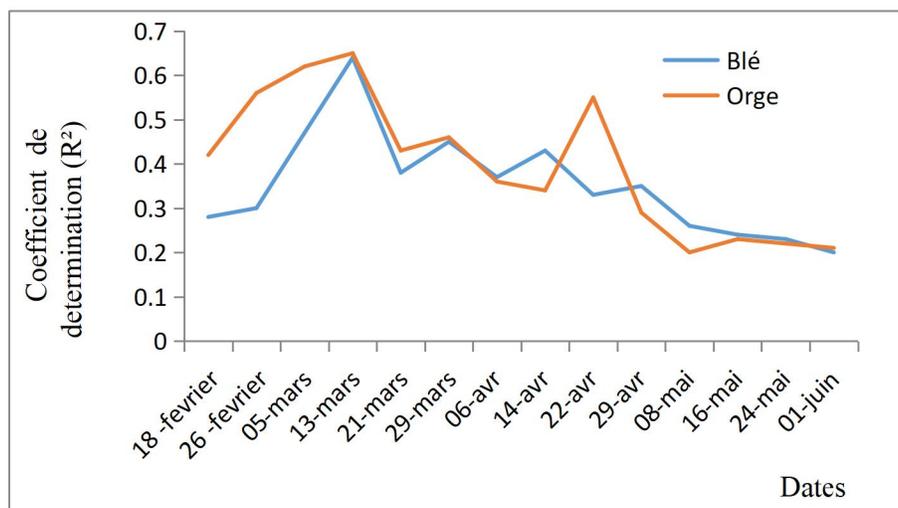
D'après les résultats obtenus durant les trois ans d'expérimentation, il est observé que le FGCC, obtenu à partir de l'application mobile Canopeo®, présente les meilleures performances d'évaluation des rendements de blé et d'orge parmi les trois méthodes étudiées au niveau des parcelles expérimentales de l'ITGC, ayant le coefficient de détermination ( $R^2$ ) les plus élevés, avec les valeurs les plus basses de RMSE et MSE. Cela indique que le FGCC est un prédicteur puissant du rendement en grains de blé et d'orge dans cette étude. Suivie par la méthode de la biomasse sèche, elle fonctionne également bien, avec des valeurs  $R$  et  $R^2$  élevées et des RMSE et MSE relativement faibles. En comparaison, la méthode NDVI montre une corrélation proche de celle du FGCC et de la biomasse sèche, avec des valeurs de RMSE et MSE très proche des deux autres méthodes, ce qui suggère qu'elle peut être aussi précise et potentiellement plus rapide et moins coûteuse que les méthodes du FGCC et de la biomasse verte dans ce contexte.

## II) Prédiction des rendements à partir des données historiques de long terme (2002-2022)

### 1. Prédiction des rendements au niveau parcellaire

#### 1.1. Prédiction des rendements en utilisant le NDVI

Les résultats de cette étude démontrent une corrélation linéaire positive entre le NDVI et le rendement pour les deux céréales (blé et orge) à l'échelle parcellaire. Les coefficients de déterminations sont présentés dans la figure (55).



**Figure (55) :** Evolution du coefficient de détermination entre le rendement en grains et le NDVI des deux céréales (orge et blé), couvrant la période de 2002-2022 au niveau parcellaire.

La corrélation entre le NDVI et le rendement en grains pour les deux céréales (orge et blé) devient de plus en plus forte au fur et à mesure de l'avancement de la campagne agricole, à partir du 26 Février et atteint son pic le 13 Mars, avec des valeurs de  $R^2$  variaient de 0.2 à 0.65 pour l'orge et entre 0.21 à 0.64 pour le blé. Ce pic coïncide avec le sommet moyen des valeurs de NDVI observées au cours des différentes années. Par la suite, la corrélation entre ces variables diminue progressivement à des dates ultérieures, ce déclin pouvant être attribué à la sénescence et le changement du couleur des feuilles lors des derniers stades de croissance des céréales (figure 55). Cela démontre que le moment optimal de prédiction du rendement en grains du blé et d'orge en utilisant MODIS-NDVI, est le 13 mars qui correspondant au stade "plein tallage".

Sur la base de résultats obtenus, on peut déduire que le NDVI contribue avec une part importante dans l'explication de la variabilité des rendements des céréales, cependant la partie restante peut être expliquée par l'effet des autres paramètres agro-climatiques (précipitations, températures, évapotranspiration...etc.), édaphiques ou techniques.

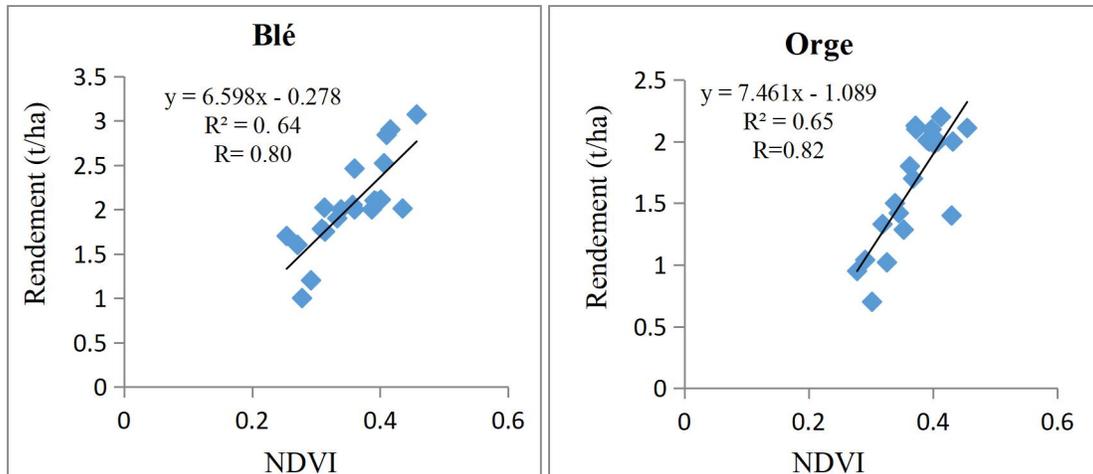
L'analyse par régression linéaire, on utilisant le NDVI du 13 mars (de 2002 à 2022) comme prédicteur montre que les coefficients de régression du blé et l'orge étaient respectivement de 6.598 et 7.461. Ce qui implique qu'une augmentation de 0.1 du NDVI est associée à une augmentation moyenne de 0.659 t/ha et 0.746 t/ha du rendement en grains pour le blé et l'orge (figure 56).

Selon le modèle obtenu, l'estimation des rendements de blé suit l'équation suivante :

$$\text{Rendement blé} = 6.598 \times \text{NDVI}_{(13\text{mars})} - 0.278$$

Pour l'orge, l'estimation des rendements suit l'équation suivante :

$$\text{Rendement orge} = 7.461 \times \text{NDVI}_{(13\text{mars})} - 1.089$$

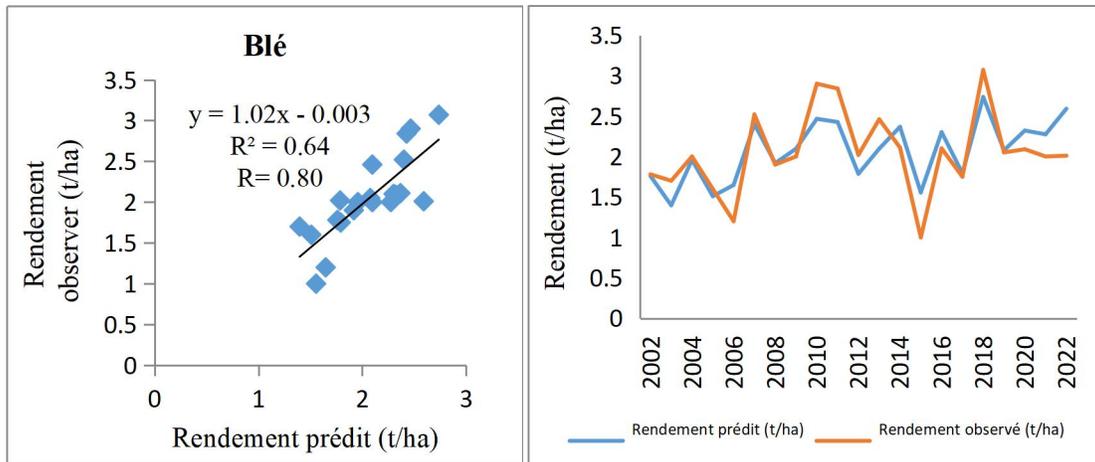


**Figure (56) :** Modèle de régression linéaire entre le rendement (de blé et d'orge) et le NDVI, le 13 mars couvrant la période 2002-2022, au niveau parcellaire.

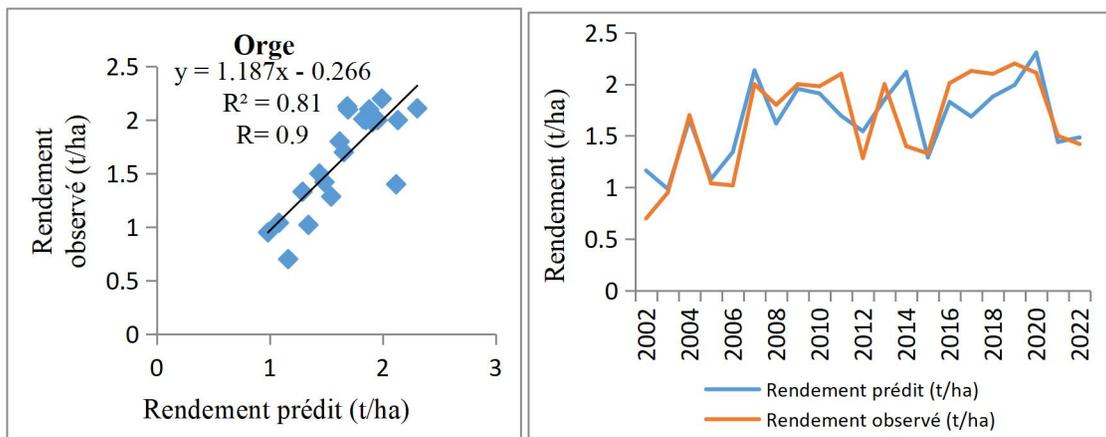
### 1.1.1. Vérification de la performance des modèles

Les résultats de la prédiction des rendements à l'aide du NDVI ont révélé une forte corrélation entre les rendements réels et prédits, les coefficients de corrélation ( $R$ ) de 0.80 pour le blé et 0.90 pour l'orge indiquent une forte corrélation entre les rendements observés et les rendements prédits au niveau de signification de 0.05 ( $\alpha = 5\%$ ). De plus les coefficients de détermination ( $R^2$ ) sont 0.64 pour le blé et 0.81 pour l'orge, ce qui signifie que 64% et 81% de la variation dans les rendements observés de blé et d'orge sont expliquées par le modèle de régression. Les faibles valeurs de RMSE (0.01 à 0.276 t/ha) et du MSE (0.061 à 0.076), indiquent que les prédictions du modèle sont précises et très proches des valeurs réelles (figures 57, 58).

Les résultats de performance des modèles sont présentés dans le tableau (11).



**Figure (57).** Corrélation entre les rendements observés et prédit du blé en utilisant le NDVI au niveau parcellaire



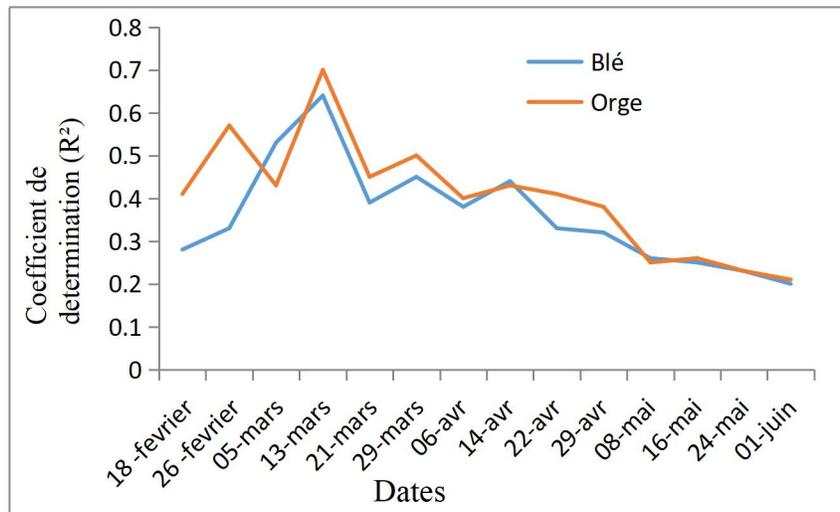
**Figure (58).** Corrélation entre les rendements observés et prédit de l'orge en utilisant le NDVI au niveau parcellaire

**Tableau (11).** Résultats de performances des modèles de prédiction du rendement au niveau parcellaire, en utilisant le NDVI.

Cultures	R <sup>2</sup>	R	RMSE	MSE	Equation
Ogre	0.81	0.90	0.246	0.061	Y=1.187x - 0.266
Blé	0.64	0.8	0.276	0.076	Y=1.02x - 0.003

### 1.2. Prédiction des rendements en utilisant le NDVI et les précipitations

La combinaison de deux variables explicatives telles que le NDVI et la pluviométrie, permet d'améliorer la précision des prédictions à différentes échelles spatiales. Étant donné que le NDVI et la pluviométrie sont des facteurs environnementaux majeurs influençant la croissance et le rendement des cultures, l'intégration de ces variables dans un modèle de régression linéaire multiple permet d'étudier leurs interactions et leur impact sur le rendement des céréales.



**Figure (59) :** Evolution du coefficient de détermination dans le cas de la combinaison du NDVI et de la pluviométrie dans la prévision du rendement des deux céréales (orge et blé) au niveau parcellaire.

La corrélation entre la combinaison du NDVI et de la pluviométrie avec le rendement au niveau des parcelles d'ITGC de Sétif augmente au fur et à mesure de l'avancement de la campagne agricole, à partir du 26 Février et atteint son pic le 13 Mars pour les deux cultures. Pour l'orge, les valeurs de  $R^2$  variaient de 0,2 à 0,7, tandis que pour le blé, les valeurs de  $R^2$  variaient de 0,20 à 0,64. Par la suite, cette relation tend à diminuer progressivement à des dates ultérieures (figure 59).

D'après ces résultats, il est possible de déduire que le moment optimal pour prédire le rendement du blé et d'orge au niveau parcellaire en utilisant la combinaison du NDVI et de la pluviométrie, est le stade de "plein tallage", plus précisément le 13 Mars pour les deux cultures. À ce stade, les conditions environnementales exercent une influence significative sur le développement final des céréales.

Une analyse de régression linéaire multiple a été effectuée afin de développer des modèles de prédiction des rendements céréaliers, on utilisant les valeurs du NDVI et de la pluviométrie du 13 Mars (entre 2002 et 2022) comme variables indépendantes pour les deux cultures, tandis que les rendements en grains de blé et d'orge ont été pris comme variables dépendantes. Une corrélation positive forte a été constaté, entre le NDVI et la pluviométrie avec le rendement en grain de blé, avec un coefficient de détermination ( $R^2 = 0.64$ ). De manière similaire, pour l'orge, on observe une corrélation positive forte avec un coefficient de détermination ( $R^2 = 0.7$ ).

D'après les régressions linéaires multiples, l'estimation des rendements de blé suit l'équation suivante :

$$\text{Rendement blé} = 7.13 \times \text{NDVI}_{(13 \text{ mars})} - 0.0004 \times \text{pluviométries}_{(13 \text{ mars})} - 0.450$$

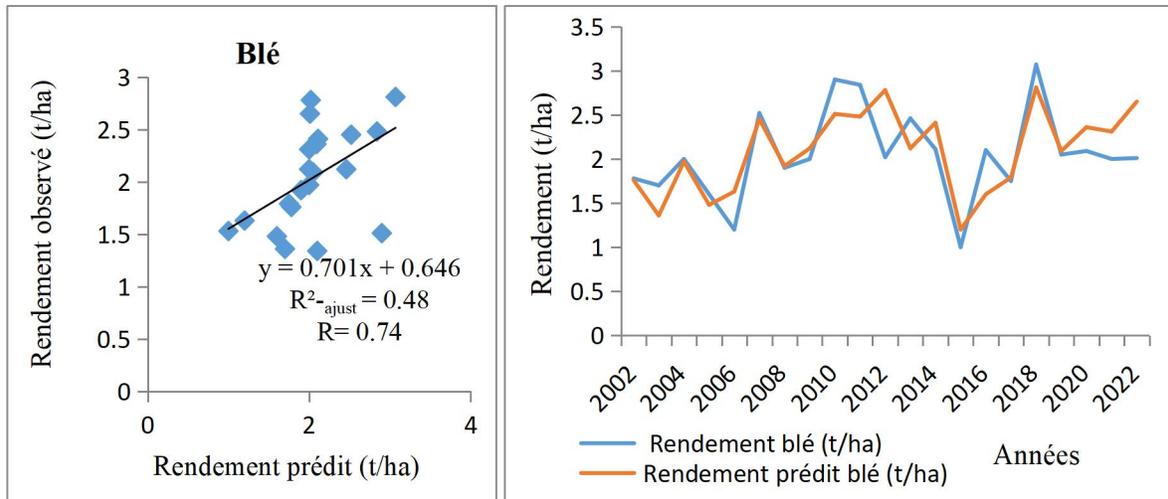
Et celle de l'orge suit l'équation suivante :

$$\text{Rendement orge} = 7.870 \times \text{NDVI}_{(13 \text{ mars})} - 0.001 \times \text{pluviométries}_{(13 \text{ mars})} - 1.115$$

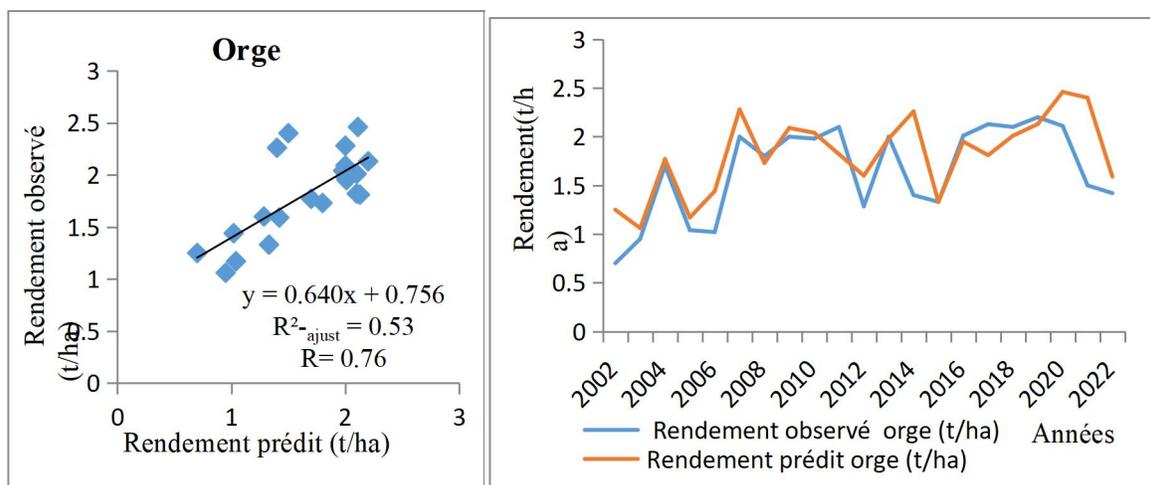
### **1.2.1. Vérification de la performance des modèles**

Les résultats de la prédiction des rendements à l'aide de la combinaison du NDVI et de la pluviométrie ont révélé une forte corrélation entre les rendements réels et prédits, les coefficients de corrélation ( $R$ ) de 0.74 pour le blé et 0.76 pour l'orge indiquent une forte corrélation entre les rendements observés et les rendements prédits au niveau de signification de 0.05 ( $\alpha = 5\%$ ). De plus les coefficients de détermination ( $R^2_{\text{ajust}}$ ) de 0.48 et 0.53 pour le blé et l'orge respectivement cela signifie que 48% et 53% de la variance dans les rendements observés de blé et d'orge sont expliquées par le modèle de régression, au niveau de signification de 0.05 ( $\alpha = 5\%$ ). Les faibles valeurs de RMSE, allant de 0.285 à 0.319 t/ha, indiquent que les prédictions du modèle sont très proches des valeurs réelles. De même les valeurs du MSE sont basses allant de 0.092 à 0.096, suggèrent que les prédictions du modèle sont généralement précises (figures 60, 61).

Les résultats sont présentés dans le tableau (12), confirmant ainsi une prédiction précise des rendements deux à trois mois avant la récolte au niveau parcellaire.



**Figure (60) :** Corrélation entre les rendements prédit et observé du blé au niveau parcellaire en utilisant un modèle de régression multiple (NDVI et pluviométrie) entre 2002-2022



**Figure (61) :** Corrélation entre les rendements prédit et observé de l’orge au niveau parcellaire en utilisant un modèle de régression multiple (NDVI et pluviométrie) entre 2002-2022

**Tableau (12) :** Résultats de performances des modèles de régression linéaire multiple associant le NDVI et la pluviométrie au niveau parcellaire.

Cultures	R <sup>2</sup> -ajust	R	RMSE	MSE	Equation
Ogre	0.53	0.76	0.303	0.092	Y= 0.640x + 0.756
Blé	0.48	0.74	0.309	0.096	Y= 0.701x + 0.646

**1.3. Prédiction des rendements avec les algorithmes d'apprentissage (Machine learning)**

Dans cette analyse on a utilisé les modèles (algorithmes) de prédiction suivants : SVM, Neural Network, KNN, Random Forest, Tree, Gradient Boosting, AdaBoost, Stochastic Gradient Descent. Afin d'entraîner les modèles sur les variables et les validé par la suite, on a divisé les données d'une façon aléatoire en deux : 70% pour l'entraînement et 30% pour le test, en suivant méthode de la validation croisée (cross-validation).

**1.3.1 Entraînement des modèles en utilisant tous les variables au niveau parcellaire**

Le tableau (13) révèle que lors de l'utilisation de toutes les variables d'entres (87 variables), le modèle "KNN" semble être le plus précis et le plus adapté pour prédire les rendements de blé au niveau parcellaire, avec un R<sup>2</sup> de 0.367 qui explique environ 36.7% de la variation des rendements avec un RMSE de 0.464t/ha et un MSE de 0.215t/ha. Ensuite, le modèle "SVM" se classe le deuxième avec (R<sup>2</sup>=0.334), suivi par le modèle "Stochastic Gradient Descent" (R<sup>2</sup>=0.305) en troisième position. En revanche, certains modèles comme "la régression linéaire" et les modèles d'arbres (Tree et Random Forest) présentent des performances plus faibles avec des R<sup>2</sup> négatifs ou très faibles.

Pour l'orge, le modèle "Gradient Boosting" semble être le plus précis et le mieux adapté pour prédire les rendements d'orge au niveau parcellaire, avec un R<sup>2</sup> de 0.528 qui explique environ 52.8% de la variation des rendements avec un RMSE de 0.347t/ha et un MSE de 0.12t/ha. Le modèle "Tree" avec un (R<sup>2</sup>=0.511) se classe en deuxième position, suivi par le modèle "Random Forest" avec (R<sup>2</sup>=0.491) en troisième position. Cependant, certains modèles tels SVM, la régression linéaire et le modèle "Stochastic Gradient Descent" présentent des performances plus faibles avec des faibles valeurs de R<sup>2</sup>.

**Tableau (13) : Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant toutes les variables au niveau parcellaire.**

Modèles	Blé			Orge		
	R <sup>2</sup>	RMSE	MSE	R <sup>2</sup>	RMSE	MSE
Régression linéaire	0.235	0.510	0.260	0.125	0.472	0.223
KNN	<b>0.367</b>	<b>0.464</b>	<b>0.215</b>	0.144	0.467	0.218

<b>Random Forest</b>	-0.006	0.584	0.314	0.491	0.36	0.130
<b>Tree</b>	-0.225	0.645	0.416	0.511	0.353	0.125
<b>SVM</b>	0.334	0.475	0.226	0.092	0.481	0.231
<b>Gradient Boosting</b>	0.159	0.534	0.285	<b>0.528</b>	<b>0.347</b>	<b>0.120</b>
<b>Neural Network</b>	0.262	0.500	0.250	0.169	0.46	0.212
<b>AdaBoost</b>	0.296	0.489	0.239	0.437	0.379	0.143
<b>Stochastic Gradient Descent</b>	0.305	0.486	0.236	0.161	0.462	0.214

### 1.3.2. Prédiction des rendements en utilisant les dix meilleurs variables

Le tableau (14) révèle que lors de l'utilisation des dix meilleurs variables (par la méthode de Gain information), le modèle "KNN" semble être le plus précis et le plus adapté pour prédire les rendements de blé au niveau parcellaire, avec un  $R^2$  de 0.386 qui explique environ 38.6% de la variation des rendements, avec un RMSE de 0.456 t/ha. Ensuite, le modèle "SVM" se classe deuxième avec ( $R^2=0.263$ ), suivi par le modèle "Random Forest" ( $R^2=0.185$ ) en troisième position. En revanche, certains modèles comme "la régression linéaire" et les "de Gradient Boosting" et "le AdaBoosting" présentent des performances plus faibles avec des  $R^2$  négatifs ou très faibles.

Pour l'orge, le modèle "Tree" semble être le plus précis et adapté pour prédire les rendements d'orge au niveau parcellaire, avec un  $R^2$  de 0.672 et un RMSE de 0.289 t/ha. Ensuite le modèle "Random Forest" avec un ( $R^2=0.529$ ) se classe en deuxième position, suivi par le modèle "AdaBoost" avec un ( $R^2=.522$ ) en troisième position. Des modèles comme "la régression linéaire", "KNN" et "SVM" présentent des performances plus faibles avec des faibles valeurs de  $R^2$ .

Il est notable que les prédictions du rendement de l'orge ont été plus précises que celles du blé, que ce soit en utilisant toutes les variables ou uniquement les dix meilleures variables.

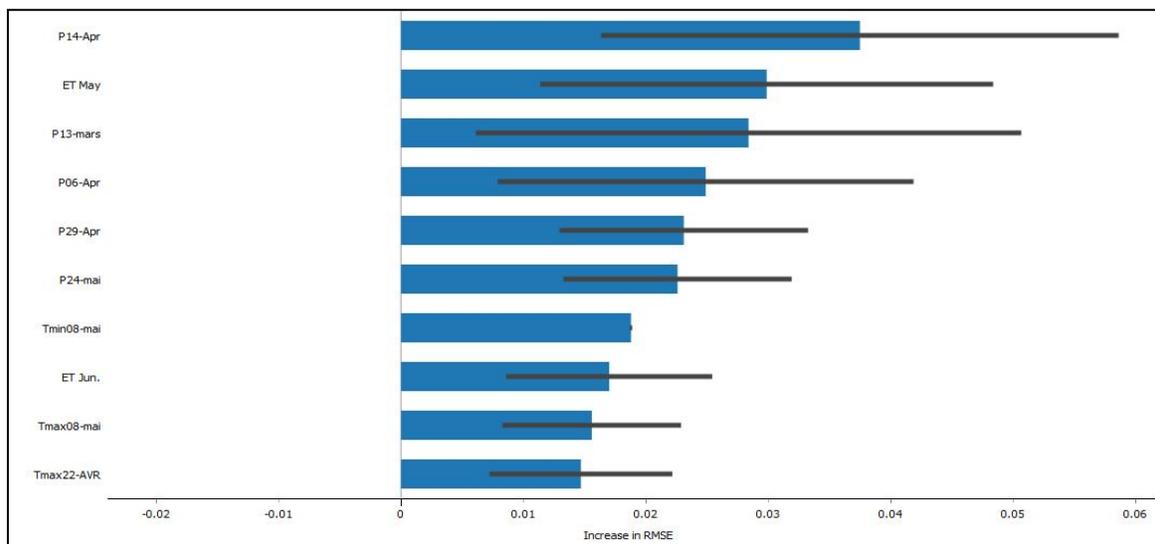
**Tableau (14):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant les dix meilleurs variables au niveau parcellaire

Modèles	Blé			Orge		
	R <sup>2</sup>	RMSE	MSE	R <sup>2</sup>	RMSE	MSE
Régression linéaire	-0.529	0.927	0.858	-1.100	0.731	0.535
<b>KNN</b>	<b>0.386</b>	<b>0.456</b>	<b>0.208</b>	0.296	0.423	0.179
<b>Random Forest</b>	0.185	0.526	0.277	0.529	0.289	0.084
<b>Tree</b>	-0.001	0.583	0.344	<b>0.672</b>	<b>0.084</b>	<b>0.289</b>
<b>SVM</b>	0.263	0.500	0.250	0.170	0.460	0.211
<b>Gradient Boosting</b>	-0.022	0.589	0.347	0.434	0.380	0.144
<b>Neural Network</b>	0.054	0.567	0.321	0.436	0.379	0.144
<b>AdaBoost</b>	-0.010	0.585	0.343	0.522	0.349	0.122
<b>Stochastic Gradient Descent</b>	0.036	0.572	0.327	0.464	0.370	0.137

➤ **Extraction des meilleurs variables pour la prévision du rendement du blé**

L'extraction des meilleurs variables pour la prévision du rendement du blé a été effectuée par la méthode de 'Gain information'. D'après la figure (62), on déduit que les précipitations du 14 Avril, ainsi que celles du 13 Mars, du 06 Avril, du 29 Avril et du 24 Mai, associées à la température minimale du 08 Mai, ainsi qu'à la température maximale du 08 Mai et du 22 Avril, et l'évapotranspiration (ET) du Mai et de Juin, sont les 10 meilleures variables pour prédire le rendement du blé en utilisant le modèle KNN au niveau parcellaire. Ci-dessous une explication de la contribution des différents variables dans la détermination du rendement :

- ✓ Les précipitations du mois de Mars (début du printemps), sont bénéfiques pour favoriser la croissance et le développement des cultures de blé. Ce qui se reflète sur la vigueur initiale des plantes et leur capacité à absorber les nutriments du sol.
- ✓ Les précipitations en Avril sont souvent essentielles, lorsque les plantes atteignent le stade de montaison, c'est-à-dire la formation des épis de blé, un niveau approprié de précipitations à ce stade peut favoriser un développement optimal des grains et une meilleure qualité de la récolte.
- ✓ Les températures appropriées à des moments spécifiques, comme le début de Mai, jouent un rôle crucial dans la maturation et la qualité des grains de blé. De même, les températures en Avril peuvent influencer le développement végétatif et la floraison du blé, ce qui est directement lié au rendement final.



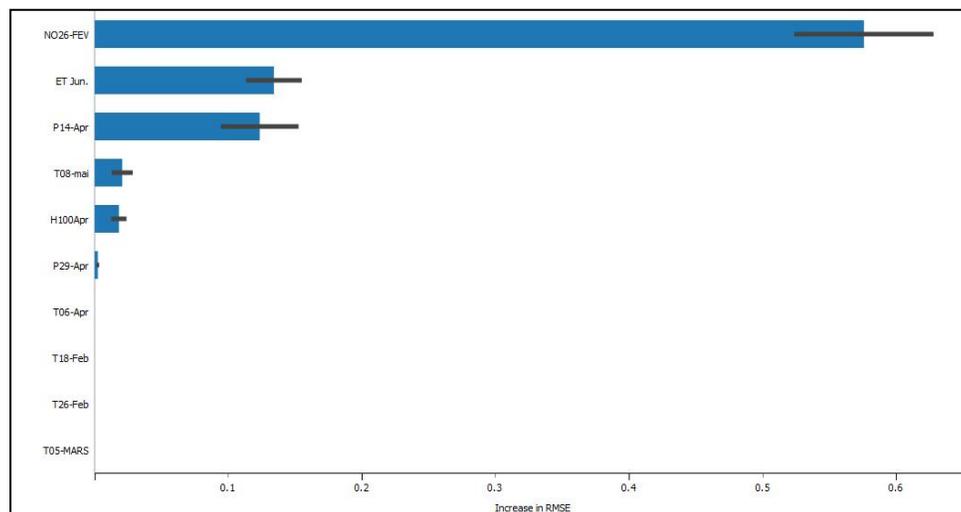
**Figure (62) :** Les dix meilleurs variables de prédiction du rendement du blé au niveau parcellaire.

### ➤ Extraction des meilleurs variables pour la prévision du rendement d'orge

En ce qui concerne l'orge, les 10 variables les plus significatives dans la prévision du rendement de l'orge avec le modèle "Tree" sont : le NDVI du 26 Février, l'évapotranspiration (ET) du mois de Juin, les précipitations du 14 Avril et du 29 Avril, la température moyenne du 08 Mai, ainsi que l'humidité à 100 cm du mois d'Avril (figure 63).

Ci-dessous une explication de l'importance des différents prédicteurs dans la détermination du rendement de l'orge au niveau parcellaire :

- ✓ Le NDVI du 26 Février, un NDVI élevé indique une végétation dense et saine, ce qui est généralement associé à des rendements plus élevés pour les cultures céréalières.
- ✓ l'évapotranspiration (ET), son impact sur le rendement de l'orge souligne l'importance de maintenir un équilibre approprié d'eau dans le sol pour favoriser une croissance optimale de cette culture
- ✓ L'humidité du sol à une profondeur de 100 cm en Avril, est un indicateur crucial qui peut affecter directement le développement et la croissance des racines des céréales. Un sol bien humidifié favorise la croissance des racines et permet aux plantes d'absorber plus efficacement l'eau et les nutriments essentiels. Par contre un sol trop sec peut entraîner un stress hydrique, affectant négativement la croissance, le développement et le rendement des céréales.

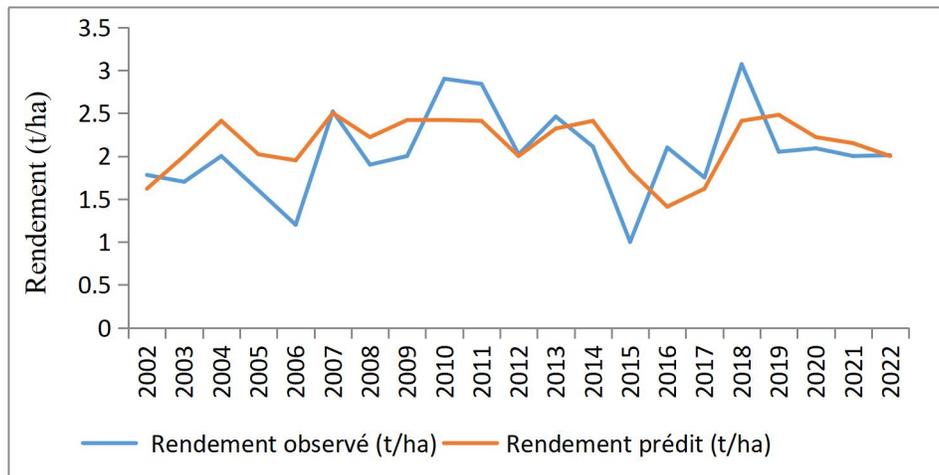


**Figure (63) :** Les dix meilleurs variables de prédiction du rendement de l'orge au niveau parcellaire

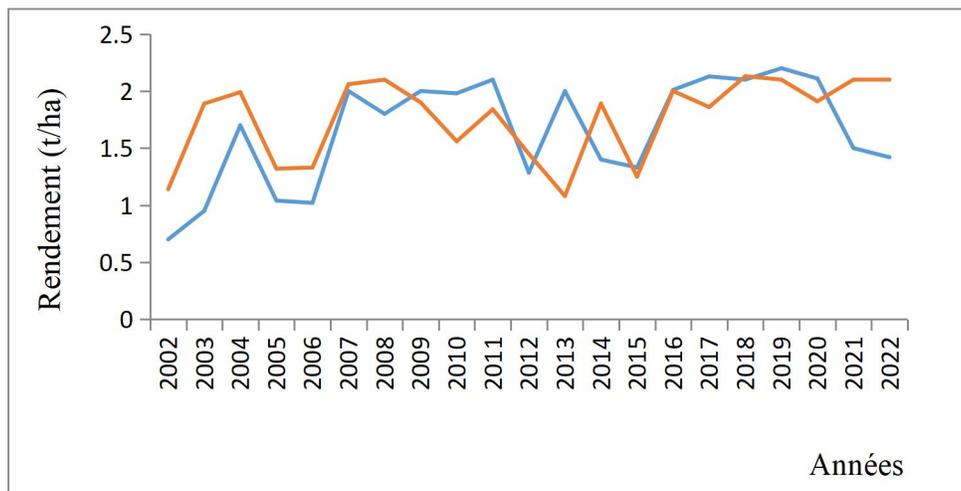
### 1.3.3. Evaluation de la prédiction des meilleurs modèles

Les résultats de la prédiction des rendements obtenus à partir des meilleurs modèles de prévision ("KNN" pour le blé et "Tree" pour l'orge) montrent une corrélation linéaire moyenne, entre les rendements observés et prédits, avec des coefficients de corrélation (R) de 0.56 pour le blé et 0.50 pour l'orge. Les valeurs de RMSE, variant de 0.414 à 0.456 t/ha, suggèrent que les prédictions du modèle sont moyennement proches des valeurs réelles. De même, les valeurs de

MSE, allant de 0.172 à 0.208, indiquent que les prédictions du modèle sont moins précises (figures 64, 65).



**Figure (64) :** Rendements prédit et observé du blé au niveau parcellaire obtenus par le modèle "KNN" entre 2002-2022

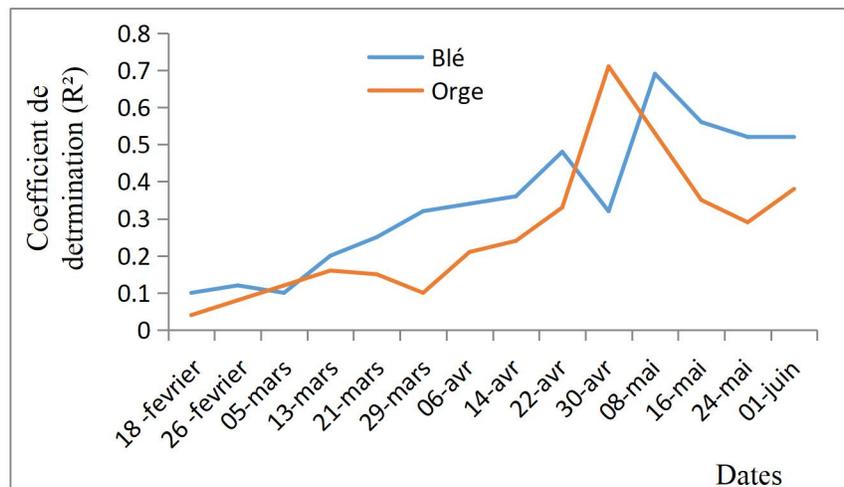


**Figure (65) :** Rendements prédit et observé d'orge au niveau parcellaire obtenus par le modèle "Tree" entre 2002-2022

## 2. Prédiction des rendements au niveau de la commune de Sétif

### 2.1. Prédiction des rendements en utilisant le NDVI

Les résultats de corrélation entre le rendement en grains de blé et d'orge avec les valeurs de NDVI montrent une forte relation linéaire au niveau de la commune de Sétif. Les coefficients de déterminations sont présentés graphiquement sur la figure (66).



**Figure (66) :** Evolution du coefficient de détermination entre le rendement et le NDVI pour les deux céréales (orge et blé), couvrant la période 2002 -2022 dans la commune de Sétif

La corrélation entre le NDVI et le rendement en grains devient de plus en plus significative au fur et à mesure de l'avancement de la campagne agricole, débutant à partir du 26 Février ( $R^2=0.1$  pour le blé,  $R^2=0.04$  pour l'orge), avec une notable diminution entre le 29 Mars pour l'orge seulement, qui est probablement attribuable à des conditions climatiques défavorables (nuages) pendant cette période. Par la suite, cette corrélation atteint son pic le 08 Mai pour le blé avec ( $R^2= 0.69$ ) et le 30 Avril pour l'orge ( $R^2= 0.71$ ) (figure 66).

On peut dire que le moment optimal de prédiction du rendement en grains du blé et de l'orge à l'aide de MODIS-NDVI au niveau de la commune de Sétif est le 08 mai pour le blé et le 30 Avril pour l'orge (environ 150 à 158 jours après le semis pour les deux cultures), qui correspond au stade "remplissage des grains".

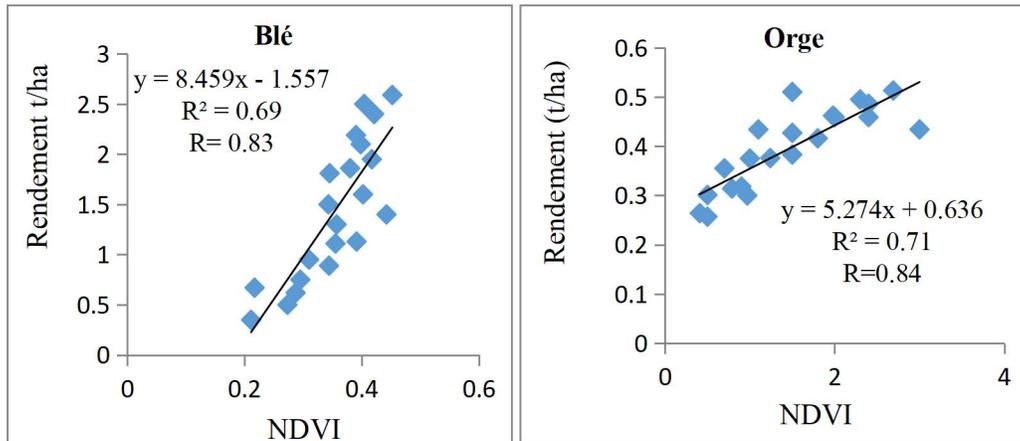
Pour élaborer les modèles de régression linéaires, les valeurs du NDVI du 08 Mai pour le blé et du 30 Avril pour l'orge (entre 2002 et 2022) ont été utilisées comme variables indépendantes, tandis que les rendements moyens de blé et d'orge de la commune de Sétif ont été considérés comme variables dépendantes. Les résultats sont présentés graphiquement dans la figure (67), indiquant une forte corrélation entre le NDVI et le rendement en grains pour les deux types de céréales, avec des coefficients de détermination  $R^2= 0.71$  pour l'orge et  $0.69$  pour le blé. Cela démontre que le NDVI a une forte influence sur le rendement de blé et d'orge au niveau communale.

L'estimation des rendements de blé suit la relation suivante :

$$\text{Rendements blé} = 8.459 \times \text{NDVI}_{(08 \text{ mai})} + 1.557$$

Pour l'orge, l'estimation des rendements suit l'équation suivante :

$$\text{Rendements orge} = 5.274 \times \text{NDVI}_{(30 \text{ avril})} + 0.636$$



**Figure (67) :** Modèle de régression linéaire entre le rendement et le NDVI le 08 Mai pour le blé et le 30 Avril pour l'orge au niveau de la commune de Sétif.

### 2.1.1. Vérification de la performance des modèles

Les résultats de la prédiction des rendements à l'aide du NDVI au niveau de la commune de Sétif ont révélé une forte corrélation entre les rendements réels et prédits, les coefficients de corrélation (R) de 0.84 pour le blé et 0.85 pour l'orge indiquent une forte corrélation linéaire entre les rendements observés et les rendements prédits, au niveau de signification de 0.05 ( $\alpha = 5\%$ ). Les faibles valeurs du RMSE, qui varient entre 0.363 et 0.472 t/ha, indiquent que les rendements simulés sont proches des rendements observés. De même, les valeurs de MSE (0.127 à 0.212) indiquent la bonne précision des modèles élaborés (figures 68, 69, tableau 15).

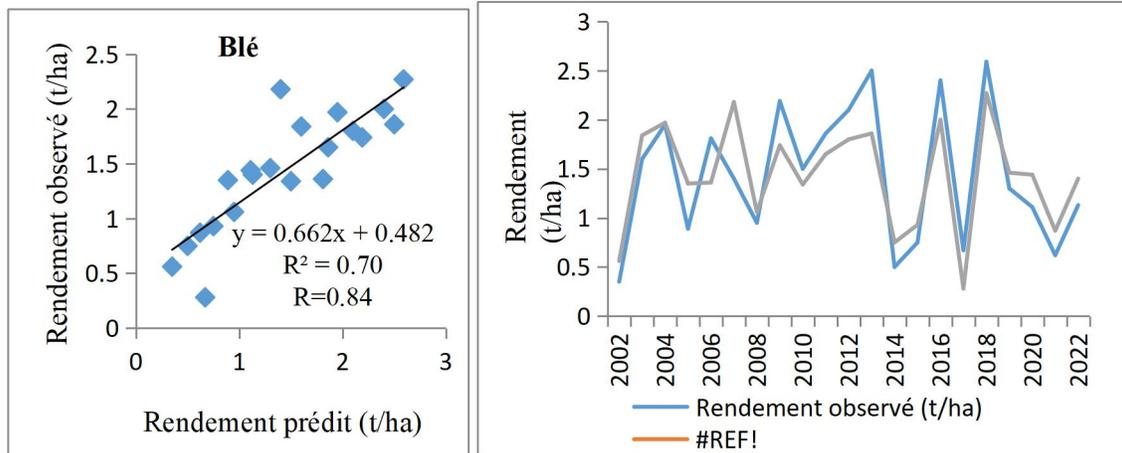


Figure (68) : Corrélation entre les rendements observés et prédits selon le modèle élaboré pour le blé en utilisant le NDVI au niveau de la commune de Sétif

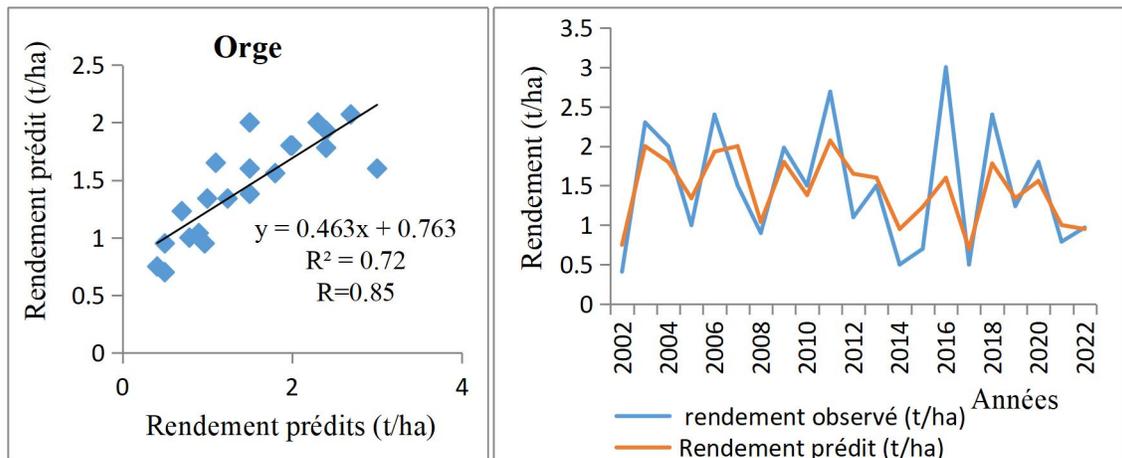


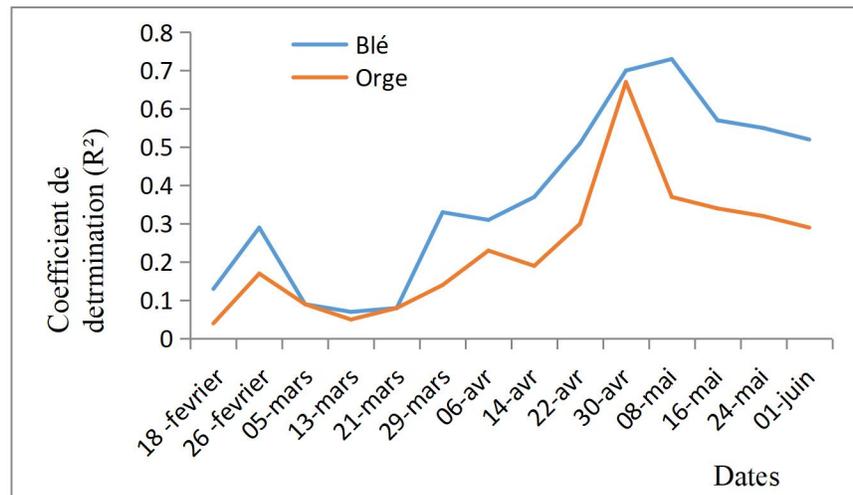
Figure (69) : Corrélation entre les rendements observés et prédits selon le modèle élaboré pour l'orge en utilisant le NDVI au niveau de la commune de Sétif

Tableau (15) : Résultats de performances des modèles au niveau de la commune de Sétif en utilisant le NDVI.

Cultures	R <sup>2</sup>	R	RMSE	MSE	Equation
Ogre	0.72	0.85	0.363	0.127	Y=0.463x-0.763
Blé	0.70	0.84	0.472	0.212	Y=0.662x-0.482

### 2.2. Prédiction des rendements en utilisant le NDVI et les précipitations

D'après la figure (70), on peut remarquer qu'il existe une corrélation linéaire significative entre l'utilisation combinée du MODIS-NDVI et la pluviométrie d'une part et le rendement des deux céréales d'hiver (blé et orge) au niveau régionale (commune de Sétif) d'autre part.



**Figure (70) :** Evolution du coefficient de détermination dans le cas de la combinaison du NDVI et de la pluviométrie dans la prévision du rendement des deux céréales (orge et blé) au niveau communale.

La corrélation entre la combinaison du NDVI et de la pluviométrie avec le rendement au niveau de la commune de Sétif devient de plus en plus forte au fur et à mesure de l'avancement de la campagne agricole, à partir du 21 Mars, atteignant un pic le 08 Mai pour le blé et le 30 Avril pour l'orge. Pour l'orge, les valeurs de  $R^2$  variaient de 0,04 à 0,67, tandis que pour le blé le  $R^2$  variait de 0,07 à 0,73. Par la suite, cette relation tend à diminuer progressivement à des dates ultérieures (figure 70).

D'après ces résultats, il est possible de déduire que le moment optimal pour prédire le rendement du blé et d'orge au niveau de la commune de Sétif en utilisant la combinaison du NDVI et de la pluviométrie, est la fin du printemps, plus précisément le 08 Mai pour le blé et le 30 Avril pour l'orge (environ 150 à 158 jours après le semis pour les deux cultures respectivement). Ces dates coïncident avec le stade de "remplissage des grains", où les conditions environnementales influent de manière significative sur le développement final des cultures céréalières.

Une analyse de régression linéaire multiple a été effectuée afin de développer des modèles de prédiction des rendements céréalières, on utilisant les valeurs du NDVI et de la pluviométrie du 08 Mai pour le blé et du 30 Avril pour l'orge (entre 2002 et 2022) comme variables indépendantes, tandis que les rendements en grains de blé et d'orge ont été pris comme variables dépendantes. Une corrélation positive forte a été constaté, entre le NDVI et la pluviométrie avec le rendement en grain de blé, avec un coefficient de détermination ( $R^2 = 0.73$ ). De manière similaire, pour l'orge, on observe une corrélation positive forte, avec un coefficient de détermination ( $R^2 = 0.67$ ).

D'après les régressions linéaires multiples, l'estimation des rendements de blé suit l'équation suivante :

$$\text{Rendement blé} = 10.946 \times \text{NDVI}_{(08 \text{ mai})} - 0.004 \times \text{pluviométries}_{(08 \text{ mai})} - 1.66$$

Et celle de l'orge suit l'équation suivante :

$$\text{Rendement orge} = 6.65 \times \text{NDVI}_{(30 \text{ avril})} - 0.001 \times \text{pluviométries}_{(30 \text{ avril})} - 1.438$$

### **2.2.1. Vérification de la performance des modèles**

Les résultats de la prédiction des rendements basée sur l'NDVI et la pluviométrie au niveau de la commune de Sétif mettent en évidence une corrélation significative entre les rendements réels et prédits, avec des coefficients de corrélation ( $R$ ) de 0.89 pour le blé et 0.87 pour l'orge, au niveau de signification de 0.05 ( $\alpha = 5\%$ ). Les valeurs réduites de RMSE, oscillant entre 0.343 et 0.408 t/ha, indiquent une proximité des prédictions du modèle avec les valeurs réelles. De même, les faibles valeurs de l'erreur quadratique moyenne entre les rendements observés et prédits (MSE), variant entre 0.113 et 0.243 t/ha, suggèrent une précision générale des prédictions du modèle (figures 71 et 72, tableau 16).

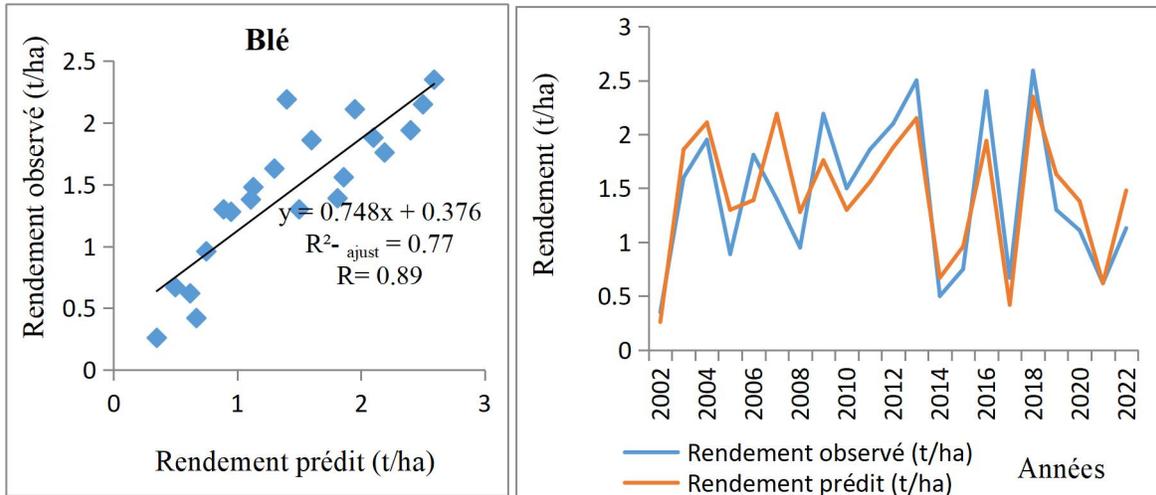


Figure (71) : Corrélation entre les rendements prédits et observés du blé au niveau de la commune de Sétif en utilisant le NDVI et la pluviométrie de la période 2002-2022.

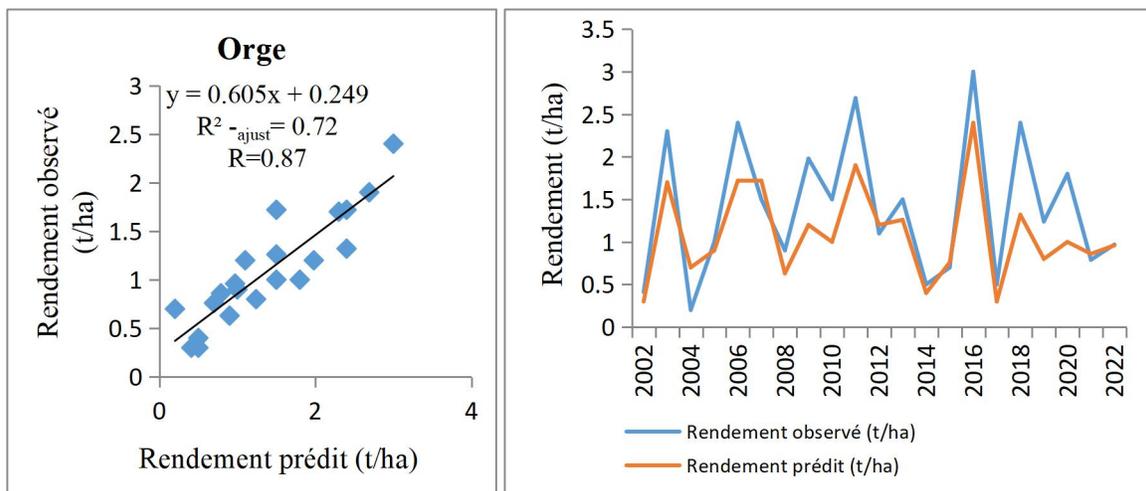


Figure (72) : Corrélation entre les rendements prédits et observés d’orge au niveau de la commune de Sétif en utilisant le NDVI et la pluviométrie de la période 2002-2022.

Tableau (16) : Résultats de performances des modèles au niveau de la commune de Sétif en utilisant le NDVI et les précipitations.

Cultures	R <sup>2</sup> -ajust	R	RMSE	MSE	Equation
Ogre	0.72	0.87	0.408	0.246	Y=0.605x + 0.245
Blé	0.77	0.89	0.343	0.113	Y=0.748x + 0.376

### **2.3. Prédiction des rendements avec les algorithmes d'apprentissage**

#### **2.3.1. Entraînement des modèles en utilisant tous les variables au niveau communal**

Le tableau (17) met en lumière que le modèle "Gradient Boosting" se distingue comme le plus précis et le mieux adapté pour la prédiction des rendements de blé au niveau de la commune de Sétif, affichant un coefficient de détermination ( $R^2$ ) de 0.405, expliquant ainsi environ 40.5% de la variation des rendements. Ensuite, le modèle "Random Forest" se classe en deuxième position avec un  $R^2$  de 0.394, suivi par le modèle "AdaBoosting" en troisième position avec un  $R^2$  de 0.157. Cependant, certains modèles tels que la "régression linéaire", le "Neural Network" et le "Stochastic Gradient Descent" présentent des performances moindres avec des  $R^2$  négatifs, ce qui indique une adéquation limitée pour la prédiction des rendements de blé dans ce contexte.

Pour ce qui est de l'orge, les résultats révèlent des défis significatifs en termes de performance des modèles de régression. Presque tous les modèles de régression montrent des valeurs négatives pour le coefficient de corrélation ( $R$ ), suggérant une corrélation inverse ou très faible entre les variables prédictives et les rendements observés. Le modèle "SVM" apparait le meilleur, avec un coefficient de corrélation ( $R$ ) relativement faible de 0.01, avec un RMSE de 0.760 t/ha et un MSE de 0.578 t/ha.

Ces constats soulignent les défis rencontrés lors de l'utilisation de toutes les variables disponibles pour entraîner les modèles à prédire le rendement de l'orge au niveau de la commune de Sétif. Les faibles performances des modèles de régression suggèrent une capacité limitée à capturer les relations significatives entre les variables utilisées et les rendements observés. Cette situation pourrait nécessiter une révision des variables utilisées ou une exploration plus approfondie des modèles afin d'améliorer la prédiction du rendement de l'orge dans ce contexte spécifique.

**Tableau (17):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant toutes les variables au niveau de la commune de Sétif

Modèles	Blé			Orge		
	R <sup>2</sup>	RMSE	MSE	R <sup>2</sup>	RMSE	MSE
<b>Regression lineaire</b>	-0.823	17.909	0.987	-0.933	1.515	2.295
<b>KNN</b>	0.157	0.573	0.328	-0.234	0.848	0.720
<b>Random Forest</b>	0.394	0.485	0.236	-0.280	0.864	0.747
<b>Tree</b>	0.06	0.605	0.366	-0.430	0.913	0.834
<b>SVM</b>	0.05	0.608	0.369	<b>0.010</b>	<b>0.760</b>	<b>0.578</b>
<b>Gradient Boosting</b>	<b>0.405</b>	<b>0.481</b>	<b>0.231</b>	-0.390	0.901	0.811
<b>Neural Network</b>	-0.744	0.823	0.678	-0.386	0.899	0.809
<b>AdaBoost</b>	0.32	0.514	0.264	-0.057	0.785	0.617
<b>Stochastic Gradient Descent</b>	-0.768	17.293	0.991	-0.924	1.695	2.874

### 2.3.2. Prédiction des rendements en utilisant les dix meilleurs variables

Le tableau (18), révèle que lors de l'utilisation des dix meilleurs variables, le modèle "Stochastic Gradient Descent" semble être le plus précis et le plus adapté pour prédire les rendements de blé au niveau de la commune de Sétif, avec un R<sup>2</sup> de 0.651, un RMSE de 0.368 et un MSE de 0.136 . Ensuite, le modèle "Ada Boost" se classe le deuxième avec R<sup>2</sup>=0.560, suivi par le modèle "Gradient Boosting" (R<sup>2</sup>=0.557) en troisième position. En revanche, certains modèles comme "la régression linéaire", "tree" et "Rondom Forest" présentent des performances plus faibles avec des R<sup>2</sup> moyennes ou négatifs.

Pour l'orge, le modèle "Gradient Boosting" semble être le plus précis et le plus adapté pour prédire les rendements au niveau de la commune de Sétif, avec un R<sup>2</sup> de 0.358 qui explique

environ 35.8% de la variation des rendements. Le modèle "Stochastic Gradient Descent" ( $R^2=0.194$ ) se classe ensuite en deuxième position, suivi par le modèle "AdaBoost" ( $R^2=0.170$ ) en troisième position. Cependant, certains modèles comme "la régression linéaire", "KNN" et "Neural Network" présentent des performances plus faibles avec des  $R^2$  négative et très faibles. Ces résultats mettent en évidence l'importance du choix du modèle dans la précision des prédictions de rendement de blé et d'orge dans la commune de Sétif. Il est recommandé d'approfondir l'exploration et l'optimisation des modèles les plus performants pour obtenir les meilleures prédictions possibles. De plus, il est remarquable que le rendement de blé a été prédit avec une plus grande précision que celui du d'orge au niveau de la commune de Sétif.

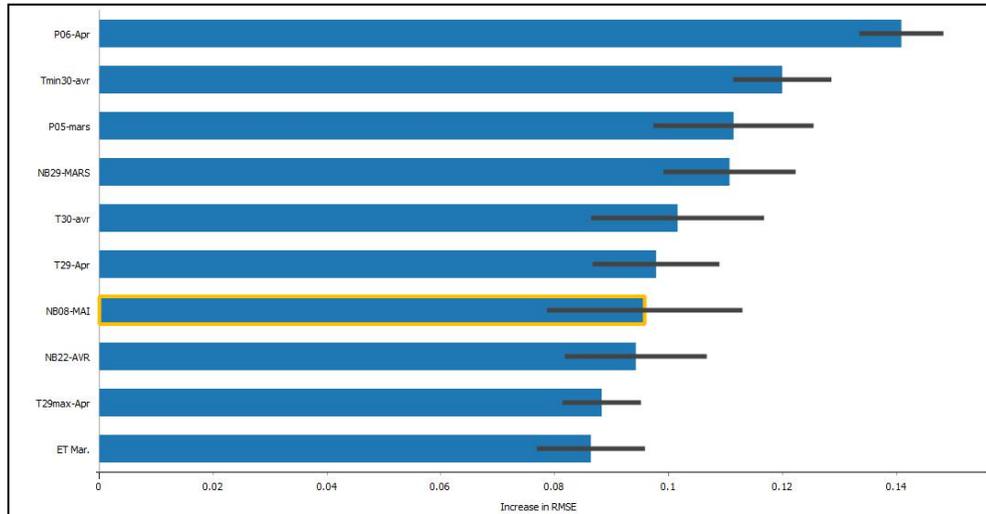
**Tableau (18) :** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant les dix meilleurs variables au niveau de la commune de Sétif

Modèles	Blé			Orge		
	R <sup>2</sup>	RMSE	MSE	R <sup>2</sup>	RMSE	MSE
<b>Regression lineaire</b>	-0.490	0.984	0.968	-0.911	1.511	2.282
<b>KNN</b>	0.375	0.493	0.243	0.072	0.736	0.542
<b>Random Forest</b>	0.305	0.520	0.270	0.118	0.718	0.515
<b>Tree</b>	0.279	0.530	0.280	0.117	0.718	0.515
<b>SVM</b>	0.397	0.484	0.234	0.146	0.706	0.499
<b>Gradient Boosting</b>	0.557	0.415	0.172	<b>0.358</b>	<b>0.612</b>	<b>0.375</b>
<b>Neural Network</b>	0.493	0.444	0.197	0.108	0.722	0.521
<b>AdaBoost</b>	0.560	0.414	0.171	0.170	0.696	0.484
<b>Stochastic Gradient Descent</b>	<b>0.651</b>	<b>0.368</b>	<b>0.136</b>	0.194	0.688	0.484

### ➤ **Extraction des meilleurs variables pour la prévision du rendement du blé**

D'après la figure (73), les précipitations du 06 avril et du 05 mars, ainsi que la température minimale du 30 avril, la température maximale du 29 avril, les températures moyennes du 30 avril, l'évapotranspiration (ET) du mois de mars, et le NDVI du 29 mars, du 08 mai et du 22 avril, se distinguent comme les dix meilleures variables combinées pour prédire le rendement du blé au niveau de la commune de Sétif. L'importance de ces variables dans la détermination du rendement du blé peut être expliquée comme suit:

- ✓ Les précipitations des mois de mars et avril sont cruciales car elles surviennent respectivement au début du cycle de croissance végétative du blé et pendant la phase de floraison et de formation des épis. Des précipitations adéquates pendant ces périodes fournissent l'humidité nécessaire pour favoriser une croissance végétative saine et la formation des épis. Un manque d'eau peut entraîner un stress hydrique et ainsi affecter négativement le rendement du blé.
- ✓ Les températures basses de la fin d'avril, notamment les gelées tardives, influencent la croissance de la plante, en particulier pendant la phase de floraison, pouvant avoir un impact négatif sur la formation et la maturation des grains de blé.
- ✓ L'évapotranspiration du mois de mars est un indicateur crucial de la quantité d'eau perdue par évaporation du sol et de transpiration des plantes. Une évapotranspiration élevée peut indiquer des besoins accrus en eau pour les cultures de blé pendant cette période critique.
- ✓ Les valeurs du NDVI des dates sélectionnées (29 mars, 08 mai et 22 avril) offrent des informations précieuses sur la densité et la santé de la végétation. Un NDVI élevé est généralement associé à une végétation dense et saine, ce qui correspond souvent à des rendements plus élevés du blé.



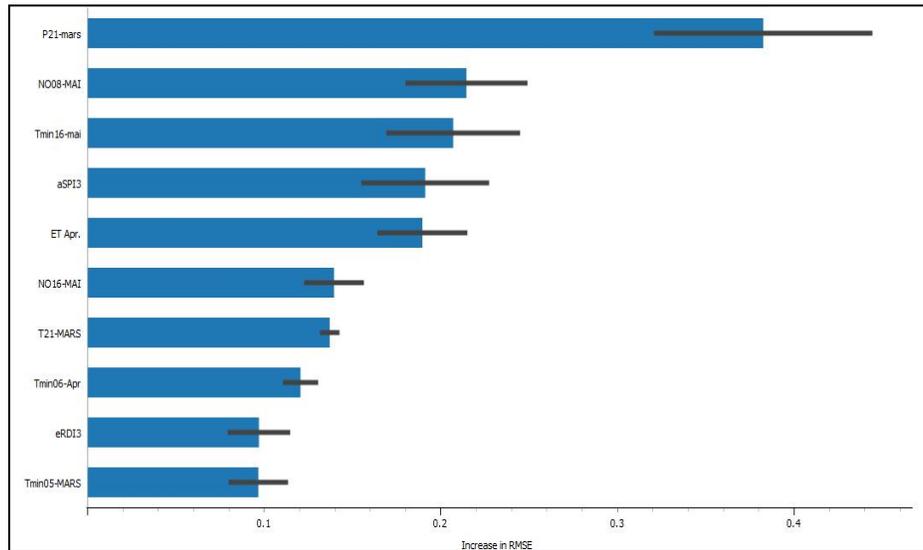
**Figure (73) :** Les dix meilleurs variables de prédiction du rendement du blé au niveau communale.

### ➤ Extraction des meilleurs variables pour la prévision du rendement de l'orge

Conformément à la figure (74), les variables les plus significatives pour la prédiction du rendement de l'orge au niveau de la commune de Sétif comprennent : les précipitations du mois de Mars, le NDVI du 08 Mai et du 16 Mai, l'évapotranspiration (ET) du mois d'Avril, la température moyenne du 21 Mars, les températures minimales du 06 Avril et du 05 Mars, ainsi que l'indice de sécheresse de reconnaissance efficace (eRDI) et l'indice de précipitation agricole standardisé (aSPI) des trois mois ( de Février à Avril). Ces facteurs sont essentiels pour prédire de manière fiable le rendement d'orge dans cette région, étant sélectionnés pour leur importance potentielle dans le processus de croissance de cette culture :

- ✓ Des précipitations adéquates au début du printemps favorisent une croissance initiale, vigoureuse des céréales, contribuant ainsi à des rendements plus élevés.
- ✓ Un NDVI élevé pour l'orge au début et à la mi-Mai indique une végétation dense et saine, généralement associée à des rendements accrus pour cette culture.
- ✓ Les températures très basses à début Avril et la mi-Mai peuvent affecter la floraison et la formation des épis l'orge. Des températures modérées sont bénéfiques pour un développement optimal de la culture.
- ✓ Une évapotranspiration élevée en Avril peut signaler des conditions sèches qui pourraient avoir un impact négatif sur le rendement de l'orge.

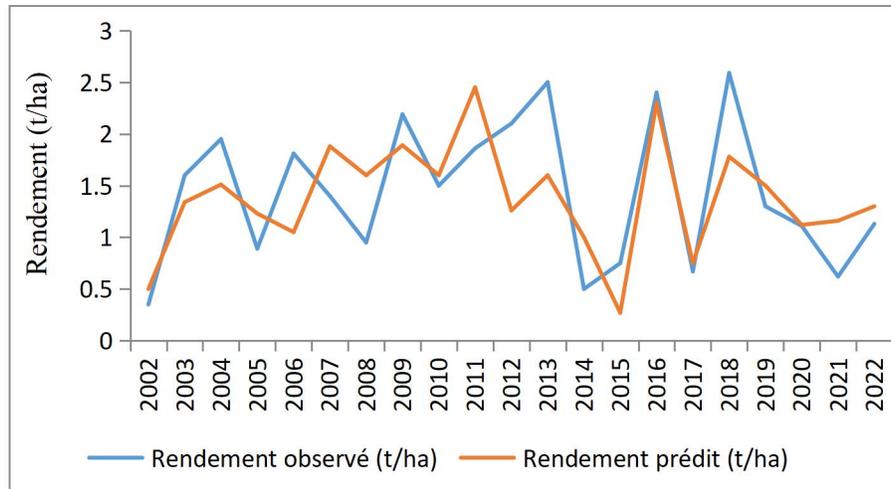
- ✓ L'indice de Précipitation Agricole Standardisé (aSPI3) et l'indice de Sécheresse de Reconnaissance Efficace (eRDI3) des mois de Février, Mars et Avril offre des connaissances sur les conditions de sécheresse ou d'humidité durant cette phase critique du cycle de croissance de l'orge, pouvant influencer le rendement.



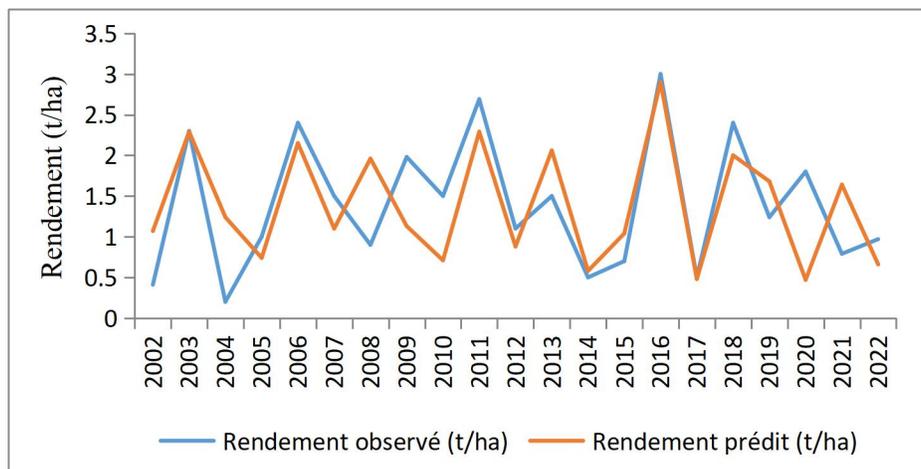
**Figure (74) :** Les dix meilleures variables de prédiction du rendement de l'orge au niveau communale

### 2.3.3. Evaluation de la prédiction des meilleurs modèles

Les résultats de la prédiction des rendements élaborés par les modèles "Stochastic Gradient Descent" pour le blé et "Gradient Boosting" pour l'orge, montrent des bonnes corrélations linéaires entre les rendements observés et prédits, avec des coefficients de corrélation (R) de 0.69 pour le blé et 0.66 pour l'orge. Avec des valeurs de RMSE qui sont relativement basses variant de 0.368 à 0.612 t/ha, suggèrent que les prédictions du modèle sont éloignées des valeurs réelles. De même, les valeurs de MSE allant de 0.136 à 0.375, ce qui suggère une précision généralement acceptable des prédictions des modèles (figures 75, 76).



**Figure (75) :** Comparaison entre les rendements du blé observés et prédits par le modèle "Stochastic Gradient Descent" au niveau de la commune de Sétif

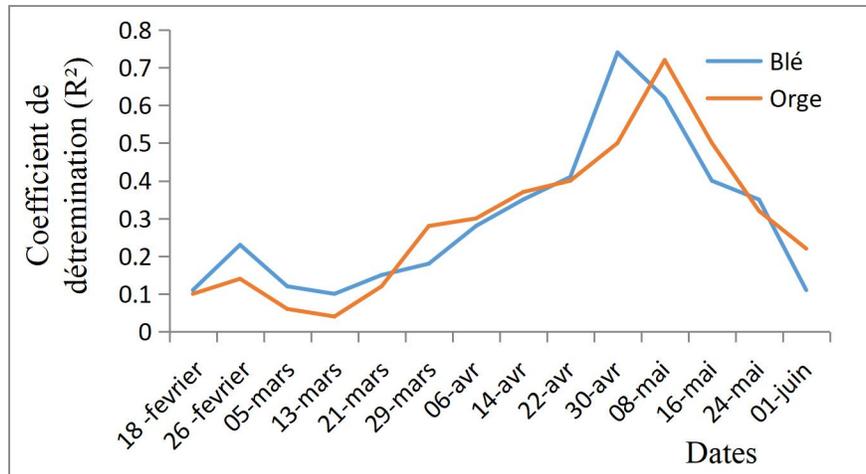


**Figure (76) :** Comparaison entre les rendements d'orge observés et prédits par le modèle "Gradient Boosting" au niveau de la commune de Sétif

### 3. Prédiction des rendements au niveau de la wilaya de Sétif

#### 3.1. Prédiction des rendements par le NDVI

Les résultats de corrélation entre le rendement en grains de blé et de l'orge avec les valeurs de NDVI révèlent une forte relation linéaire positive au niveau de la wilaya de Sétif. L'évolution des coefficients de déterminations sont présentés dans la figure (77).



**Figure (77) :** Evolution du coefficient de détermination entre le rendement en grains et le NDVI pour les deux céréales (orge et blé) couvrant la période de 2002 à 2022 dans la wilaya de Sétif

La corrélation entre le NDVI et le rendement en grains devient de plus en plus significative au fur et à mesure de l'avancement de la campagne agricole, débutant à partir du 21 mars ( $R^2=0.15$  pour le blé,  $R^2=0.12$  pour l'orge). Cette corrélation atteint ensuite son pic le 30 Avril pour le blé ( $R^2= 0.74$ ) et le 08 Mai pour l'orge ( $R^2= 0.72$ ). Cependant, cette relation entre les variables diminue progressivement à des dates ultérieures, probablement en raison de la saturation du NDVI lors des derniers stades de croissance des céréales (voir figure 77).

D'après ces résultats, le moment optimal pour prédire le rendement en grains du blé et de l'orge à l'aide du MODIS-NDVI au niveau de la wilaya de Sétif est la fin du printemps, plus précisément le 08 Mai pour l'orge et le 30 Avril pour le blé (environ 150 à 158 jours après le semis pour les deux cultures), correspondant au stade "remplissage des grains".

Une analyse de régression linéaire a été menée pour élaborer des modèles de prédiction des rendements céréaliers. Les valeurs du NDVI du 30 Avril pour l'orge et du 08 Mai pour le blé (entre 2002 et 2022) ont servi comme variables indépendantes, tandis que les rendements moyens de blé et d'orge de la wilaya de Sétif ont été considérés comme variables dépendantes. Les résultats, présentés graphiquement dans la figure 78, indiquent une forte relation entre le NDVI et le rendement en grains pour les deux types de céréales, avec des coefficients de détermination  $R^2=0.74$  pour le blé et  $0.72$  pour l'orge. Cela signifie que 74% et 72% de la variabilité du

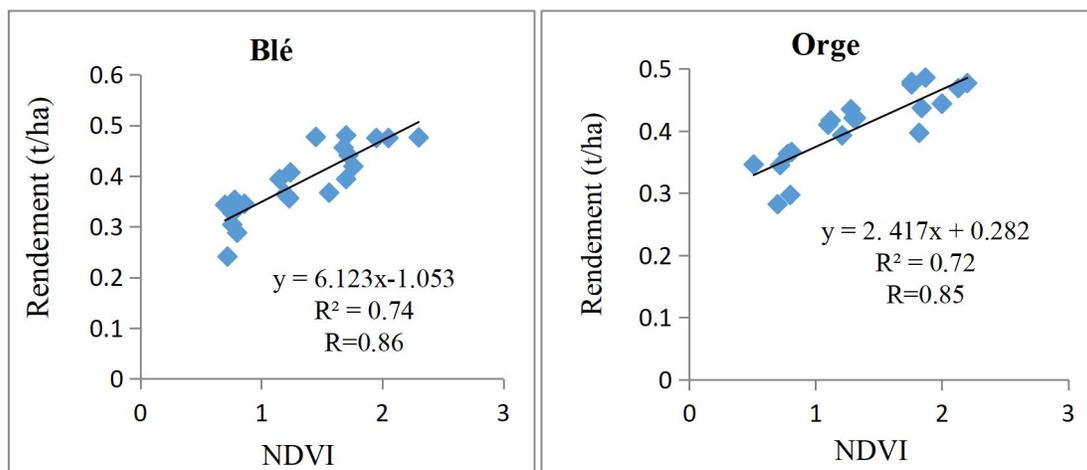
rendement d'orge et de blé peuvent être expliqués par le NDVI, soulignant ainsi son influence significative sur les rendements. Les coefficients de régression pour le blé et l'orge indiquant que chaque augmentation de 0.1 du NDVI est associée à une augmentation moyenne de 0.612 t/ha et 0.241 t/ha du rendement en grains pour le blé et l'orge.

Selon le modèle de régression élaboré, l'estimation des rendements de blé suit l'équation suivante :

$$\text{Rendement blé} = 6.123 \times \text{NDVI}_{(30 \text{ avril})} - 1.053$$

Pour de l'orge Le calcul des rendements suit l'équation suivante :

$$\text{Rendement orge} = 2.417 \times \text{NDVI}_{(08 \text{ mai})} + 0.282$$

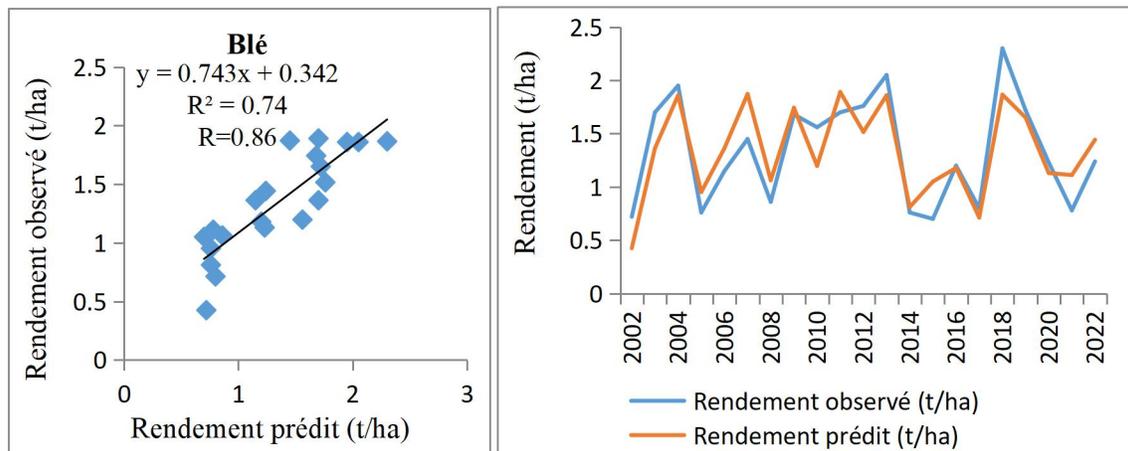


**Figure (78) :** Modèle de régression linéaire entre le rendement et le NDVI, le 30 Avril pour le blé et le 08 Mai pour l'orge au niveau de la wilaya de Sétif.

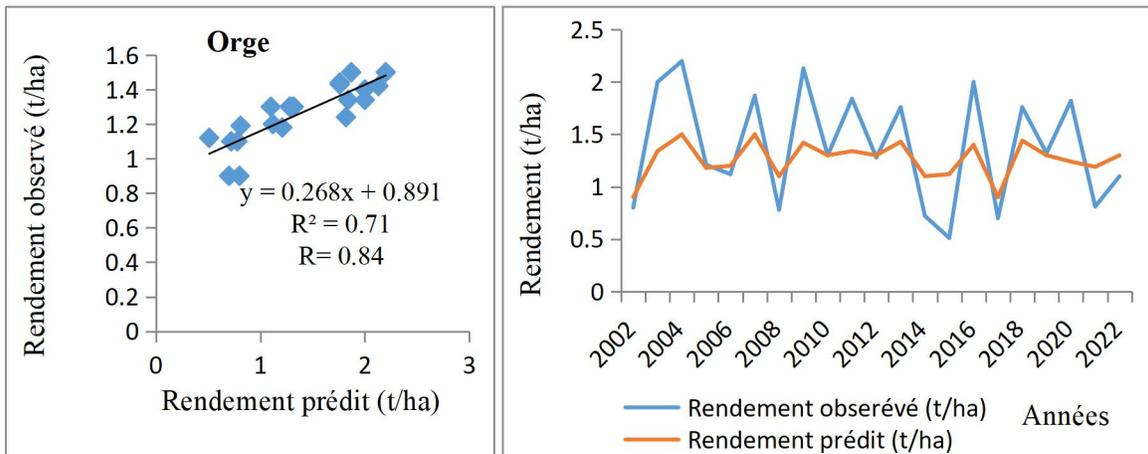
### 3. 1.1. Vérification de la performance des modèles

Les résultats de la prédiction des rendements basée sur le NDVI dans la wilaya de Sétif mettent en évidence une forte corrélation entre les rendements réels et prédits. Les coefficients de corrélation (R) de 0.86 pour le blé et 0.84 pour l'orge dénotent une corrélation linéaire solide entre les rendements observés et prévus, au niveau de signification de 0.05 ( $\alpha = 5\%$ ). De plus, les coefficients de détermination ( $R^2$ ) de 0.74 et 0.71 pour le blé et l'orge, respectivement, indiquent que 74% et 71% de la variabilité des rendements observés peuvent être expliqués par les modèles de régression.

Les valeurs faibles de RMSE, se situant entre 0.338 et 0.433 t/ha, montrent que les estimations du modèle sont étroitement alignées avec les valeurs réelles. De même, les valeurs de l'erreur quadratique moyenne entre les rendements observés et prédits (MSE), allant de 0.113 à 0.179, soulignent une bonne précision générale des prédictions du modèle. Ces résultats sont représentés dans les figures 79 et 80, ainsi que dans le tableau 19.



**Figure (79) :** Corrélation entre les rendements observés et prédits du blé en utilisant le NDVI



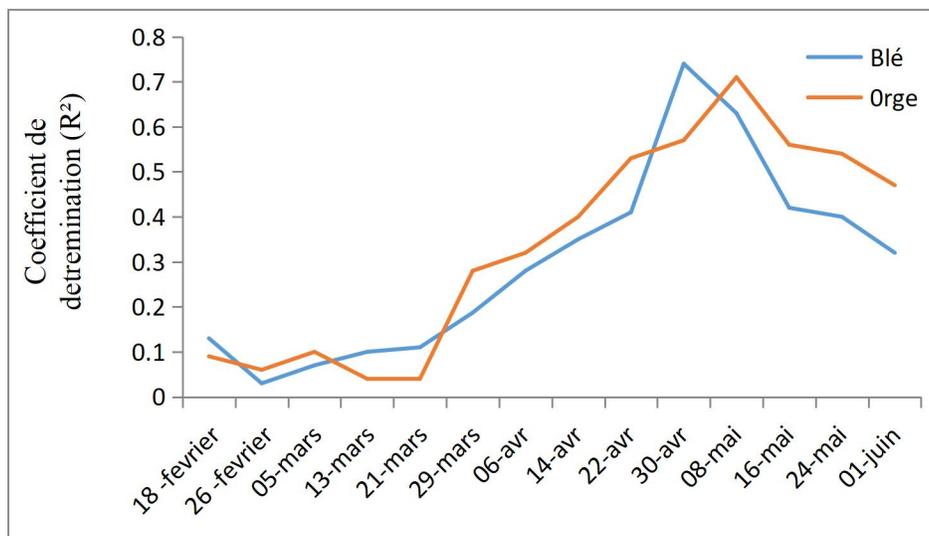
**Figure (80) :** Corrélation entre les rendements observés et prédits d'orge en utilisant le NDVI

**Tableau (19).** Résultats de performances des modèles au niveau de la wilaya de Sétif en utilisant le NDVI.

Cultures	R <sup>2</sup>	R	RMSE	MSE	Equation
Ogre	0.71	0.84	0.433	0.179	Y= 0.268 x + 0.891
Blé	0.74	0.86	0.338	0.113	Y= 0.743 x + 0.342

**3.2. Prédiction des rendements en utilisant le NDVI et les précipitations**

D’après la figure (81), on remarque qu’il existe une corrélation linéaire significative entre l’utilisation combinée du MODIS-NDVI et la pluviométrie d’une part et le rendement en grains pour les deux céréales d’hiver (blé et orge) d’autre part au niveau de la wilaya de Sétif. L’évolution des coefficients de déterminations sont présentés graphiquement dans la figure (81) ci- dessous.



**Figure (81) :** Evolution des coefficients de détermination entre le rendement, pour les deux céréales (orge et blé) en utilisant l’NDVI et la pluviométrie, couvrant la période de 2002- 2022 au niveau de la wilaya de Sétif.

On peut remarquer que le coefficient de détermination augmente de plus en plus au fur et à mesure de l’avancement de la campagne agricole, à partir du 21 Mars, atteignant un pic le 08 Mai pour l’orge et le 30 Avril pour le blé. Pour l’orge, les valeurs de R<sup>2</sup> variaient de 0.04 à 0.71.

Pour le blé, les valeurs de  $R^2$  variaient de 0.03 à 0.74. Par la suite, cette relation tend à diminuer progressivement à des dates ultérieures.

Sur la base de ces résultats, le moment idéal pour prédire le rendement d'orge au niveau de la wilaya de Sétif est 08 Mai, le 30 Avril pour le blé. Ces dates correspondent au stade de "remplissage des grains".

Les résultats indiquent une forte corrélation positive ( $R = 0.86$ ) entre le NDVI et la pluviométrie avec le rendement en grain de blé. Le coefficient de détermination ( $R^2 = 0.74$ ) montre que 74% de la variation du rendement en grain de blé peut être expliquée par la combinaison du NDVI et de la pluviométrie. De manière similaire, pour l'orge, on observe une forte corrélation positive ( $R = 0.85$ ) entre le NDVI et la pluviométrie avec le rendement en grain. Avec un ( $R^2 = 0.71$ ) révèle que 71% de la variation du rendement en grain d'orge peut être expliquée par le NDVI et la pluviométrie combinés, au niveau de signification de 0.05 ( $\alpha = 5\%$ ).

D'après les résultats de la régression linéaire multiple, l'estimation des rendements de blé suit l'équation suivante :

$$\text{Rendement blé} = 6.21x \text{ NDVI}_{(30 \text{ avril})} - 0.002 x \text{ pluviométries}_{(30 \text{ avril})} - 1.045$$

Et celle de l'orge suit l'équation suivante :

$$\text{Rendement orge} = 7.85x \text{ NDVI}_{(08 \text{ mai})} - 0.005 x \text{ pluviométries}_{(08 \text{ mai})} - 1.82$$

### 3.2.1. Vérification de la performance des modèles

Les résultats de la prédiction des rendements à l'aide du NDVI et de la pluviométrie au niveau de la wilaya de Sétif ont mis en évidence une forte relation linéaire entre les rendements anticipés et réels, avec des coefficients de corrélation ( $R = 0.88$ ) pour le blé et 0.87 pour l'orge, au niveau de signification de 0.05 ( $\alpha = 5\%$ ), avec un  $R^2_{\text{ajusté}} = 0.74$  et 0.73 pour le blé et l'orge respectivement. Cela indique que 74% et 73% de la variation des rendements observés de blé et d'orge peuvent être expliqués par les modèles de régression. Les valeurs de RMSE, comprises entre 0.277 et 0.346 t/ha. De même, les valeurs de l'erreur quadratique moyenne entre les rendements observés et prédits (MSE), oscillant entre 0.077 et 0.113, indiquent une bonne précision des prédictions des modèles élaborés (figures 82, 83, tableau 20).

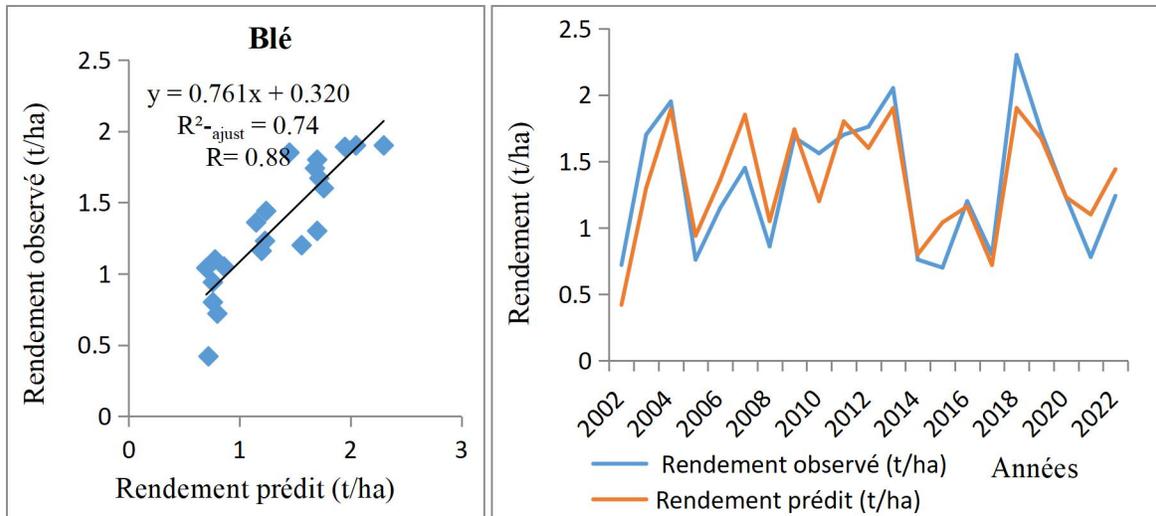


Figure (82): Relation entre les rendements prédits et observés du blé au niveau de la wilaya de Sétif en utilisant le NDVI et la pluviométrie entre 2002 et 2022.

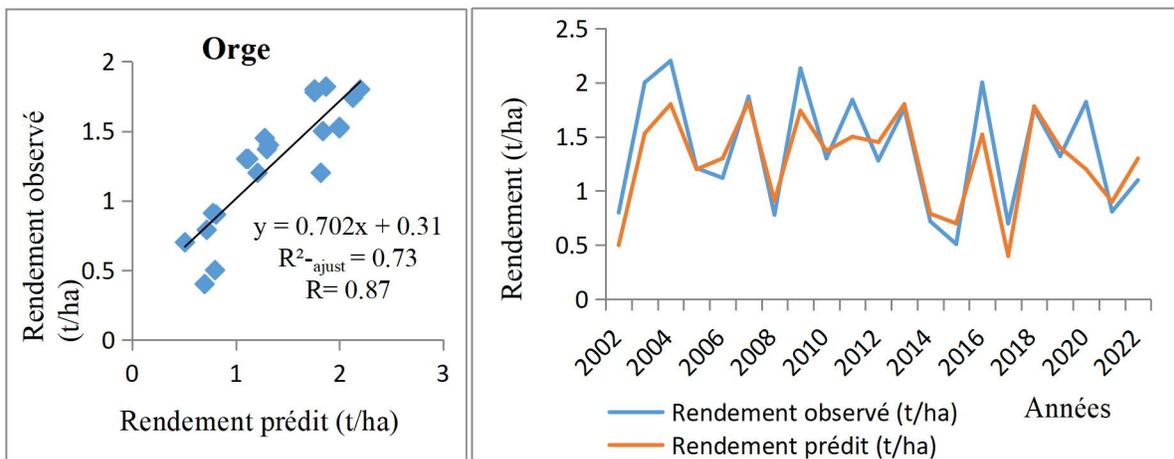


Figure (83) : Relation entre les rendements prédits et observés d’orge au niveau de la wilaya de Sétif en utilisant le NDVI et la pluviométrie entre 2002 et 2022

Tableau (20): Résultats de performances des modèles le NDVI et les précipitations au niveau de la wilaya de Sétif

Cultures	R <sup>2</sup> -ajust	R	RMSE	MSE	Equation
Ogre	0.73	0.87	0.343	0.113	Y= 0.702x + 0.31
Blé	0.74	0.88	0.277	0.077	Y= 0.761x + 0.320

On a pu remarquer que l'ajout de la pluviométrie comme facteur supplémentaire de prédiction a permis une légère amélioration des performances des modèles de prévision du rendement, car elle influence directement l'hydratation des plantes, la disponibilité en eau dans le sol et, par conséquent, la croissance et le développement des cultures.

### **3.3. Prédiction des rendements avec les algorithmes d'apprentissage**

#### **3.3.1. Entraînement des modèles en utilisant tous les variables au niveau wilaya**

Le tableau (21) met en évidence que lors de l'utilisation de toutes les variables disponibles, le modèle "AdaBoost" se distingue comme le plus précis pour prédire les rendements de blé dans la wilaya de Sétif, avec un  $R^2$  de 0.58, expliquant environ 58% de la variation des rendements, avec un RMSE de 0.321t/ha et un MSE de 0.103t/ha. Ensuite, le modèle "Stochastic Gradient Descent" prend la deuxième place avec un  $R^2$  de 0.51, suivi du modèle "de la régression linéaire" en troisième position avec un  $R^2$  de 0.456. Cependant, certains modèles comme "Neural Network", "SVM" et "Random Forest" affichent des performances inférieures, ce qui suggère une adéquation moindre pour la prédiction des rendements de blé dans cette étude.

Concernant l'orge, il semble que les résultats de l'analyse statistique pour la prédiction des rendements au niveau de la wilaya de Sétif en utilisant toutes les variables présentent des défis significatifs en termes de performance des modèles de régression. La plupart des modèles de régression montrent des valeurs négatives pour le coefficient de détermination, indiquant ainsi une corrélation inverse ou très faible entre les variables prédictives et les rendements observés.

En revanche, le modèle "AdaBoost" apparaît le meilleur modèle et présente un coefficient de détermination faible de 0.097, avec un RMSE de 0.472t/ha et un MSE de 0.223.

Ces résultats mettent en évidence les défis rencontrés lors de l'utilisation de toutes les variables disponibles pour prédire le rendement d'orge au niveau de la wilaya de Sétif. Les performances faibles ou négatives des modèles de régression suggèrent une capacité limitée à capturer les relations significatives entre les variables utilisés et les rendements observés. Cette situation pourrait nécessiter une sélection des variables utilisées et une exploration plus approfondie des modèles pour améliorer la prédiction du rendement de l'orge à cet échelle.

**Tableau (21):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant tout les variables au niveau de la wilaya de Sétif

Modèles	Blé			Orge		
	R <sup>2</sup>	RMSE	MSE	R <sup>2</sup>	RMSE	MSE
<b>Regression lineaire</b>	0.456	0.366	0.134	-0.363	0.580	0.337
<b>KNN</b>	0.356	0.398	0.158	-0.028	0.504	0.254
<b>Random Forest</b>	0.326	0.407	0.166	-0.537	0.616	0.280
<b>Tree</b>	0.336	0.404	0.163	-0.870	0.680	0.462
<b>SVM</b>	0.281	0.420	0.177	-0.032	0.489	0.239
<b>Gradient Boosting</b>	0.436	0.372	0.138	-0.031	0.709	0.502
<b>Neural Network</b>	0.277	0.421	0.178	-0.556	0.620	0.384
<b>AdaBoost</b>	<b>0.580</b>	<b>0.321</b>	<b>0.103</b>	<b>0.097</b>	<b>0.472</b>	<b>0.223</b>
<b>Stochastic Gradient Descent</b>	0.510	0.347	0.120	-0.272	0.561	0.314

### 3.3.2. Prédiction des rendements en utilisant les dix meilleurs variables

Le tableau (22) révèle que lors de l'utilisation les dix meilleurs variables, le modèle "Stochastic Gradient Descent" semble être le plus précis pour prédire les rendements de blé au niveau de la wilaya de Sétif, avec un R<sup>2</sup> de 0.783, avec un RMSE de 0.231t/ha et un MSE de 0.053. Ensuite, le modèle " Neural Network" se classe le deuxième avec (R<sup>2</sup>=0.609), suivi par le modèle "KNN" (R<sup>2</sup>=0.568) en troisième position. Certains modèles comme "la régression linéaire" présente des performances plus faibles avec des R<sup>2</sup> négatifs.

Pour l'orge, le modèle "Tree" semble être le plus adapté pour prédire les rendements d'orge au niveau de la wilaya de Sétif, avec un R<sup>2</sup> de 0.390 avec un RMSE de 0.388 t/ha et un MSE de 0.151. Puis, le modèle "Stochastic Gradient Descent" (R<sup>2</sup>=0.363) se classe en deuxième position, suivi par le modèle "Neural Network" (R<sup>2</sup>=0.331) en troisième position. Cependant, certains

modèles comme "la régression linéaire", "KNN" et "AdaBoost" présentent des valeurs plus faibles ou négatives.

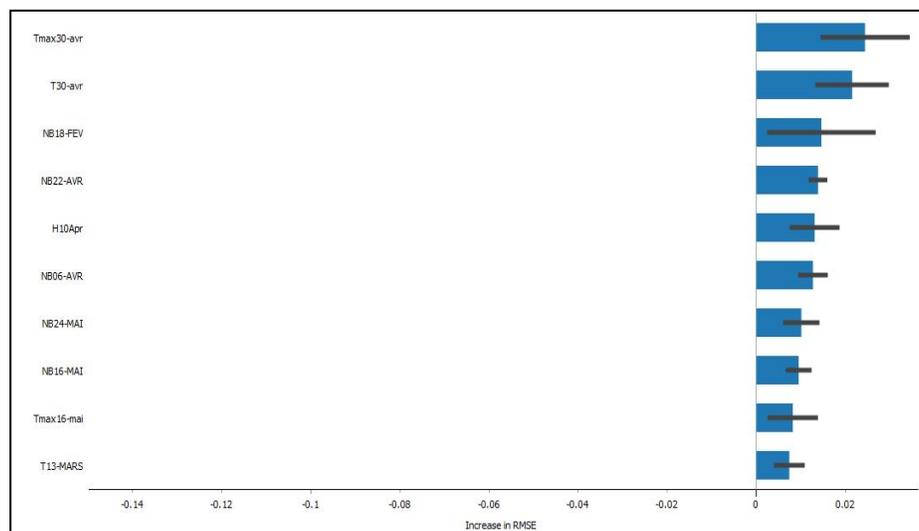
Ces résultats mettent en évidence l'importance du choix du modèle dans la précision des prédictions, ainsi que la sélection des meilleures variables dans la prévision du rendement de blé et d'orge dans la wilaya de Sétif. De plus, il a été constaté que le rendement de blé a été prédit avec une plus grande précision que celui d'orge.

**Tableau (22):** Résultats d'analyse de performance des modèles de prédiction des rendements de blé et d'orge en utilisant les dix meilleurs variables au niveau de la wilaya de Sétif

Modèles	Blé			Orge		
	R <sup>2</sup>	RMSE	MSE	R <sup>2</sup>	RMSE	MSE
Régression linéaire	-0.018	0.500	0.25	-0.997	0.703	0.494
KNN	0.568	0.326	0.106	0.096	0.473	0.223
Random Forest	0.535	0.338	0.114	0.317	0.411	0.169
Tree	0.532	0.339	0.115	<b>0.390</b>	<b>0.388</b>	<b>0.151</b>
SVM	0.546	0.334	0.112	0.178	0.451	0.203
Gradient Boosting	0.495	0.352	0.124	0.177	0.451	0.203
Neural Network	0.609	0.310	0.096	0.331	0.407	0.165
AdaBoost	0.560	0.329	0.108	-0.079	0.516	0.267
Stochastic Gradient Descent	<b>0.783</b>	<b>0.231</b>	<b>0.053</b>	0.363	0.397	0.157

### ➤ Extraction des meilleurs variables pour la prévision du rendement du blé

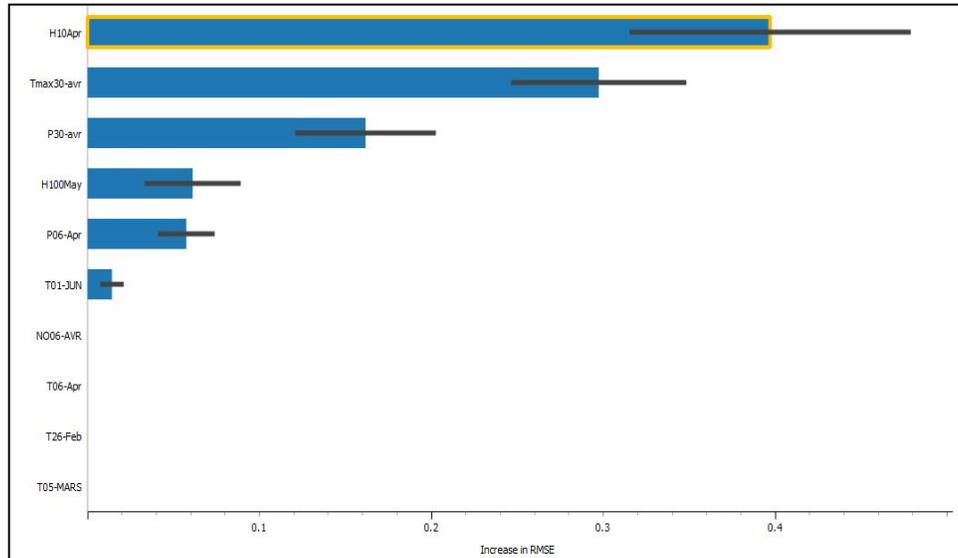
D'après la figure (84), on observe que les températures moyennes du 30 Avril et du 13 Mars, les températures maximales du 30 Avril et du 16 Mai, ainsi que le NDVI du 18 et 22 Février, du 06 et 22 Avril, et du 16 et 24 Mai, ainsi que l'humidité à 10 cm du mois d'Avril, constituent les dix meilleures variables combinées pour prédire le rendement du blé au niveau de la wilaya de Sétif.



**Figure (84) :** Les dix meilleurs variables sélectionnés pour prédire le rendement du blé au niveau de la wilaya de Sétif

### ➤ Extraction des meilleurs variables pour la prévision du rendement de l'orge

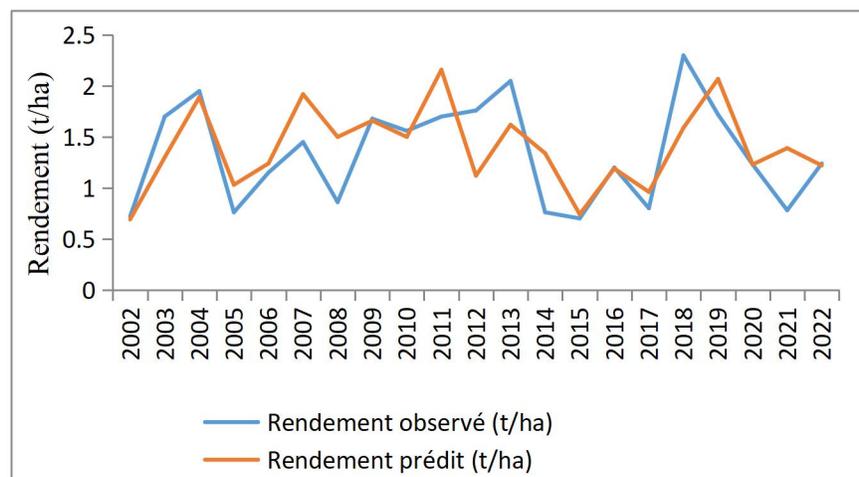
Les variables les plus significatives pour prédire précisément le rendement d'orge au niveau de la wilaya de Sétif sont: les précipitations du 06 et 30 Avril, le NDVI du 06 Avril, les températures moyennes du 26 Février, 05 Mars, 06 Avril et du 01 Juin, les températures maximales du 30 Avril et du 05 Mars, ainsi que l'humidité à 10 cm du mois d'Avril et à 100 cm du mois de Mai (figure 85).



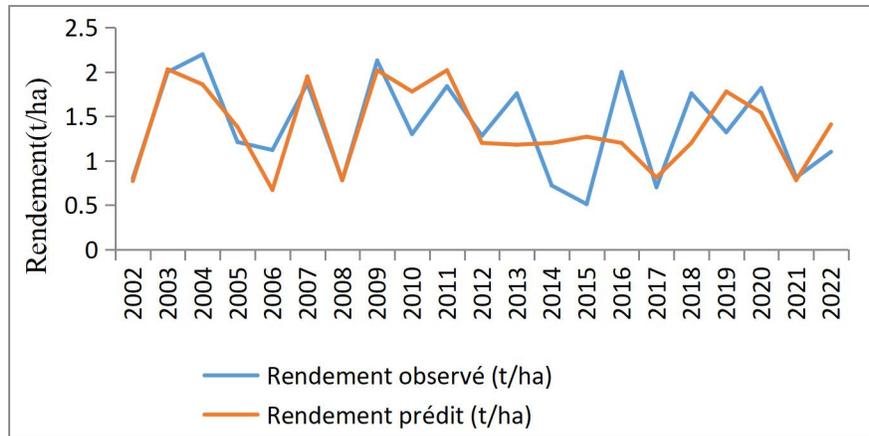
**Figure (85) :** Les dix meilleurs variable sélectionnés pour prédire le rendement d’orgeau niveau de la wilaya de Sétif

### 3.3.3. Vérification de la performance des modèles

Les résultats de la prédiction des rendements obtenus par le modèle "Stochastic Gradient Descent" du blé et "Tree" d’orge montrent une forte corrélation entre les rendements observés et prédits, avec des coefficients de corrélation (R) de 0.7 pour le blé et 0.68 pour l'orge. Les valeurs de RMSE, variant de 0.388 à 0.231 t/ha, suggèrent que les prédictions du modèle sont proche des valeurs réelles. De même, les valeurs de MSE, qui sont faible allant de 0.053 à 0.151, indiquent que les prédictions du modèle sont précises (figures 86, 87).



**Figure (86) :** Comparaison entre les rendements observés et prédits du blé par le modèle "Stochastic Gradient Descent" au niveau de la wilaya de Sétif



**Figure (87) :** Comparaison entre les rendements observés et prédits d'orge par le modèle "Tree" au niveau de la wilaya de Sétif

# **Discussion générale**

### **Discussion générale**

Prédire avec précision les rendements des cultures est crucial pour une planification agricole efficace, assurant ainsi la sécurité et la disponibilité des aliments. La télédétection par satellite est largement utilisée pour prévoir le rendement des céréales, en raison de sa capacité à être appliquée à l'échelle mondiale.

Les résultats de cette étude, démontre que l'utilisation du MODIS-NDVI seul (données entre 2002 à 2022) pour prédire le rendement du blé et d'orge au début du printemps (13 Mars), qui coïncide avec le stade "début tallage", est une période optimale pour effectuer des estimations fiables des rendements au niveau parcellaire, environs deux à trois mois avant la récolte. Tandis qu'au niveau communal et au niveau wilaya, les dates optimales pour prédire les rendements des deux cultures sont le 30 Avril pour le blé et le 8 Mai pour l'orge, correspondant aux stades "épiaison et remplissage des grains" environs deux mois avant la récolte, les valeurs du coefficient de corrélation varient de 0.7 à 0.86, alors que les coefficients de détermination se situent entre 0.65 et 0.74.

Mkhabela et *al.* (2010) ont prouvé que MODIS-NDVI pouvait prédire efficacement les rendements des cultures dans les prairies canadiennes, environ un à deux mois avant la récolte. Leurs résultats ont démontré une corrélation significative entre MODIS-NDVI et le rendement en grains pour toutes les cultures et zones agro-climatiques étudiées, avec des coefficients de détermination ( $R^2$ ) allant de 0.48 à 0.90 pour l'orge et de 0.47 à 0.80 pour le blé. Il est intéressant de noter que la force de la relation était similaire, voire plus forte, par rapport aux résultats obtenus dans notre étude. Dans une étude menée par Panek et Gozdowski (2021), une analyse de régression linéaire a été utilisée, pour étudier la corrélation entre l'indice de végétation par différence normalisé (NDVI) obtenu à partir des données satellite MODIS et le rendement en grains de blé et d'orge dans 20 pays européens entre 2010 et 2018. Ils ont trouvé une forte corrélation entre le NDVI et le rendement des céréales au début du printemps pour plusieurs pays, qui est similaire aux résultats de notre étude. Cependant, la force de cette relation était légèrement inférieure à celle observée dans notre étude, avec un  $R^2$  de 0.610 pour le blé et de 0.614 pour l'orge. Les résultats de la régression ont montré qu'une augmentation de 0.1 unité du NDVI est liée à une augmentation de 1.35 à 1.65 t/ha du rendement en grains des céréales. Wang et *al.* (2019) ont utilisé un modèle d'approche Carnegie-Ames-Stanford (CASA) amélioré, combiné à des séries chronologiques d'images de télédétection par satellite obtenues à partir de MODIS,

pour estimer le rendement du blé d'hiver dans certaines régions de Chine. L'étude a rapporté un coefficient de détermination de  $R^2 = 0.56$  entre le rendement estimé et mesuré du blé d'hiver, inférieur à celui trouvé dans cette recherche, une erreur quadratique moyenne (RMSE) de 1.22 t/ha, qui est supérieur à celui trouvé dans ce travail. Aussi, Gop et Savenkov (2016) ont trouvé une corrélation significative entre le rendement du blé de printemps et le NDVI ( $R^2 = 0.859$ ). L'étude a démontré que le NDVI était responsable de 85 % de la variation du rendement du blé de printemps. En moyenne, une augmentation de 0.1 unité du NDVI était associée à une augmentation d'environ 6.7 t/ha des rendements en grains du blé de printemps. Dans une autre étude menée par Tuğaç et *al.* (2022), la corrélation la plus élevée entre le NDVI et le rendement a été observée pendant la période de floraison ( $R^2 = 0.63$ ). Ils ont également constaté que le modèle MLP pour MODIS offrait les meilleures performances de prédiction, avec une erreur quadratique moyenne (RMSE) variant de 0.23 à 0.65 t/ha. Selon Meshaba et *al.* (2017), la corrélation entre le NDVI et le rendement du blé était fortement positive avec une valeur  $R^2$  de 0.73 et un RMSE de 0.41 t/ha. Pismennaya et *al.* (2021) ont également examiné la corrélation entre le MODIS-NDVI et le rendement du blé d'hiver dans les zones arides de la région centrale du Pré-Caucase, en utilisant des données de 2017 à 2020. Leurs résultats ont montré une corrélation très forte et positive ( $R^2 = 0.78$ ) entre le rendement du blé d'hiver et le NDVI. De plus, ils ont observé une augmentation moyenne de 0.20 t/ha du rendement en grains de blé pour chaque augmentation de 0.1 de la valeur NDVI. En Europe centrale, Panek et Gozdowski (2020) ont trouvé une forte relation entre le rendement en céréales et MODIS-NDVI au printemps (Avril), trois à quatre mois avant la récolte. L'augmentation du NDVI au début du printemps de 0.1 unité augmente le rendement en grains des céréales d'environ 1.1 à 2.6 t/ha.

L'utilisation combinée du MODIS-NDVI et des données de précipitations (régression linéaire multiple) pour prédire les rendements à trois échelles, a révélé les mêmes périodes optimales d'estimation des rendements, avec une légère amélioration de la corrélation (R) par rapport à l'utilisation du NDVI seul, les valeurs de R varient de 0.7 à 0.86, tandis que les  $R^2$  se situent entre 0.64 et 0.82.

Vannoppen et *al.* (2020), ont étudié la combinaison des NDVI avec des variables météorologiques pour évaluer leur corrélation avec le rendement du blé en Lettonie (Latvia). Les résultats ont montré qu'un modèle de rendement par régression linéaire expliquait 71 % de la variabilité avec une erreur résiduelle de 0.55 t/ha. Lorsque les données météorologiques ont été

intégrées comme variables prédictives au modèle empirique de rendement du blé, une approche aléatoire de forêt 'Random forest' a donné de meilleurs résultats par rapport à une approche de régression linéaire. La variance expliquée a augmenté jusqu'à 97 % et l'erreur résiduelle a diminué à 0.17 t/ha, ce qui est similaire aux nos résultats. Aussi, Balaghi *et al.* (2006), ont développé des modèles empiriques de régression pour prévoir les rendements de blé au niveau provincial et national au Maroc. En utilisant le NDVI décadaire (10 jours), les cumuls de précipitations et les températures mensuelles moyennes. Les rendements provinciaux en blé ont été évalués avec des erreurs variant de 80 à 762 kg ha<sup>-1</sup>, selon les provinces. Au niveau national, le rendement du blé a été prévu à la troisième décade d'avril avec une erreur de 73 kg ha<sup>-1</sup>, en utilisant le NDVI et la pluviométrie. Cependant, des prévisions plus précoces sont possibles, à partir de la deuxième décade de Mars avec une erreur de 84 kg ha<sup>-1</sup>, au moins 1 mois avant la récolte. Au niveau provincial et national, la majeure partie de la variation du rendement était due au NDVI. Dans l'étude menée par Sun *et al.* (2022), une approche basée sur l'apprentissage automatique (Random Forest, RF) a été employée pour estimer le rendement du blé d'hiver en Chine en intégrant des données satellitaires, des données climatiques. Les résultats indiquent que la méthode RF présente une meilleure précision dans l'estimation du rendement par rapport à la régression linéaire multiple. L'intégration de données climatiques et géographiques permet d'améliorer significativement la précision de l'estimation du rendement, avec un (R<sup>2</sup>) de 0.87 et une erreur quadratique moyenne (RMSE) de 488.59 kg/ha. Ce modèle permet d'obtenir une meilleure prévision des rendements à grande échelle deux mois à l'avance.

On a démontré dans cette étude aussi, que l'application mobile Canopeo®, utilisée pour mesurer la couverture fractionnée du couvert vert (FGCC), constitue une alternative simple et rentable par rapport au NDVI et à la méthode traditionnelle de mesure de la biomasse verte. Ces résultats sont cohérents avec plusieurs autres études. Mcglinch *et al.* (2021). ont étudié la corrélation entre le nombre de tiges, la couverture fractionnée du couvert vert (FGCC), l'indice de végétation par différence normalisée (NDVI) et le rendement en grains d'orge. Les résultats ont indiqué que les mesures du nombre de tiges et du FGCC présentaient la plus grande variabilité du rendement en grains d'orge, avec des valeurs R<sup>2</sup> de 0.67 à 0.65, respectivement, au stade plein tallage. À l'inverse, le NDVI a montré une influence moins marquée sur la variabilité du rendement, avec un R<sup>2</sup> de 0.50. Reed *et al.*(2021), ont souligné que les mesures FGCC et NDVI sont fortement corrélées au rendement de blé et fournissaient des estimations similaires avec une

valeur  $R^2$  de 0.76. Dans une étude menée par Goodwin *et al.* (2018), des modèles de régression ont été développés pour prédire le rendement en grains de blé d'hiver à différents stades de croissance dans l'Ohio. Les résultats ont indiqué que, au cours du stade « plein tallage », l'indice de végétation par différence normalisée (NDVI) et la couverture fractionnée du couvert vert (FGCC) avaient le plus influence significative sur la variabilité du rendement, présentant des valeurs  $R^2$  respectives de 0.49 et 0.45. En 2021, Schmitz et Kandel, ont trouvé que les valeurs du NDVI prise au stade R5 (début des stades de croissance du développement des graines), est le prédicteur le plus précis du rendement du soja, avec un ( $R^2$ ) de 0.65. Suivi du FGCC avec un  $R^2 = 0.52$ .

Dans d'autres études, bien que la couverture fractionnée du couvert vert (FGCC) explique une partie légèrement inférieure de la variabilité du rendement du blé par rapport au NDVI, elle s'avère être une approche plus économique et plus rentable, étant donné que Canopeo® utilisée pour mesurer le FGCC est actuellement disponible en téléchargement gratuit (Patrignani et Ochsner, 2015). Il est toutefois essentiel de noter que pour obtenir des résultats précis, les mesures FGCC doivent être effectuées dans un environnement exempt de mauvaises herbes. Cette précaution est nécessaire car l'application mobile Canopeo® intègre la biomasse des mauvaises herbes vertes dans la mesure FGCC.

Les résultats des modèles d'apprentissage (machine learning) montrent que différents algorithmes d'apprentissage automatique peuvent exceller à différentes échelles et pour différentes cultures. Il n'existe pas un modèle unique universellement supérieur, et le choix du modèle optimal dépend du contexte spécifique et des objectifs de l'étude.

A l'échelle parcellaire, le modèle "KNN" (k-Nearest Neighbors) s'est distingué en démontrant la meilleure performance pour la prédiction du rendement du blé, expliquant environ 38.6% de la variation des rendements avec un RMSE de 0.456 t/ha. Cette observation suggère l'efficacité du KNN dans l'identification des relations locales entre les variables prédictives et les rendements à cette échelle. Quant à l'orge, le modèle "Tree" semble être le plus précis et adapté pour prédire les rendements d'orge au niveau parcellaire, avec un  $R^2$  de 0.672 et un RMSE de 0.289 t/ha, mettant en avant sa capacité à combiner plusieurs modèles faibles pour mieux saisir les complexités des facteurs influençant le rendement de l'orge.

A l'échelle communale, le modèle "Stochastic Gradient Descent" s'est démarqué dans la prédiction du rendement du blé, avec un  $R^2$  de 0.651, un RMSE de 0.368 t/ha et un MSE de 0.136.

Son approche itérative d'optimisation des paramètres du modèle semble avoir mieux appréhendé les relations entre les variables prédictives et les rendements à cette échelle. De même, le modèle "Gradient Boosting" a affiché les meilleures performances pour la prédiction du rendement de l'orge, avec un  $R^2$  de 0.358 et un RMSE de 0.612 t/ha, confirmant sa robustesse à cette échelle.

À l'échelle de la wilaya, le modèle "Stochastic Gradient Descent" a montré une forte performance dans la prédiction du rendement du blé, avec un  $R^2$  de 0.783, un RMSE de 0.231 et un MSE de 0.053, soulignant sa capacité à fournir des prédictions précises à des échelles spatiales plus larges. Quant à l'orge, le modèle "Tree" s'est avéré le plus efficace, avec un  $R^2$  de 0.390, un RMSE de 0.388 et un MSE de 0.151, bénéficiant de sa structure arborescente pour capturer les interactions complexes entre les variables prédictives et le rendement de l'orge.

Les modèles d'apprentissage automatique, malgré leur potentiel théorique pour capturer des relations complexes, ont montré des performances variables et globalement inférieures aux méthodes basées sur le NDVI, le FGCC et la régression linéaire pour la prédiction du rendement des céréales dans cette étude.

Plusieurs facteurs pourraient expliquer cette sous-performance. Premièrement, la résolution spatiale moyenne des données météorologiques (tmin, tmax, tmoy, precipitation) qui est de  $0.1^\circ$  (10km) et par la suite qui ont influé sur le calcul de l'évapotranspiration et les indices de sécheresse (aSPI et eRDI), de même l'humidité du sol à une résolution spatiale estimée à  $0.96^\circ$  (9.6km) pourrait ne pas être suffisamment fine pour capturer la variabilité locale des conditions environnementales qui influencent le rendement des cultures à l'échelle parcellaire.

Deuxièmement, le grand nombre de variables d'entrée (87) utilisé dans certains modèles d'apprentissage automatique pourrait avoir conduit à un sur-apprentissage, limitant ainsi la capacité des modèles à généraliser à de nouvelles données. Troisièmement, la période d'étude relativement courte (2002-2022) pourrait ne pas fournir suffisamment de données pour entraîner efficacement les modèles d'apprentissage automatique, qui nécessitent généralement de grands ensembles de données pour atteindre des performances optimales.

Bien que tous les méthodes de prédiction utilisées dans ce travail (FGCC, NDVI, biomasse des cultures, les facteurs climatiques et agro-climatiques) aient démontré leurs efficacités dans l'évaluation des rendements en grains des céréales (blé et orge), il est essentiel de prendre en compte d'autres facteurs pouvant influencer le rendement final des céréales, après la date de prévision. Les facteurs environnementaux, telles que, les maladies ou les infestations de

ravageurs peuvent également avoir un impact sur le rendement final. De plus, les images satellitaires sont sensibles à divers effets atmosphériques, notamment les nuages, qui peuvent compromettre la qualité des données et affecter par la suite les modèles de rendement des cultures développés.

# **Conclusion générale et Perspectives**

### **Conclusion générale et Perspectives**

La prévision des rendements des cultures est une tâche critique et complexe dans l'agriculture moderne en raison de divers défis. Ces défis incluent les impacts du changement climatique mondial, y compris les événements météorologiques extrêmes tels que la sécheresse, les inondations et autres catastrophes naturelles, ainsi que l'augmentation de la population mondiale et de la demande alimentaire.

Les résultats de trois ans d'expérimentation mettent en évidence l'efficacité de la couverture fractionnée du couvert vert (FGCC), mesurée à l'aide de l'application Canopeo®, en tant que prédicteur fiable du rendement en grains des céréales deux à trois mois avant la récolte. Comparativement aux données NDVI et à la méthode traditionnelle de mesure de la biomasse sèche, le FGCC a démontré une corrélation plus forte avec les rendements observés. Cette conclusion suggère que le FGCC, en tant qu'outil de mesure simple et rentable, peut être utilisé efficacement pour évaluer et prédire les rendements des cultures céréalières. En plus Canopeo® s'avère facile et accessible avec seulement un téléphone intelligent requis. Tirant parti de l'analyse d'images numériques, il offre une méthode précise, non destructive et efficace pour prévoir le rendement en grains de céréales tout en facilitant l'enregistrement, le traitement et l'archivage d'un grand volume d'images. Cette application pourrait être utilisée par les agriculteurs et les conseillers agricoles pour estimer le rendement des céréales avant la récolte. Aussi permettre aux agriculteurs de prendre des décisions agronomiques, gestion éclairées, de déterminer la faisabilité des applications d'azote (N) et de fongicides liées à la rentabilité potentielle de la culture. Pour l'avenir, des perspectives intéressantes incluent l'exploration de l'application du FGCC dans d'autres environnements agricoles et cultures, ainsi que l'intégration de données supplémentaires pour affiner les modèles de prédiction. Il serait également bénéfique d'étudier les variations saisonnières du FGCC et son impact sur la fiabilité des prévisions de rendement. En développant davantage notre compréhension et notre utilisation du FGCC, nous pourrions améliorer la gestion agricole et la planification des cultures, contribuant ainsi à une agriculture plus efficace et durable.

Cependant, des recherches plus approfondies sont essentielles pour valider l'exactitude de cette approche dans divers contextes. Une limitation notable de cette application est qu'elle ne peut pas faire la différence entre les cultures et les mauvaises herbes, elle nécessite des mises à jour, comme les algorithmes d'apprentissage automatique, qui peuvent analyser les modèles

visuels et distinguer les caractéristiques des cultures et des mauvaises herbes. Entraînez l'algorithme à l'aide d'un ensemble de données diversifié pour améliorer la précision.

On a démontré aussi l'efficacité de l'utilisation du MODIS-NDVI seul ou en combinaison avec les données des précipitations entre 2002 et 2022, pour prédire le rendement des cultures céréalières (blé et orge) à trois échelles d'étude (parcelle, commune et wilaya). Ces méthodes ont permis d'obtenir des prévisions fiables deux à trois mois avant la date de récolte. L'intégration de variables météorologiques en plus du NDVI semble améliorer la capacité à capturer mieux des bonnes relations avec le rendement des céréales, aboutissant à des prédictions plus fiables.

Malgré ces limitations à cause de la résolution spatiale faible à moyenne, le sur-apprentissage et la courte durée d'entraînement des données (vingt ans), certains modèles de machine learning ont montré des résultats prometteurs, notamment "KNN" au niveau parcellaire et "Stochastic Gradient Descent" aux niveaux communal et wilaya pour le blé. Tandis que pour l'orge, le modèle "Tree" a montré des bonnes performances au niveau parcellaire et wilaya. Cependant, Le modèle "Gradient Boosting" a enregistré des bonnes performances au niveau communale.

Les résultats ont montré que la sélection des variables d'entrée joue un rôle crucial dans la performance des modèles d'apprentissage automatique. L'utilisation des dix meilleures variables, identifiées par la méthode de "gain d'information", a généralement conduit à de meilleures performances par rapport à l'utilisation de toutes les variables disponibles. Cela suggère que, avec des améliorations en termes de résolution des données, de sélection des variables et de taille de l'échantillon, les modèles d'apprentissage automatique pourraient jouer un rôle plus important dans la prédiction du rendement des cultures à l'avenir. Des recherches supplémentaires sont nécessaires pour explorer et optimiser ces modèles afin de maximiser leur potentiel pour la gestion agricole et la sécurité alimentaire.

De plus, il est crucial de tenir compte des conditions météorologiques extrêmes et d'autres facteurs environnementaux qui pourraient influencer les rendements des cultures, afin de renforcer la robustesse des modèles de prédiction.

Pour l'avenir, plusieurs perspectives intéressantes d'utilisation du NDVI dans la prédiction des rendements des céréales, sera possible d'améliorer la robustesse et l'efficacité des modèles de prédiction, ce qui contribuera à une gestion agricole plus précise et durable tels que :

## *Conclusion générale et perspectives*

---

- ✓ Intégration de données multi-sources: Combiner le NDVI avec d'autres sources de données telles que des images satellitaires supplémentaires, des données topographiques et des informations sur le sol pour une analyse complète des facteurs influençant les rendements des céréales.
- ✓ Utilisation des données NDVI à haute résolution spatiale pour une analyse plus précise.

# **Références bibliographiques**

## Références

- Abbassenne, F. (1997). Etude génétique de la durée des phases de développement et leur influence sur le rendement et ses composantes chez le blé dur (*Triticum durum* Desf.). Thèse de Magister INA, 81p.
- Adeniyi, O. D., Szabo, A., Tamás, J., et Nagy, A. (2020). Wheat Yield Forecasting Based on Landsat NDVI and SAVI Time Series. *Preprints*. <https://doi.org/10.20944/preprints202007.0065.v1>.
- agricole 2<sup>ème</sup> Ed, Ballier France. 351p.
- Aidani, H. (2015). Effet des attaques de Capucin des grains sur les céréales stockées Estimation sur la perte pondérale et le pouvoir germinatif cas de blé dur dans la région de Tlemcen. 82.
- Algérie Eco. (2022). <https://www.algerie-eco.com/2022/07/13/cereales-lalgerie-a-importe-106-millions-de-tonnes-durant-la-campagne-2021-2021/>.
- Alsalam, B. H. Y., Morton, K., Campbell, D., et Gonzalez, F. (2017). Autonomous with vision based on-board decision making for remote sensing and precision agriculture. In *Aerospace Conference*, pages 1–12.
- Ammar, M. (2014). Organisation de la chaîne logistique dans la filière céréales en Algérie : état des lieux et perspectives. *Master of Sciences*, CIHEAMIAMM, 127 pp.
- Araus, J. L., Villegas, D., Aparicio, N., Garcia del Moral, L. F., El Hani, S., Rharrabti, Y., Ferrio, J. P. et Royo, C. 2003. Environmental factors determining carbon isotope discrimination and yield in durum wheat under Mediterranean conditions. *Crop Sci*, 43:170-180.
- Assoko, T. et Joel. (2022). L'Algérie veut compter sur son stock de sécurité.
- Atzberger, C., Vuolo, F., Klisch, A., Rembold, F., Meroni, M., Marcio, P. M., et Formaggio, A. (2016). Agriculture. In P. S. Thenkabail (Eds.), *Remote Sensing Handbook*, 71-103.
- Boudjema, M. A. (2017). Parc national de Tlemcen. Dynamique de la couverture végétale et perspective. Mémoire de master, univ. Abou Bakr Belkaid – Tlemcen, 50,51p.
- Baghem, O. (2012). Effet des Techniques Culturelles sur la Biodiversité Faunistique des céréales dans la zone Semi-aride. Thèse de Magister, Université de Sétif, 75 p.

Baio, F. H. R., Silva, E. E., Souza, M. A. V., Zanin A. R. A., Teodoro, P. E. (2018). Vegetation indices to estimate spray application rates of crop protection products in corn. *Agronomy Journal*, 110: 1-15. <https://doi.org/10.2134/agronj2017.12.0718>.

Balaghi, R., Bernard, B., Herman, E., Mohammed, J. (2006). Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco, *International Journal of Applied Earth Observation and Geoinformation*, <https://doi.org/10.1016/j.jag.2006.12.001>.

Baldy, C. (1984). Utilisation efficace de l'eau par la végétation en climats méditerranéens. Bull. Soc. Bot. Fr 131(2,3,4). *Actual. Botan* (491-499).

Baldy, C. (1986). Comportement des blés dans les climats méditerranéens. *Ecologia Mediterranea*, 12: 73-88.

Bannari, A., Morin, D., Bonn, F., Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, Vol. 13, p. 95-120.

Baret, F., et Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, 35: 161-173.

Basso, B., Cammarano, D., Carfagna, E. (2013). Review of Crop Yield Forecasting Methods and Early Warning Systems. First Meet. Sci. Advis. Comm. Glob. Strateg. to Improv. *Agric. Rural Stat*.

Basso, B., Liu, L. (2019). Seasonal crop yield forecast: Methods, applications, and accuracies. *Advances in Agronomy*. <https://doi.org/10.1016/bs.agron.2018.11.002>.

Battinger, R. (2002). Les grandes cultures. Conseil Européen des Jeunes Agriculteurs. 15p.

Becker-Reshef, I., Justice, C., Barker, B., Humber, M., Rembold, F., Bonifacio, R., Zappacosta, M., Budde, M., Magadzire, T., Shitot, C., Pound, J., Constantino, A., Nakalembe, C., Mwangi, K., Sobue, S., Newby, T., Whitcraft, A., Jarvis, I., et Verdin, J. (2020). Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM Crop Monitor for Early Warning. *Remote Sens. Environ.* 237, 111553. <https://doi.org/10.1016/j.rse.2019.111553>.

Bégué, A., Arvor, D., Bellon, B., Betbeder, J., de Aballeyra, D., Ferraz, R. P. D., Verón, S. (2018). Remote Sensing and Cropping Practices: A Review. *Remote Sensing*, 10(1), 99. <http://doi.org/10.3390/rs10010099>.

Belagrouz, A. (2013). Analyse du Comportement du Blé Tendre, Variété El WIFAK (*Triticum aestivum* L.) Conduite en Labour Conventionnel, Travail Minimum et Semis Direct sur Les Hautes Plaines Sétifiennes. Mémoire Magistère, université Ferhat Abbas Setif. 88p.

Bencharif, A., Rastoin, J.L. (2007). Concepts et méthodes de l'analyse de filières agroalimentaires : application par la chaîne globale de valeur au cas des blés en Algérie » Montpellier

Benhammed, A., Djekoune, A., Bouzerzour, H et Hassoun, K.L. (2005). Genotype x year interaction of barley grain yield its relationship with plant height earliness and climate factors under semi-arid growth conditions. *Dirasat, Agricultural Science*, 32: 239-247. blé : 256 p.

Bennour, A. (2015). Contribution au recalage des images satellitaires de haute résolution pour le suivi à long terme des changements écologiques, université Badji Mokhtar Annaba. Thèse de Doctorat, pp 8, 9.

Bonn, F. et Rochon, G. Précis de télédétection Volume 1 : Principes et Méthodes, Presses de l'Université du Québec/AUPELF, 1992

Bourahla, S., Hadji, A. (2009) : L'effet de stress hydrique sur la teneur de chlorophylle de blé dur (*Triticum durum* Desf.). Mémoire de DES en biologie. Msila. PP : 4-5.

Bourihane, D., Mekkaoui, Z. 2013. Analyse des déterminants de la production du blé en Algérie Cas des wilayas Tiaret, Sétif et Médéa. L'échantillon 1990 – 2009. Mémoire de master. Option : Economie Appliquée et Ingénierie Financière. Université Abderrahmane mira de Bejaia : 86 p.

Büchi, L., Wendling, M., Mouly, P., et Charles, R. (2018). Comparison of visual assessment and digital image analysis for canopy cover estimation. *Agronomy Journal*, 110: 1289–1295. <https://doi.org/10.2134/agronj2017.11.0679>.

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., et al. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agric. For. Meteorol*, 274:144–459. <https://www.sciencedirect.com/science/article/pii/S0168192319301224>.

Caloz, R. (1992). Télédétection satellitaire, Cours Ecole Polytechnique Fédérale de Lausanne, Département Génie Rural, Lausanne, 136 p.

Caloz, R., Blaser, T.J et Willemin, G. (1993). « Création d'une orthoimage à l'aide d'un - modèle numérique d'altitude : influences des modes de rééchantillonnage radiométrique », Télédétection et cartographie. Ed. AUPELF-UREF. Les presses de l'université du Québec., p. 17-30.

Campbell, J. B. (1987) Introduction to Remote Sensing. The Guilford Press, New York, USA, 551.

Casadesús, J., Kaya, J., Bort, J., Nachit, M. M., Araus, J. L., Samor, G., Ferrazzano, F., Maalouf, M., Maccaferri, V., Martos, H., Ouabbou, et Villegas, D. (2007). Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in

water limited environments. *Ann. Appl. Biol*, 150:227–236. <https://doi:10.1111/j.1744-7348.2007.00116.x>.

Caturegli, L., Corniglia, M., Gaetani, M., Grossi, N., Magni, S., Migliazzi, M., Volterrani, M. (2016). Unmanned Aerial Vehicle to Estimate Nitrogen Status of Turfgrasses. *PLoS ONE*, 11(6). <https://doi.org/10.1371/journal.pone.0158268>.

Chennafi, H., Aïdaoui, A., Bouzerzour, H. & Saci, A. (2006). Yield responses of durum wheat (*Triticum durum* Desf.) cultivar Waha to deficit irrigation under semi arid growth conditions. *Asian J. Plant Sci*, 5: 854-860.

Chevallier, R. (1971). La photographie aérienne. Armand Colin, Paris, 233 p.

Chlingaryan, A., Sukkarieh, S., Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* <https://doi.org/10.1016/j.compag.2018.05.012>.

Clement-Grandcourt, M. et Prats, J. (1971). Les céréales Collections d'enseignement

Clerget, Y., 2011. La biodiversité des céréales et leur utilisation par l'homme. Bulletin de la Société d'Histoire Naturelle du Pays de Montbéliard. 16 p.

Commission Européenne, 2010. Rapport de la Commission au Conseil et au Parlement Européen concernant la mise en œuvre des applications de télédétection ainsi que l'utilisation des ressources financières mises à sa disposition au titre du règlement (CE) n° 78/2008 du Conseil (rapport intermédiaire). 12 p. <https://doi.org/10.1016/j.biosystemseng.2004.09.003>.

Dali, S., Jouadi, H., Louinis, M. et Nekkab, D., et al., 2010. Evaluation des contrats de performance des wilayas. *Céréaliculture : revue technique et scientifique de l'Institut Technique des Grandes Cultures*, 55 : 192.

De Rango, F., Palmieri, N., Santamaria, A.F., and Potrino, G. (2017). Asimulator for uavs management in agriculture domain. In *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, pages 1–8.

Debaeke, P., Purch, J. et Casal, M. L. (1996). Elaboration du rendement de blé d'hiver en condition de déficithydrigue. Etude en Lysimètre. *Agronomie*, 16 : 3-23

Despinasse, Y. (2015). Diversité chimique et caractérisation de l'impact du stress hydrique chez les lavandes. Thèse de doctorat. Université Jean Monnet-Saint-Etienne. 167p.

DiMiceli, C., Carroll, M., Sohlberg, R., et Townshend, J. (2020). Annual global automated MODIS vegetation continuous fields (MOD44B) at 250m spatial resolution for data years beginning day 65, 2000–2017, Collection 6 percent tree cover. University of Maryland. Data set.

Djermoun, A. (2009). La production céréalière en Algérie: les principales caractéristiques. *Nature and Technology*, 1- 45.

DSASI, (2020). Statistiques de Série B : Rapport annuel. Direction de Statistiques Agricoles et des Systèmes d'Information. Ministère de l'Agriculture et du Développement Rural, MADR, Algérie.

Dugué, M.J. (2012). Caractérisation des stratégies d'adaptation au changement climatique en agriculture paysanne. Retrieved from : <https://www.avsf.org>.

Eckert, S., Hüsler, F., Liniger, H., Hodel, E. (2015). Trend analysis of MODIS NDVI time series for detecting land degradation and regeneration in Mongolia. *J. Arid Environ*, 113: 16-28. <https://doi.org/10.1016/j.jaridenv.2014.09.001>.

Ed-Daoudi, R., Alaoui, A., Ettaki, B., Zerouaoui, J. (2023). Improving Crop Yield Predictions in Morocco Using Machine Learning Algorithms. *J. Ecol. Eng*, 24: 392–400. <https://doi.org/10.12911/22998993/162769>.

Fang, H., Liang, S., Hoogenboom, G., Teasdale, J., & Cavigelli, M. (2008). Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model. *Int. J. Remote Sens.* 2 : 3011–3032.

FAO, (2020). Statistiques sur les productions ; ressources STAT en ligne. [www.fao.org](http://www.fao.org)

FAOstat, (2023). Food and Agriculture Organization of United Nations, Rome, <http://faostat3.fao.org.org>.

Feillet, P. (2000). Le grain de blé, composition, utilisation. Ed, INRA, paris, 308p.

Fellahi, Z. (2017). Analyse génétique d'un croisement line x testeur, réponse à la sélection et tolérance des stress du blé tendre (*Triticum aestivum* L.) sous condition semi-aride, thèse de doctorat en science, Univ.Ferhat Abbas Sétif 1, 227p.

Fellahi, Z., Hannachi, A., et Bouzerzour, H. (2019). Expected genetic gains from mono trait and indexes based selection in advanced bread wheat (*Triticum aestivum* L.) *population*. *Revista Facultad Nacional de Agronomía*. <https://doi.org/10.15446/rfnam.v73n2.77806>.  
Forecast, <https://www.fao.org/worldfoodsituation/csdb/en/>.

Foudili, D. et Gasmi, A. (2017). Stress de la sécheresse chez quatre variétés du Blé dur (*Triticum durum* Desf.): Un examen sur quelques caractéristiques morphologiques et sur les pigments. Mémoire de Master, Univ MohamedBoudiaf - M'sila, 52 p.

Gadal, S. (2003). La question des liens multi niveaux en analyse de données spatiales : Exemple de Casablanca, Séminaire GDR 2340 SIGMA « Multi-Echelles », Université d'AVIGNON, Centre d'analyse et de mathématique sociale, Ecole des hautes études en sciences sociales de Paris.

Gate, P. (1995). Ecophysiologie du blé. Ed. ITCF. Technique et Documentation. Lavoisier, Paris. P: 419.

Gates, D. M., and L. D. Miller. (1977). Contribution of the soil spectra to grass canopy spectral reflectance. *Photogramm. Eng. Remote Sens*, 43: 721–726.

Gausman, H. W. (1985). Plant leaf optical properties in visible and near infrared light, Graduate Studies, Texas Tech University (No. 29). Lubbock, Texas: Texas Tech Press, pp 78-88.

Geng, La., Mengdi, Li., Guoping, et Zhang G.Y. (2022). Barley: a potential cereal for producing healthy and functional foods. *Food quality and safety*, 6:1-13, <https://doi.org/10.1093/fqsafe/fyac012>.

Giglio., Louis., Csiszar., Ivan., et Justice., Christopher,O. (2006). Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. *Journal of geophysical research : Biogeosciences*, 111(G2).

GIMMS|Global Agricultural Monitoring System. Available online: <https://glam1.gsfc.nasa.gov/> (accessed on 12 January 2021).

Girard, M. C., et Girard, C. (2010). Traitement des données de télédétection . *Environnement et ressources naturelles* (2nd ed.).

Global Croplands. Available online: <https://croplands.org/app/map?lat=0&lng=0&zoom=2> (accessed on 12 January 2021).

Goodwin, A. W., Lindsey, L. E., Harrison, S. K and Paul P. A. (2018). Estimating wheat yield with normalized difference vegetation index and fractional green canopy cover. *Crop, Forage and Turfgrass Management*, 4: 1–6. <https://doi.org/10.2134/cftm2018.04.0026>.

Goodwin, A. W., Lindsey, L. E., Harrison, S. K and Paul, P. A. (2018). Estimating wheat yield with normalized difference vegetation index and fractional green canopy cover. *Crop, Forage and Turfgrass Management*, 4: 1–6. <https://doi.org/10.2134/cftm2018.04.0026>.

Google Earth Pro, 2023.

Gop N.V., et Savenkov O. A. (2019). Relationships between the NDVI, Yield of Spring Wheat, and Properties of the Plow Horizon of Eluviated Clay-Illuvial Chernozems and Dark Gray Soils. *Eurasian Soil Science*, 2019, Vol. 52, No. 3, pp. 339–347. <http://dx.doi.org/10.1134/S1064229319030050>.

Gurayerus, L. (2008). Reconnaissance d'objets cartographiques dans les images satellitaire à haute résolution, Thèse de Doctorat à l'université PARIS DESCARTES.

Hamadache, A. (2013). *Eléments de Phytotechnie Générale-Grandes Cultures-Le blé* : 256 p.

Hamimed, A. (2009). Suivi de l'état hydrique d'une région semi-aride dans l'ouest Algérien à partir des images de télédétection haute et faible résolution. Thèse de doctorat, en sciences hydrauliques, p179.

Hargreaves, G.H. Samani, Z.A. (1985). Reference crop evapotranspiration from temperature. *Trans ASAE* 1(2):96–99.

Hénonin, J., Adiku, S. G. K., Ametsitsi, G. K. A., et Forkuor, G. (2021). Land cover classification and change detection using Landsat 8 data in the Volta region of Ghana. *Remote Sensing*, 13(11), 2137.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1-2), 195-213.

Idrissi, A., Nadem, S., Boudhar, A., Benabdoulouhab, T. (2022). Review of wheat yield estimating methods in Morocco. *Afr. J. Land Policy Geospat. Sci*, 5: 818-831.

ITGC, 2016. Evaluation de l'encadrement phytosanitaire des céréales durant la période 2010-2015. Céréaliculture : *technique et scientifique de l'Institut Technique des Grandes Cultures*. Numéro 66. 35 p.

J. Canny, A. J. (1986). Computational approach to edge detection, *IEEE trans. on pattern analysis and machine intelligence*, 8: 679-698.

Johnson, D.M. (2014). An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sens. Environ.* <https://doi.org/10.1016/j.rse.2013.10.027>.

Johnston, A.E. (2002). Phosphorus Nutrition of Cereal Crops: A Sustainable Approach. *Journal of Agricultural Science*, 139(2): 129-143.

Junior, S C. A., Nanni, M. R., Shakir, M. et al. (2018). Soybean varieties discrimination using non-imaging hyperspectral sensor. *Infrared Physics and Technology*, 89: 338-350. <https://doi.org/10.1016/j.infrared.2018.01.027>.

Kalarasse, A. (2018). Effet des altérations de la semence sur le développement de la culture et la qualité du rendement chez le blé.

Kluser, S. (2000). Elaboration d'une carte de couverture du sol par interprétation semi. Thèse de doctorat. Univ Guelma

Kayad, A., Rodrigues, FA. Jr., Naranjo, S., Sozzi, M., Pirotti, F., Marinello, F., et al. (2022). Radiative transfer model inversion using high-resolution hyperspectral airborne imagery—Retrieving maize LAI to access biomass and grain yield. *Field Crops Res.* 282:108-449. <https://doi.org/10.1016/j.fcr.2022.108449> PMID: 35663617.

Kergomard, C. (1996). La télédétection aérospatiale : une introduction. *Paris: Masson*, Collection Sciences Sup.

Kolai, T. (2008). Climat et dysfonctionnement des agro-systèmes céréalier cas des wilayas Sétif, Bordj BouArréridj et Mila. Mémoire de Magister, Institut national agronomique EL Harrach, 68 p.

Kribaa, M., Hallaire, V., Curmi, P., Lahmar, R. (2001). Effects of various cultivation methods on the structure and hydraulic properties of soil in semi-arid climate. *Soil and Tillage Research*, 60: 43-53.

Kumari, N., Srivastava, A., Dumka, U.C. (2021). A Long-Term Spatiotemporal Analysis of Vegetation Greenness over the Himalayan Region Using Google Earth Engine. *Climate*, 9: 109. <https://doi.org/10.3390/cli9070109>.

L'Agence Nationale d'Intermédiation et de Régulation Foncière (ANIREF), (2023). Monographie wilaya de Sétif (on line). Disponible sur : <https://www.aniref.dz/documentspdf/monographies/MONOGRAPHIE%20WILAYA%20SETIF.pdf>. Consulté le 20.03.2024.

Li, J., et Liu, Y. (2020). A preliminary study on the correlation between MODIS data and maize yield. *Agricultural and Forest Meteorology*, 291: 108-085.

Li, Y., Qian, M., Liu, P., Cai, Q., Li, X., Guo, J., Yan, H., Yu, E., Yuan, K., Yu, J., et al. (2018). The recognition of rice images by uav based on capsule network. *Cluster Computing*, 1–10, 2018.

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors (Switzerland)*, 18: 1-29. <https://doi.org/10.3390/s18082674>.

Liu, H., Zhang, X., Xu, Y., Ma, F., Zhang, J., Cao, Y., et al. (2020). Identification and validation of quantitative trait loci for kernel traits in common wheat (*Triticum Aestivum* L.). *BMC PlantBiol*, 20:529. <https://doi:10.1186/s12870-020-02661-4>.

Lobell, D. B., Hammer, G. L., Mclean, G., et al. (2013). The critical role of extreme heat for maize production in the United States. *Nat. Clim. Change*, 3: 497–501. <https://doi.org/10.1038/nclimate1832>.

Lobell, D.B., Schlenker, W., Roberts, C, J. (2011). Climate trends and global crop production since 1980. *Science*, 333: 616–620. <https://doi.org/10.1126/science.1204531>.  
l'Université du Québec/AUPELF, 1992.

Lykhovyd, P. V. (2020). Sweet corn yield simulation using normalized difference vegetation index and leaf area index. *Journal of Ecological Engineering*, 21: 228-236. <https://doi.org/10.12911/22998993/118274>.

- MacDonald, R.B., Hall, F.G. (1980). Global crop forecasting. *Science*, 208: 670-679.
- Makhlouf, A., Dehbi, F., Hanachi, A., Harbi, M. (2012). Réponses de blé dur aux basses températures en relation avec la capacité de production. *Agriculture*, 3 : 13-23.
- Martignac, C. (2006). L'université d'Antananarivo et le CIRAD-initialisation à la télédétection (Bulletin de l'information Géographique Appliqué aux activités de recherche-Developpement BIG N°3).
- Mashaba, Z., Chirima, G., Botai, J. O., CombrinckL, Munghemezulu C., et Dube, E. (2017). Forecasting winter wheat yields using MODIS NDVI data for the Central Free State region. *South African Journal of Science*, 113 : 11-12. [dx.doi.org/10.17159/sajs.2017/20160201](https://doi.org/10.17159/sajs.2017/20160201).
- Mashaba, Z., Chirima, G., Botai, J. O., CombrinckL, Munghemezulu C., & Dube, E. (2017). Forecasting winter wheat yields using MODIS NDVI data for the Central Free State region. *South African Journal of Science*, 113 (11/12). [dx.doi.org/10.17159/sajs.2017/20160201](https://doi.org/10.17159/sajs.2017/20160201).
- Mcglinch, G. J., Jacquemin, S. J., Lindsey, L. E. (2021). Evaluating winter malting barley grain yield with fractional green canopy cover. *Crop, Forage and Turfgrass Mgmt*. 7:e20079. <https://doi.org/10.1002/cft2.20079>.
- McKee, T.B., Doeskin, N.J., Kleist, J. (1993). The relationship of drought frequency and duration to time scales, in proceedings of the. In: 8th conference on applied climatology. Anaheim, CA, Jan. 17–23, 1993. American Meteorological Society, Boston, pp 179–184.
- Mekhlouf A., Bouzerzour H., Benmahammed A., Hadj Sahraoui A. & Harkati N. (2006). Adaptation des variétés de blé dur (*Triticum durum* Desf.) au climat semi-aride. *Sécheresse* 17, 507-513.
- Mekliche, A. (1983). Contribution à l'établissement de la fertilisation azotée du blé d'hiver dans le haut Chélif. Thèse de magistère. I.N.A. Alger .81p.
- Mengel, G.E et al. (1987). Potassium and Plant Health. *Plant Physiology*, 85(3): 962–967
- Mhhabela, M., Bullock, P., Gervais, M., Finlay, G., & Sapirstein ,H. (2010). Assessing indicators of agricultural drought impacts on spring wheat yield and quality on the Canadian Prairies. *Agr. Forest Meteorol.* 150, 399–410.
- Ministère de l'Agriculture, (2017).
- Ministère de l'intérieur. (2016). Commune de Sétif (on line). Disponible sur : <https://interieur.gov.dz/images/pdf/setif2016.pdf>. Consulté le 20.03.2024.
- Ministre de l'agriculture, du développement rural et de pêche, 2022.

Moule, C. 1997. Céréale : Caractéristique généraux des céréales, Tome 1, Ed, la maison Rustique, paris, pp 5-6.

Nagy, A., Szabó, A., Adeniyi, O. D., et Tamás J. (2021). Wheat Yield Forecasting for the Tisza River Catchment Using Landsat 8 NDVI and SAVI Time Series and Reported Crop Statistics. *Agronomy*, 11(4), 652. <https://doi.org/10.3390/agronomy11040652>.

Nagy, A., Szabó, A., Adeniyi, O. D., et Tamás, J. (2021). Wheat Yield Forecasting for the Tisza River Catchment Using Landsat 8 NDVI and SAVI Time Series and Reported Crop Statistics. *Agronomy*, 11(4), 652. <https://doi.org/10.3390/agronomy11040652>.

ONS, 2021. <http://www.ons.dz/>.

Paden. Guide pour l'évaluation du rendement et l'estimation de le production horticole, Unité de Coordination du Programme (UCP).

Panek E, Gozdowski, D. (2020). Analysis of relationship between cereal yield and NDVI for selected regions of Central Europe based on MODIS satellite data. *Remote Sens. Appl. Soc. Environ.* 17, 100286. <https://doi.org/10.1016/j.rsase.2019.100286>.

Panek, E., et Gozdowski, D. (2021). Relationship between MODIS Derived NDVI and Yield of Cereals for Selected European Countries. *Agronomy*, 11(2), 340. <https://doi.org/10.3390/agronomy11020340>.

Patrignani, A., and Ochsner, T. E. (2015). Canopeo: A powerful new tool for measuring fractional green canopy cover. *Agronomy Journal*, 107:2312–2320. <https://doi.org/10.2134/agronj15.0150>.

Perthuis, C. d. (2022). Récupéré sur futuribles: <https://theconversation.com/les-impacts-de-laguerre-en-ukraine-sur-les-marches-agricoles-et-la-securite-alimentaire-178628>.

Pismennaya, E. V., Azarova1, M. Yu., Golosnoy, E.V., Odintsov, S.V., et Kipa, L.V. (2021). Relationship between NDVI index obtained from MODIS and winter wheat yield. *Earth and Environmental Science*, 848 012110. [doi:10.1088/1755-1315/848/1/012110](https://doi.org/10.1088/1755-1315/848/1/012110).

Pottier, E., Michaud, A., Farrié, J.P., Plantureux, S. et Baumont, R. (2012). Les prairies permanentes françaises au cœur d'enjeux agricoles et environnementaux. *Innovations agronomiques* 25 (2012), 85-97.

Ranjitha, G., Srinivasan, M.R., Rajesh, A. (2014). Detection and estimation of damage caused by Thrips *Thrips tabaci* (Lind) of cotton using hyperspectral radiometer. *Agrotechnology* 3, 1–5.

Raymonde, M.R. (1997). « Cartographie des zones sensibles a l'érosion dans les rougiers de camarés ( sud-aveyron) application a une région test sur la commune de Montlaur ». Mémoire de diplôme d'études supérieures spécialisées (DESS) Aménagement Rural et Développement Local Laboratoire d'études du Comportement des Sols Cultivés, Montpellier.

- Reed, V., Arnall, D. B., Finch, B., and Souza J. L. B. (2021). Predicting Winter Wheat Grain Yield Using Fractional Green Canopy Cove (FGCC). *International Journal of Agronomy*: 1443191, 8 pages. <https://doi.org/10.1155/2021/1443191>.
- Reyniers, M., Vrindts, E., et De Baerdemaeker, Josse. (2004). Optical Measurement of Crop Cover for Yield Prediction of Wheat. *Biosystems Engineering*. 89. 383-394.
- Richardson, A. J., Wiegand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, 43: 1541-1552. <https://www.asprs.org/wp-content/uploads/pers/1977>.
- Robin, M. (2002). Télédétection. Des satellites aux SIG. Coll Fac Géographie, Nathan Université, 2<sup>e</sup> édition.
- Rouabhi, A. (2017). Spatiotemporal Characterization of the Annual Rainfall in Setif Region-algeria. *Agriculture*. Volume 4, Numéro 3, Pages 31-38.
- Rouse, J. W., Haas, R. H., Schell, J. A. et Deering, D. W. (1974). *Monitoring* vegetation systems in the Great Plains with ERTS, United States. <https://doi.org/citeulike-article-id:12009708>.
- Ruiter, H., Macdiarmid, J. I., Matthews, R.B., Kastner, T., Lynd, L.R., et Smith P. (2017). Total Global Agricultural Land Footprint Associated with UK Food Supply 1986–2011. *Glob. Environ. Chang*, 43: 72–81, <https://doi.org/10.1016/j.gloenvcha.2017.01.007>.
- Sabins, Jr. (1987). Remote sensing – principles and interpretation. Second edition (New York: W. H. Freeman and Company), 449p.
- Samson, M.F et Desclaux, D. (2006). Amélioration de la valeur technologique et commerciale du blé dur : vers une réduction des taux de moucheture et de mitadin. Colloque régional du 21 Juin 2006. Campus INRA Montpellier : 5 p.
- Samuel, A. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Dev.* 3.
- Sarwar, M.H., Sarwar, M. F., Sarwar, M., Qadri, N. A., Moghal, S. (2013). The importance of cereals (Poaceae: Gramineae) nutrition in human health: A review. *Journal of cereals and oilseeds*, 4(3): 32-35.
- Schmitz, P. K., Kandel, H. J. (2021). Using Canopy Measurements to Predict Soybean Seed Yield. *Remote Sens.* 13, 3260. <https://doi.org/10.3390/rs13163260>.
- Semcheddine, N. (2015). Evaluation de la tolérance à la sécheresse chez le blé dur (*Triticum durum* Desf.) par l'étude du bilan hydrique et des paramètres phéno-morpho-physiologiques. Thèse de doctorat en sciences, UnivFerhat Abbas Sétif 103p.

Seo, B., Jihye, L., Kyung, Do. L., Sukyoung, H., Sinkyu, K. (2019). Improving remotely sensed crop monitoring by NDVI-based crop phenology estimators for corn and soybeans in Iowa and Illinois, USA. *Field Crops Res*, 238: 113–128. <https://doi.org/10.1016/j.fcr.2019.03.015>.

-Shanmugapriya, P., Rathika, S., Ramesh, T., et Janaki, P. (2019). Applications of remote sensing in agriculture-a review, *International Journal of Current Microbiology and Applied Sciences*, 8: 2270-2283. <https://doi.org/10.20546/ijcmas.2019.801.238>.

Sharma, A., Jain, A., Gupta, P., Chowdary, V. (2021). Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.3048415>.

Shepherd, M. J., Lindsey, L. E., et Lindsey, A. J. (2018). Soybean canopy cover measured with Canopeo compared with light interception. *Agricultural and Environmental Letters*, 3: 18-31. <https://doi.org/10.2134/aer2018.06.0031>.

Shewry, P. (2009). Wheat Journal of experimental Collection sciences techniques agricoles, 17<sup>ème</sup>éd, Paris, 464 p. *botany*. 60(6), pp15-37.

Skakun, S., Franch, B., Roger, J.C., Vermote, E., Becker-Reshef, I., Justice, C., Santamaria-Artigas, A. (2016). Incorporating yearly derived winter wheat maps into winter wheat yield forecasting model. *International Geoscience and Remote Sensing Symposium (IGARSS)*. <https://doi.org/10.1109/IGARSS.2016.7730869>.

Soltner, D. (1990). Les grandes productions végétales : céréales, plantes sarclées,

Soltner, D. (2005). La base de la production végétale Tom I. Le sol et son amélioration 24<sup>ème</sup>Edi. Collection science et technique agricole. 472P.

Soltner. (1988). Les grandes productions végétales. Les collections sciences et techniques agricoles, Ed. 16<sup>ème</sup> éditions 464P.

Son, N., Chen, C., Minh, V., Trung, N. (2014). A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agric. For. Meteorol*, 197: 52- 64. <https://doi.org/10.1016/j.agrformet.2014.06.007>.

Soudani, K. (2006). Introduction générale a la télédétection. Cours en ligne Faculté des Sciences d'Orsay – Université Paris Sud XI.

Sun, y., Zhang, S., Tao, F., Aboelenein, R., Amer, A. (2022), Improving Winter Wheat Yield Forecasting Based on Multi-Source Data and Machine Learning. *Agriculture*, 12, 571. <https://doi.org/10.3390/agriculture12050571>.

Tattaris, M., Reynolds, M. P. et Chapman, S. C. (2016). A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. *Frontiers in Plant Science*, 7(1131). <https://doi.org/10.3389/fpls.2016.01131>

Tenreiro, T. R., Garc'ia-Vila, M., G'omez J. A., et al. (2021). Using NDVI for the assessment of canopy cover in agricultural crops within modelling research. *Computers and Electronics in Agriculture*, 182: 10-038. <https://doi.org/10.1016/j.compag.2021.106038>.

Thompson, L.M. (1969). Weather and Technology in the Production of Corn in the U. S. Corn Belt 1. *Agron. J.* <https://doi.org/10.2134/agronj1969.00021962006100030037x>.

Tigkas, D., Vangelis, H., Tsakiris, G. (2015). DrinC: a software for drought analysis based on drought indices. *Earth Sci Inf*, 8(3):697–709.

Tsakiris, G. Vangelis, H. (2005). Establishing a drought index incorporating evapotranspiration. *Eur Water* 9(10):3–11.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ*, 8: 127-150.

Tucker, C. J., Pinzon, J.E., Brown, M. En Slayback, D.A., Pak, E.W., Mahoney, R El Saleous N. (2005). An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *Int. J. Remote Sens*, 26: 4485–4498. <https://doi.org/10.1080/01431160500168686>.

Tuğaç, M. G., Özbayoğlu, A. M., Torunlar, H., et Karakurt, E. (2022). Wheat Yield Prediction with Machine Learning based on MODIS and Landsat NDVI Data at Field Scale. *International Journal of Environment and Geoinformatics (IJECEO)*, 9(4):035-047. [doi.10.30897/ijegeo.1128985](https://doi.org/10.30897/ijegeo.1128985).

Vannoppen, A., Gobin, A., Kotova, L., Top, S., Cruz, L., Viksna, A., Aniskevich, S., Bobylev, L., Buntmeyers, L., Caluwaerts, S et al. (2020). Wheat Yield Estimation from NDVI and Regional Climate Models in Latvia. *Remote Sens.*, 12, 2206.

Vatin, N., El Hajj, M., et Baghdadi, N. (2021). Mapping land cover from SPOT5 images using machine learning algorithms. *International Journal of Remote Sensing*, 42(4): 1223-1242.

Verschuur, J., Li, S., Wolski, P., et al. (2007). Climate change as a driver of food insecurity in the Lesotho-South Africa drought. *Sci. Rep.* 11: 38-52. <https://doi.org/10.1038/s41598-021-83375-x>.

Vozhehova, R., Maliarchuk, M., Biliaieva, I., Lykhovyd, P. V., Maliarchuk, A., et Tomnytskyi, A., (2020). Spring row crops productivity prediction using normalized difference vegetation index. *Journal of Ecological Engineering*, 21(6), 176-182. <https://doi.org/10.12911/22998993/123473>.

Wang, M., Tao, F.L., Shi, W.J. (2014). Corn yield forecasting in northeast china using remotely sensed spectral indices and crop phenology metrics. *J. Integr. Agric.* [https://doi.org/10.1016/S2095-3119\(14\)60817-0](https://doi.org/10.1016/S2095-3119(14)60817-0).

Wang, P., Zhou, G., Liu, X., Li, Y., Wu, X., Li, Z., ... et Ma, W. (2021). Quantifying the Impact of Climate Variability and Human Activities on Surface Water Dynamics in Arid Regions Using Landsat Data. *Remote Sensing*, 13(4), 682.

Wang, Y., Xu, X., Huang, L., Yang, G., Fan, L., Wei, P., & Chen, G. (2019). An Improved CASA Model for Estimating winter wheat yield from remote sensing images. *Remote Sens*, 11(9), 1088. <https://doi.org/10.3390/rs11091088>.

Wang., Qunming., Shi., Wenzhong., Atkinson., Peter, M., et Zhao., Yuanling. (2015). Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*, 166: 191–204.

Wittich, KP., Hansing, O. (1995). Area-averaged vegetative cover fraction estimated from satellite data. *Int J Biometeorol*, 38: 209-215. <https://doi.org/10.1007/BF01245391>.

Xiong, Y., West, C. P., Brown, C. P., et Green, P. E. (2019). Digital image analysis of old world bluestem cover to estimate canopy development. *Agronomy Journal*, 111: 1247–1253. <https://doi.org/10.2134/agronj2018.08.0502>.

Xu, L., Jia, K., Jiang, L., Qiu, Y., & Zhang, Y. (2020). Quantitative Monitoring of Vegetation Dynamics Using Landsat Time Series: A Case Study in Central Asia. *Remote Sensing*, 12(8), 1330.

Yan, H., Harrison, M. T., Lui, K., et al. (2022). Crop traits enabling yield gains under more frequent extreme climatic events. *Sci. Total Environ*, 808: 152-170. <https://doi.org/10.1016/j.scitotenv.2021.152170>.

Zhang, H. K., et Roy, D. P. (2020). MODIS-based vegetation greenness changes over the past decade in United States croplands. *Remote Sensing*, 12(1), 100.

Zhang, P.P., Zhou, X.X., Wang, Z.X., Mao, W., Li, W.X., Yun, F., Guo, W.S., Tan, C.W . (2020). Using HJ-CCD image and PLS algorithm to estimate the yield of field-grown winter wheat. *Sci. Rep*, 10: 51-73. <https://doi.org/10.1038/s41598-020-62125-5>.

# **Annexes**

## Annexes

**Tableau (23)** : la moyenne de la production de blé et d'orge au niveau de la wilaya de Sétif entre 2002 et 2022 (DSA Sétif. 2023).

Années	Blé (t/ha)	Orge (t/ha)
2002	0.72	0.8
2003	1.7	2
2004	1.95	2.2
2005	0.76	1.21
2006	1.15	1.12
2007	1.45	1.87
2008	0.86	0.78
2009	1.68	2.13
2010	1.56	1.3
2011	1.7	1.84
2012	1.76	1.28
2013	2.05	1.76
2014	0.76	0.72
2015	0.7	0.51
2016	1.2	2
2017	0.8	0.7
2018	2.3	1.76
2019	1.72	1.32
2020	1.23	1.82
2021	0.78	0.81
2022	1.24	1.1

**Tableau (24)** : la moyenne de la production de blé et d'orge au niveau de la commune de Sétif entre 2002 et 2022 (DSA Sétif. 2023).

Années	Blé (t/ha)	Orge (t/ha)
2002	0.35	0.41
2003	1.6	2.3
2004	1.95	0.2
2005	0.89	1
2006	1.81	2.4
2007	1.4	1.5
2008	0.95	0.9
2009	2.19	1.98
2010	1.5	1.5
2011	1.86	2.69
2012	2.1	1.1
2013	2.5	1.5
2014	0.5	0.5
2015	0.75	0.7
2016	2.4	3
2017	0.67	0.5
2018	2.59	2.4
2019	1.3	1.24
2020	1.11	1.8
2021	0.62	0.79
2022	1.13	0.97

**Tableau (25)** : la moyenne de la production de blé et d'orge au niveau des parcelles expérimentales de l'ITGC de Sétif entre 2002 et 2022 (DSA Sétif. 2023).

Années	Blé (t/ha)	Orge (t/ha)
2002	1.78	0.7
2003	1.7	0.95
2004	2	1.7
2005	1.6	1.04
2006	1.2	1.02
2007	2.52	2
2008	1.9	1.8
2009	2	2
2010	2.9	1.98
2011	2.84	2.1
2012	2.02	1.285
2013	2.46	2
2014	2.11	1.4
2015	1	1.33
2016	2.1	2.01
2017	1.75	2.128
2018	3.07	2.1
2019	2.05	2.2
2020	2.09	2.11
2021	2	1.5
2022	2.01	1.42

## ملخص

يشكل اليوم تغير المناخ العالمي تهديدا كبيرا للأمن الغذائي، مما يؤثر على إنتاج المحاصيل الزراعية. ولمواجهة هذا التحدي، من الضروري وضع نظام تشغيلي للتنبؤ بمحصول الحبوب في بداية الموسم الزراعي. إن ظهور تقنيات الاستشعار عن بعد في النظم الزراعية الدقيقة جعل التنبؤ بمحصول الحبوب ممكنا، وبالتالي توفير معلومات قيمة عن التغيرات الزمنية والمكانية في ظروف إنتاج الحبوب في الأراضي الزراعية الكبيرة وصغيرة المساحة. الهدف من هذا الدراسة هو تقييم مدى فعالية استخدام معطيات MODIS-NDVI و FGCC التي تم الحصول عليها من خلال تطبيق Canopeo® إضافة إلى الكتلة الحيوية الخضراء للتنبؤ بمحصول الحبوب من القمح والشعير من 2020 إلى 2022 على مستوى الحقل من ناحية، ومن ناحية أخرى تقييم إمكانية استخدام قيم MODIS-NDVI وحدها (1)، والجمع بين بيانات NDVI مع كميات الأمطار (2)، واستخدام نماذج التعلم الآلي (3)، للتنبؤ بمحصول الحبوب (القمح والشعير) ما بين 2002 و 2022 على ثلاث مستويات: الحقول التجريبية لـ ITGC، بلدية وولاية سطيف. و من خلال النتائج التي تم الحصول عليها يتضح أن FGCC يقدم أفضل أداء في تقييم إنتاجية القمح والشعير من بين الطرق الثلاث التي تمت دراستها على مستوى ITGC ( $R^2$  يتراوح من 0.78 إلى 0.79، RMSE بين 0.051 و 0.107، MSE من 0.004 إلى 0.018)، بالإضافة إلى أن الاستخدام المشترك لبيانات NDVI مع كميات الأمطار يقدم نتائج التنبؤ الأكثر دقة ( $R$  من 0.64 إلى 0.77، RMSE بين 0.277 إلى 0.343، MSE من 0.092 إلى 0.246). تشير هذه النتائج إلى أن تطبيق Canopeo® أثبت أنه وسيلة سريعة، موثوقة وفعالة لتقدير إنتاج حبوب القمح والشعير في قطع الأراضي الصغيرة. ومن ناحية أخرى، فإن الاستخدام المشترك لـ NDVI مع هطول الأمطار يكون أكثر فعالية ودقة مقارنة بالطرق الثلاث المستعملة. أظهرت نماذج التعلم الآلي أداءً جيدًا في توقع إنتاج القمح، خاصة نموذج " Stochastic Gradient Descent" الذي حقق قيم  $R^2$  تبلغ 0.651 و 0.783 وقيم RMSE تبلغ 0.368 طن/هكتار و 0.231 طن/هكتار على مستوى البلدية والولاية، على التوالي. بينما، بالنسبة لشعير، أظهر النموذج "Tree" أداءً جيدًا على مستوى الحقل ( $R^2 = 0.672$ ، RMSE = 0.289 طن/هكتار) والولاية ( $R^2 = 0.390$ ، RMSE = 0.151 طن/هكتار). ومع ذلك، سجل نموذج "Gradient Boosting" أداءً أقل على مستوى البلدية ( $R^2 = 0.358$ ، RMSE = 0.375 طن/هكتار).

**الكلمات المفتاحية:** مردود، الحبوب، FGCC، الكتلة الحيوية، الاستشعار عن بعد، NDVI، التعلم الآلي.

## Résumé

Aujourd'hui, le changement climatique mondial constitue une menace importante pour la sécurité alimentaire, affectant le rendement des cultures. Pour faire face à ce défi, il est essentiel de mettre en place un système opérationnel de prévision du rendement céréalier en début de saison agricole. L'avènement des technologies de télédétection dans les systèmes agricoles de précision a rendu possible la prévision des rendements céréaliers, fournissant ainsi des informations précieuses sur les variations temporelles et spatiales des conditions de production céréalières sur les terres cultivées à grande et à petite échelle. Le but de cette recherche était d'une part, évaluer le potentiel de l'utilisation des données MODIS-NDVI, FGCC obtenue de l'application Canopeo® et de la biomasse sèche pour prédire les rendements céréaliers du blé et de l'orge entre 2020 à 2022 au niveau parcellaire. D'autre part évaluer le potentiel d'utilisation des valeurs MODIS-NDVI seul (1), combinaison des données NDVI avec les précipitations (2), et l'utilisation des modèles d'apprentissage automatique (3), pour prédire le rendement des céréales (blé et orge) entre 2002 et 2022 dans trois échelles d'étude : les parcelles expérimentales, la commune et la wilaya de Sétif. Les résultats obtenus montrent que FGCC présente les meilleures performances d'évaluation des rendements de blé et d'orge parmi les trois méthodes étudiées au niveau parcellaire, avec ( $R^2$ = variait de 0.78 à 0.79, RMSE= entre 0.051 et 0.107, MSE= de 0.004 à 0.018), ainsi que l'utilisation combinée du NDVI au données de pluviométrie présente les résultats de prédiction les plus précis avec ( $R$ = de 0.64 à 0.77, RMSE= entre 0.277 à 0.343, MSE= de 0.092 à 0.246 ). Ces résultats indiquent que l'application Canopeo® s'est avérée être une méthode rapide et fiable pour estimer le rendement en grains de blé et d'orge sur des petits parcelles. En revanche, l'utilisation combinée du NDVI avec les précipitations est plus efficace et précise que les trois autres méthodes utilisées. Les modèles de machine learning ont montré des bonne performances de prédiction pour le rendement de blé, notamment "Stochastic Gradient Descent" avec des  $R^2$  de 0.651 et 0.783 et des RMSE de 0.368t/ha et 0.231t/ha, au niveau communal et wilaya, respectivement. Tandis que pour l'orge, le modèle 'Tree' a montré des bonnes performances au niveau parcellaires ( $R^2$ =0.672, RMSE=0.289t/ha) et wilaya ( $R^2$ =0.390, 0.151t/ha). Cependant, le modèle "Gradient Boosting" a enregistré des performances moindre au niveau communale ( $R^2$ =0.358, RMSE=0.375t/ha).

**Mots clé :** Rendement, Céréales, FGCC, Biomasse, Télédétection, NDVI, Apprentissage automatique.

## Abstract

Today, global climate change poses a significant threat to food security, affecting crop yields. To face this challenge, it is essential to put in place an operational system for forecasting cereal yield at the start of the agricultural season. The advent of remote sensing technologies in precision agricultural systems has made the prediction of grain yields possible, providing valuable information on temporal and spatial variations in grain production conditions on large- and small-scale croplands. The aim of this research was, on the one hand, evaluate the potential of using MODIS-NDVI, FGCC data obtained from the Canopeo® application and biomass to predict cereal yields of wheat and barley between 2020 to 2022 at field level. On the other hand, evaluate the potential for using MODIS-NDVI values alone (1), combination of NDVI data with precipitation (2), and the use of machine learning models (3), to predict cereal yield (wheat and barley) between 2002 and 2022 in three study levels: the ITGC experimental plots, the commune and the wilaya of Sétif. The results obtained show that FGCC presents the best performance in evaluating wheat and barley yields among the three methods studied at the ITGC plot level with ( $R^2$ = varied from 0.78 to 0.79, RMSE= between 0.051 and 0.107, MSE= from 0.004 to 0.018), as well as the combination of NDVI and rainfall data presents the most accurate prediction results with ( $R$ = from 0.64 to 0.77, RMSE= between 0.277 to 0.343, MSE= from 0.092 to 0.246). These results indicate that the Canopeo® application has proven to be a rapid and reliable method for estimating wheat and barley grain yield on small plots. On the other hand, the combined use of NDVI with precipitation is more efficient and precise than the other three methods used. Machine learning models have shown good prediction performance for wheat yield, particularly "Stochastic Gradient Descent," with  $R^2$  values of 0.651 and 0.783 and RMSE values of 0.368 t/ha and 0.231 t/ha at the communal and wilaya levels, respectively. Meanwhile, for barley, the 'Tree' model exhibited good performance at the parcel level ( $R^2=0.672$ , RMSE=0.289 t/ha) and wilaya level ( $R^2=0.390$ , RMSE=0.151 t/ha). However, the "Gradient Boosting" model recorded lower performance at the communal level ( $R^2=0.358$ , RMSE=0.375 t/ha).

**Keywords:** Yield, cereals, FGCC, Biomass, Remote sensing, NDVI, Machine learning.



ISSN: 2455-9377

# Cereal yield forecasting in semi-arid region of Algeria using MODIS-NDVI

Hakima Boulaaras\*, Tarek Bouregaa

Plant and Animal Production Improvement Laboratory, Department of Agronomy, Faculty of Natural and Life Sciences, Ferhat Abbas University Setif 1, 1900 Setif, Algeria

## ABSTRACT

The prediction of cereals yields today is very important for global food security and helps decision-makers in the import-export operations of countries, especially with the rise world population. The advent of remote sensing technologies in precision farming systems has made cereal yield predictions possible, providing valuable insights into the temporal and spatial variations in cereal conditions across both large and small-scale crop lands. Among the various vegetation indices used to analyze these conditions, the normalized difference of vegetation index (NDVI) has emerged as a key indicator. The main objective of this study is to evaluate the possibility of using MODIS-NDVI data to forecast the yield of cereal crops (wheat and barley) in semi-arid region of Algeria (Setif). Additionally, identify the optimal timing for reliable and accurate crop yield forecasts. The remote sensing data utilized in this study covered the growing seasons from February to June, from 2002 to 2022. The results indicated a strong correlation between cereal grain yield and NDVI from late February to mid-March, with  $R^2$  values ranging from 0.55 to 0.82 for the two cereal species. The RMSE of the NDVI based prediction model ranged from 0.01 t ha<sup>-1</sup> to 0.276 t ha<sup>-1</sup>. The approximate average increase in the grain yield of barley and wheat lies between 0.659 to 0.746 t ha<sup>-1</sup> with an increase of 0.1 in NDVI value. These results demonstrate the effectiveness of using MODIS-NDVI data for cereal yield forecasting in semi-arid region of Algeria, offering valuable predictions two to three months before the harvest.

**KEYWORDS:** Predicting, Yield, Remote sensing, NDVI

**Received:** August 25, 2023  
**Revised:** January 18, 2024  
**Accepted:** January 19, 2024  
**Published:** February 02, 2024

**\*Corresponding author:**  
Hakima Boulaaras  
E-mail: boulaarashakima@gmail.com

## INTRODUCTION

Wheat, along with rice and maize, is one of the main three world food crops (Cai *et al.*, 2019). Soft wheat is one of the most important food crops that feed 40% of the world population (Liu *et al.*, 2020). Without forgetting, barley grain which ranks fourth in terms of quantity produced and area cultivated in the world after wheat, rice and corn (Geng *et al.*, 2022). In Algeria cereals play a significant role in the dietary habits of the population, encompassing production and processing activities such as semolina production and bakery in the food industry (Ammar, 2014). According to the Algérie Eco (2022), the area occupied by cereals is 3.5 million ha which is very small compared to the total area of Algeria (238 million ha). The national agricultural production is heavily influenced by its climatic conditions, which are primarily characterized by annual fluctuations in precipitation, water scarcity, and high temperatures during crop growth periods, these factors have a negative impact on production (Mekhlouf *et al.*, 2012). In 2022, Algeria imported 10.6 million tons of cereals. The majority of these imports were soft wheat, accounting for almost 6.1 million tons, followed by maize with 2.6 million tons (a decrease from 4.8 million tons in the previous campaign), durum wheat with

nearly 1.4 million tons, and 571,000 tons of barley (Algérie Eco, 2022). For this reason, accuracy and timeliness of regional crop yield estimation is crucial for ensuring national and international food security (Becker-Reshef *et al.*, 2020), it is also beneficial for policymakers in making informed decisions regarding import and export policies and determining acceptable support prices for the market (Dorosh & Salamn, 2006). In particular, weather variability and biological stresses (including pathogens and arthropods) have an increasing impact on food security (Al-Ani *et al.*, 2011; Khalaf *et al.*, 2019, 2023; Adhab & Alkuwaiti, 2022), the importance of accurate and timely regional crop yield estimation has become even more significant (FAO, 2018). Although traditional field surveys and crop statistics are useful for accurately estimating crop yield, they prove to be insufficient when predicting crop yield for large regions due to constraints such as budget, time, and shortage of skilled manpower (Fang *et al.*, 2008). Using Artificial Intelligence (AI) and computerization have contributed to the field of biotechnology and agriculture and supported the sustainability endeavor (Anaz *et al.*, 2023). Advancements in satellite sensor technology have led to the development of remote sensing, which is a science and technique focused on acquiring information about on-land objects from satellite imagery without the need for direct

Copyright: © The authors. This article is open access and licensed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted, use, distribution and reproduction in any medium, or format for any purpose, even commercially provided the work is properly cited. Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

contact (Sabins, 1987). Today, remote sensing is widely used for monitoring and predicting crop yields across region of varying sizes due to its large coverage area, non-invasive nature, and ability to provide rapid and long-term time series data. This makes it an important tool for policymakers and stakeholders in ensuring food security and developing effective agricultural policies (Zhang *et al.*, 2020). The application of vegetation indices (VIs) derived from satellite images is considered the most promising and convenient method for forecasting crop yield using remote sensing data, they are effective indicators of vegetation status and have a positive correlation with crop yield. Among the various VIs, the Normalized Difference Vegetation Index (NDVI) is frequently used for studying vegetation dynamics because of its high correlation with photosynthetic capacity, leaf area index, biomass, and net primary productivity (Li *et al.*, 2014). The NDVI is also a popular choice for crop yield prediction due to its accessibility and ease of use (Phiri *et al.*, 2020). The Normalized Difference Vegetation Index (NDVI), which was first introduced by Rouse *et al.* (1974), defined as the ratio between the difference in near-infrared and red spectra reflections from the Earth's surface and their sum. The NDVI scale ranges from -1 to 1, with higher positive values indicating greater vegetation coverage and activity (Fang *et al.*, 2004). Negative NDVI values indicate the presence of clouds, snow, water, or a bright, non-vegetated surface (Yin & Williams, 1997). In recent years, the focus of remote sensing-based yield forecasting research has shifted towards the use of National Aeronautics and Space Administration's (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS) and other sensors with different spatial resolutions. The MODIS data has a spatial resolution of 250 m, 500 and 1000 m (Atzberger *et al.*, 2016). Remote sensing studies used the empirical regression models linking historical crop yield as dependent variable and administrative units-averages of seasonal satellite data for cultivated region as independent variable (Becker-Reshef *et al.*, 2010). Numerous research studies have proved the effectiveness of remote sensing in predicting crop yields, such as, Mulianga *et al.* (2013) used the MODIS-NDVI data in the study on the sugarcane yield estimation on large territories. Kouadio *et al.* (2014) applied MODIS-NDVI and EVI data to forecast spring wheat yield at the ecodistrict scale. Huang *et al.* (2013) used time series data of NDVI values in their regression model to predict rice yield. Nagy *et al.* (2018) developed regression models using 15 different peak-season MODIS-derived NDVI time series to predict wheat and maize yields. The reported yield values were regressed against the NDVI data, and they found that MODIS-NDVI data could effectively predict crop yield for the Tisza river catchment area 6-8 weeks before harvest. Similarly, Lykhovyd (2020) and Vozhehova *et al.* (2020) applied NDVI-based regression models for forecasting yield of spring row crops at the field scale. The combination of crop models and remote sensing data has increasingly been used to forecast crop yield.

This study fills a significant research gap by introducing a new methodology for accurately forecasting cereals grain yield in Algeria's semi-arid region using MODIS-NDVI remote sensing data. The study's objectives are two-fold. To begin, it intends to assess the feasibility of using MODIS-NDVI data at various dates between 2002 and 2022 to forecast cereal yields before

harvest, specifically wheat and barley, in semi-arid region of Algeria. Second, it seeks to determine the best time of year for accurate prediction of cereal grain yield at a regional level in Algeria, given that previous studies have produced inconsistent results regarding the best time for prediction in this specific semi-arid area.

## MATERIALS AND METHODS

### Study Area

The research was conducted at the Technical Institute of Large Crops (ITGC) in Setif, Algeria. The experimental site is located at latitude 36°10'17' North, longitude 5°21'55' East, and an altitude of 1080 m (Google Earth Pro, 2023). The experimental site is located in the central zone of the high plains, which is favorable for cereal cultivation (Figure 1). The climate site was characterized by hot and dry summers and cold and humid winters (Chennafi *et al.*, 2006). The annual precipitation reaches 458 mm (Rouabhi, 2017), which mainly occurs between January to April and an average annual temperature of 13.5 °C (Climate Data, 2022). The experimental site is characterized by flat, relatively infertile land and a high risk of late frost and drought towards the end of the crop cycle. The physic-chemical analysis shows that the soil has a silty-clayey texture and an average organic matter content of 2.13%. The Bulk density of is 1.51 g/cm<sup>3</sup>, with a field capacity of 23% and a wilting point of 10%.

### Data Collection

#### Crop yield data

The crop grain yield data (t ha<sup>-1</sup>) of wheat and barley were collected from the Technical Institute of Large Crops (ITGC) of Setif, which cover a period of twenty years (2002-2022).

#### MODIS-NDVI data

The time series of average NDVI for the study area were obtained from the Global Agricultural Monitoring (GLAM) system (<https://glam1.gsfc.nasa.gov/>), hosted by the USDA and NASA. The data was downloaded on January 12<sup>th</sup>, 2021 (GIMMS, 2021). The GLAM system was developed as part of the Global Agricultural Monitoring project. This initiative has the objective of regularly assessing worldwide forecasts of agricultural production and conditions affecting global food security in an unbiased and timely manner. The GLAM system provides 8-day composited NDVI data sets that are derived from MODIS sensors onboard the Terra satellite platform. These data sets have a spatial resolution of 250 or 500 m and are based on the MOD09 product (MODIS collection 6). Our study focused on the growing season in Algeria, which spans from February end to June 1<sup>st</sup>, and covers data collected from 2002 to 2022. To obtain the NDVI values, we used the GFSAD30 2015 Crops crop mask developed by the NASA Global Food Security-Support Analysis Data project, which has a spatial resolution of 30 m (<https://croplands.org>) (USGS, 2021). To ensure high-quality data, the collected information underwent radiation, atmospheric, and geometric corrections. These measures were taken to make the data more accurate and reliable for use in studying regional vegetation.

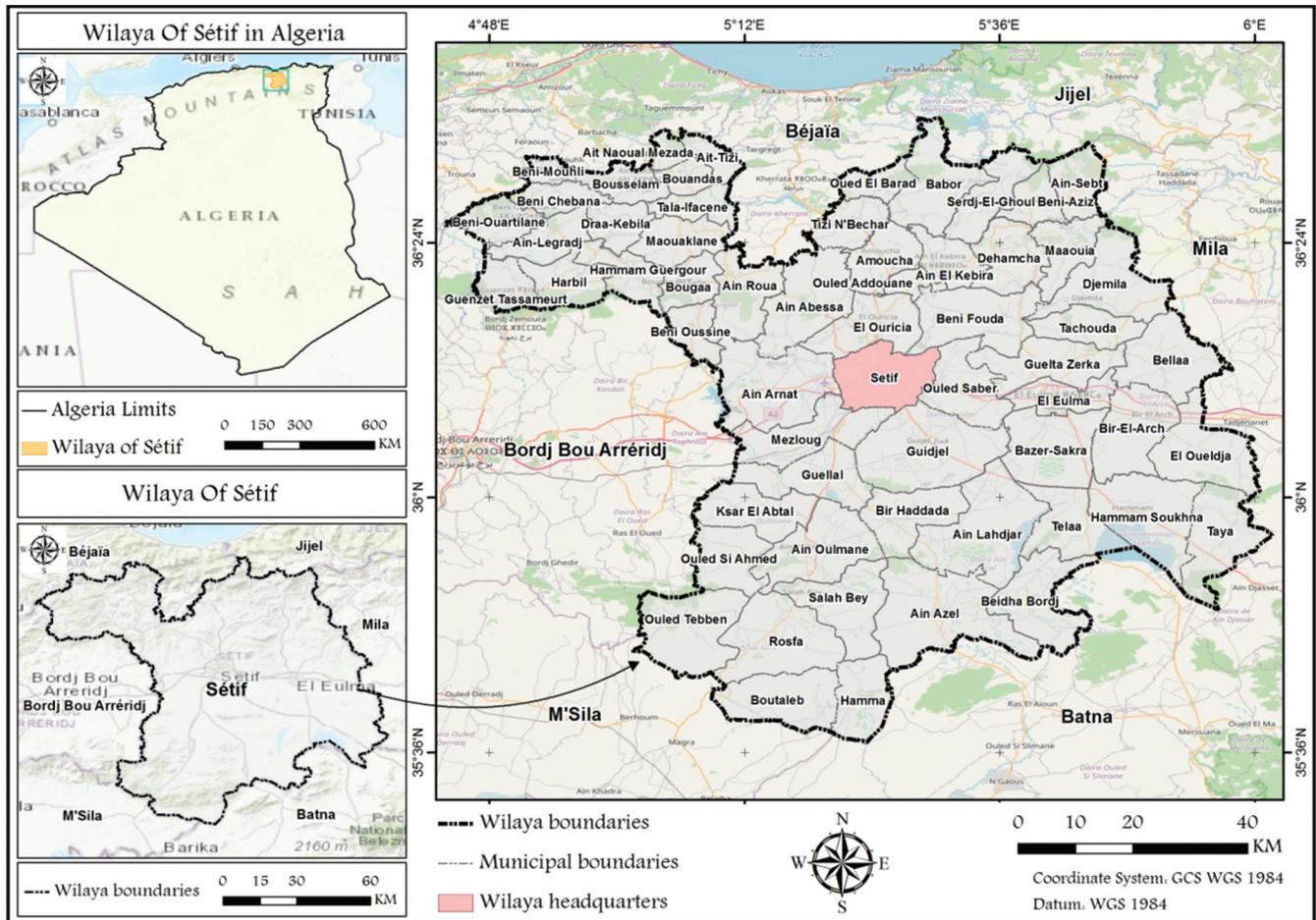


Figure 1: Geographical location of the experimental site

## Statistical Analysis

In this study we employed a separate correlation and linear regression analyses for each crop. The independent variable was the NDVI values, while the dependent variable was the yield data of two cereal crops. Regression analysis aims to identify trends in the relationship and describe the relationship mode with a particular function, thereby quantifying causal relationships. The regression coefficient measures the average change in the explanatory variable per unit change in the response variable. Meanwhile, the linear correlation coefficient, determines the percentage of variance in the response variable that is explained by the factor variable, thereby indicating its reliability. We can represent this relationship using the following equation:

$$Y = \beta_0 + \beta_1 X \quad (1)$$

Where  $\beta_1$  represents the regression coefficient. Parameter  $\beta_0$  can usually only be interpreted mathematically if the variable  $X$  is set to 0, then  $\beta_0$  is the estimate given 0 in  $X$ .

To assess the performance of the developed models, widely employed statistical metrics were used in this study. The coefficient of determination ( $R^2$ ) was used to measure the degree of linear relationship between observed and forecasted

cereal yield. The mean squared error (MSE) was used to measure the average of the squares of the errors. Meanwhile, the Root Mean Square Error (RMSE) measured the discrepancy of the forecasted yield around observations. All statistical analyses were carried out using SPSS (version 19). The  $R^2$ , RMSE, MSE was calculated using equations (2), (3) and (4).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_i^1)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (4)$$

## RESULTS

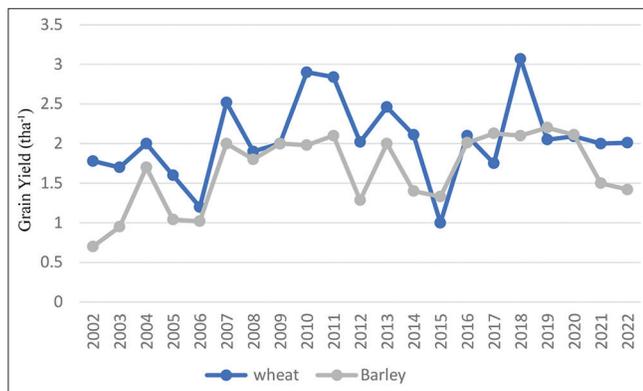
### Temporal Variability of Cereal Grain Yield

The average grain yields of two cereals varied over the study period (2002-2022) are presented in Figure 2. Wheat had

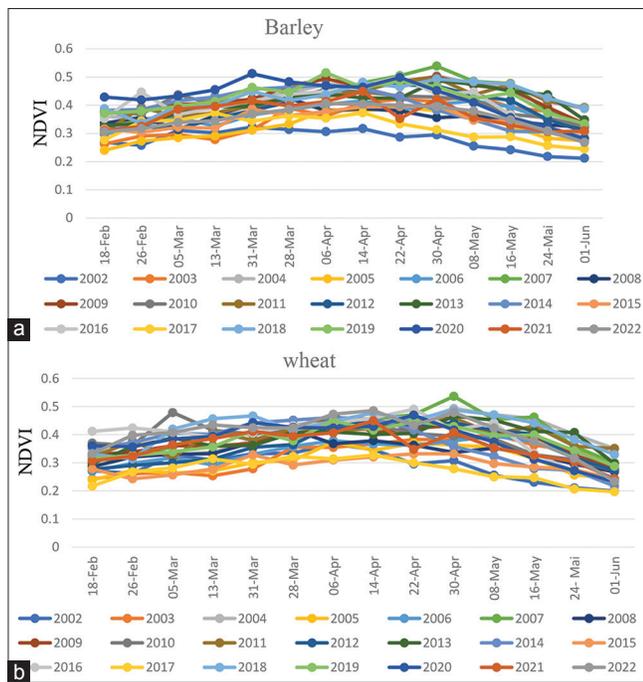
the highest averaged grain yield in 2018 with 3.07 t ha<sup>-1</sup> while barley had the highest averaged grain yield in 2019 with 2.2 t ha<sup>-1</sup>. Conversely, the lowest grain yield for wheat was observed in 2015 with 1 t ha<sup>-1</sup>, for barley it was in 2002 with 0.7 t ha<sup>-1</sup>. Differences in the mean grain yield of two cereals in arid and semi-arid regions of Algeria across years were primarily due to weather conditions, such as Variability of rainfall, very low temperatures during winter or droughts during spring and early summer.

### NDVI Temporal Variability from 2002 to 2022

Figure 3 illustrate the temporal patterns of vegetation index throughout the growth period of barley and wheat crops respectively. The NDVI values were lowest during the transplanting phase and gradually increased as the vegetative parts grew. They reached their peak during the late vegetative



**Figure 2:** Temporal variability of cereals grain yield (wheat and barley) from 2002 to 2022



**Figure 3:** NDVI temporal variability for a) barley and b) wheat from 2002 to 2022

phase and remained high until the flowering phase, which occurred between March and April.

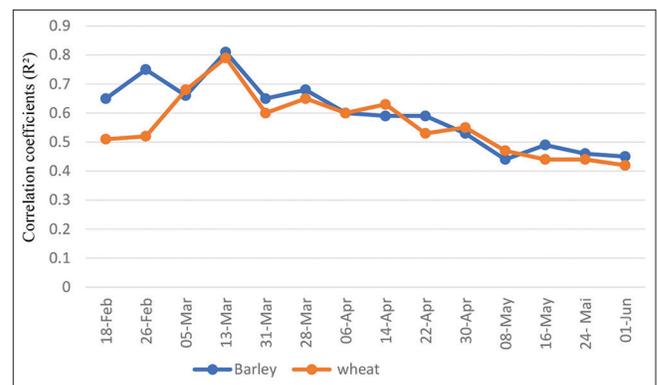
During the post-flowering phase, (i.e., the ripening phase), the vegetation index values started to decrease and reached their minimum at the fully ripened harvesting phase in June. The NDVI values ranged from 0.212 to 0.539 for all study years for barley, and from 0.197 to 0.537 for wheat. The NDVI values varied from one year to another, depending on factors such as rainfall, temperature during the seasons and sowing dates.

### Relationship between NDVI at Different Dates and Cereals-grain Yield

The NDVI is an effective tool for measuring the impact of various environmental factors and their interactions with crops. It provides valuable information on the combined effects of weather conditions, crop varieties, soil types, cultivation methods, and other factors. The results of our study demonstrate a strong linear relationship between MODIS-NDVI and grain yield for the two winter cereals (wheat and barley) at the regional level. The correlation coefficients are presented graphically in Figure 4.

The highest correlation between NDVI and cereals grain yield occurs between 26 February and 13 March ( $R^2$  ranged from 0.71 to 0.8 for barley,  $R^2$  ranged from 0.55 to 0.82 for wheat). The peaks of correlation correspond to the NDVI peaks during the growing season. We can observe that at later dates (growing season progresses), the relationships and the prediction accuracy were weaker, which may have been caused by NDVI saturation in the later growth stages of cereals. Based on these results, the best time to predict cereals grain yield accurately using MODIS-NDVI in semi-arid region of Algeria is the beginning of spring, specifically 13 March (120<sup>th</sup> after sowing). We can observe that at later dates, the relationships and the prediction accuracy were weaker, which may have been caused by NDVI saturation in the later growth stages of cereals.

A linear regression analysis was conducted to examine the relationship between NDVI and cereals grain yield (wheat and barley). The means NDVI values from 18 February to

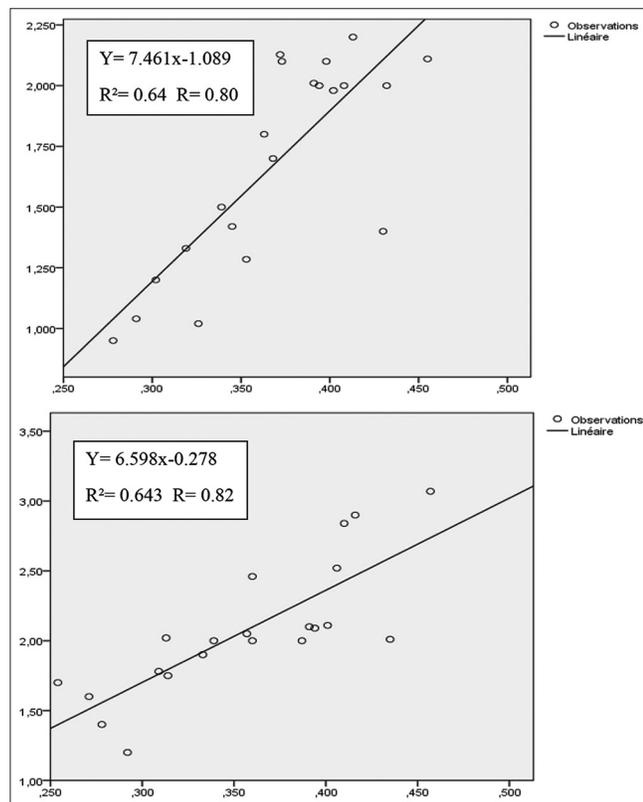


**Figure 4:** Correlation coefficients between grain-yield and NDVI for the two cereals (barley, and wheat) from February 18<sup>th</sup> to June 1<sup>st</sup>, covering the period from 2002 to 2022

01 June (2002-2022) were used as independent variables, while the dependent variable was grain yield for wheat and barley. The results are graphically presented in Figure 5, indicating a strong relationship between NDVI in early spring (13 March) and grain yield for two cereals. The regression coefficients for wheat and barley were 6.598 and 7.461 respectively which implies that an increase of 0.1 in NDVI is associated with an average increase of 0.659 t ha<sup>-1</sup> and 0.746 t ha<sup>-1</sup> in grain yield for wheat and barley respectively. The strength of the relationship is supported by strong Pearson's correlation coefficients (R) of 0.82 and 0.80 for wheat and barley, respectively.

### Model Performance Verification

The accuracy of the models was assessed by comparing the predicted yields with the actual yields obtained in the study area. Four measures of forecast accuracy were used: root mean square errors (RMSE), mean square error (MSE), correlation coefficients (R) and the coefficient of determination (R<sup>2</sup>) for each cereals crop (wheat and barley). The results showed a strong correlation between the measured and predicted yield, with correlation coefficients of 0.80, 0.901 for wheat and barley, respectively. And low RMSE values ranged from 0.01 to 0.276 t ha<sup>-1</sup>, the MSE values ranged from 0.061 to 0.076, the results are presented in the Table 1, these results indicating that the predicted values are close to the actual observed values, which confirm that the yield was predicted with great accuracy, three months before harvest which implies the proper functioning of the created model (Figures 6 & 7).



**Figure 5:** Linear regression model and correlation of barley (above), wheat (At the bottom), yield with the MODIS-NDVI for March 13<sup>th</sup>

## DISCUSSION

Forecasting crop yields is a critical and complex task in modern agriculture due to various challenges. These challenges include the impacts of global climate change, such as extreme weather events like droughts, floods, and other natural disasters, as well as the increasing global population and demand for food. Accurately predicting crop yields is crucial for effective agricultural planning, maintaining food safety and availability. Satellite remote sensing is widely used for forecasting cereal yield production, given its ability to be utilized at a global level. According to our results, we have demonstrated that the early spring stage of development is critical for achieving high grain yield for the three dominant cereals (durum wheat, soft wheat, and barley) in Algeria's most valuable semi-arid areas. Mkhabela *et al.* (2010) found that MODIS-NDVI could effectively predict crop yields across the Canadian Prairies with a lead time of one to two months before harvest. The results indicated that a power function best described the relationship between MODIS-NDVI and grain yield for all the crops and agro-climatic zones studied, with coefficient of determination (R<sup>2</sup>) ranging from 0.48 to 0.90 for barley and 0.47 to 0.80 for wheat. Interestingly, the strength of the relationship was similar or even stronger when compared to the findings of our study, with R<sup>2</sup> = 0.64 for barley, and R<sup>2</sup> = 0.643 for wheat. In a study on predicting the grain yields of wheat Adeniyi *et al.* (2020), proves that the use of Normalized Difference of Vegetation Index (NDVI) derived from Landsat 8 time series data, from 2013 to 2019 growing seasons, are effective in predicting winter wheat yield in Jász-Nagykun-Szolnok county (Northern Great Plain region of central Hungary). The highest determination coefficient (R<sup>2</sup> = 0.569) was found on the 160<sup>th</sup> day, which is lower than the value obtained in the current study (R<sup>2</sup> = 0.643). The study reported an average increase of 0.1 t/ha in grain yield of wheat with an increase of 0.1 in NDVI value, which is lower than the result obtained in the current study. Panek and Gozdowski (2021), employed a linear regression analysis to investigate the correlation between normalized difference vegetation index (NDVI) obtained from MODIS satellite data, and grain yield of wheat and barley in 20 European countries between 2010 and 2018. They found a strong relationship between NDVI and cereals grain yield in early spring for several countries, including Croatia, Czechia, Germany, Hungary, Latvia, Lithuania, Poland and Slovakia, which is similar to the results of our study. The strength of the relationship was also similar to our study, with an R<sup>2</sup> of 0.610 for wheat and 0.614 for barley. The results of the regression showed that a 0.1 unit increase in NDVI is related to a 1.35-1.65 t ha<sup>-1</sup> increase in grain yield of cereals. Wang *et al.* (2019), employed an enhanced Carnegie-Ames-Stanford approach (CASA) model, combined with time-series satellite remote sensing images obtained from MODIS, to estimate the

**Table 1:** Model performance results expressed as the correlation coefficients (R), coefficients of determination (R<sup>2</sup>), root mean square errors (RMSE), and mean squared errors (MSE)

Crop	R <sup>2</sup>	R	RMSE	MSE	Equation
Barley	0.811	0.901	0.01	0.061	Y=1.187x-0.266
wheat	0.640	0.8	0.276	0.076	Y=1.02x-0.003

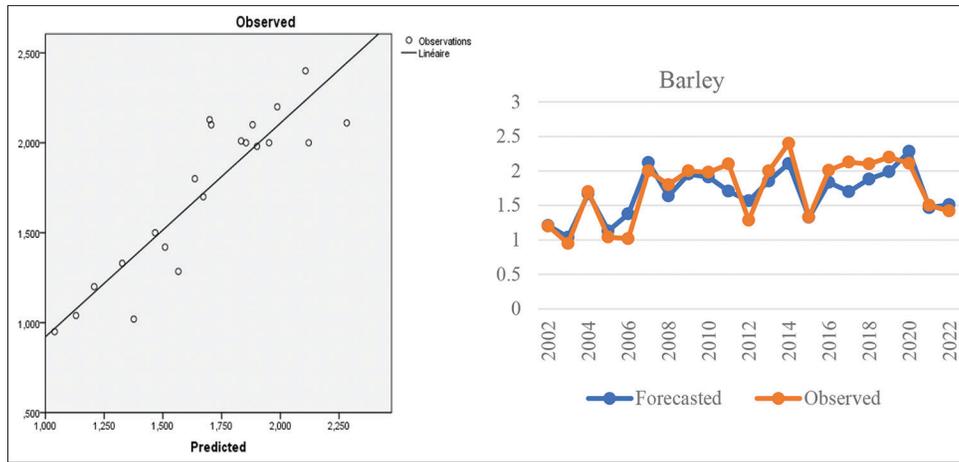


Figure 6: The scatter plot between observed and predicted values according to the created model for barley

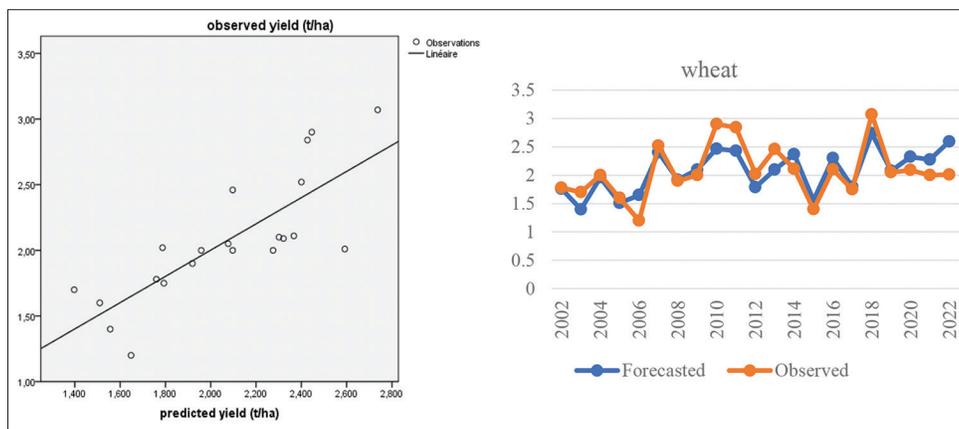


Figure 7: The scatter plot between observed and predicted values according to the created model for wheat

yield of winter wheat in selected regions of China. The study reported a determination coefficient of  $R^2 = 0.56$  between the estimated and measured winter wheat yield, which is lower than that found in our study ( $R^2 = 0.640$ ), a root mean square error (RMSE) of  $1.22 \text{ t ha}^{-1}$ , which is higher than that found in our work (RMSE =  $0.276 \text{ t ha}^{-1}$ ). Nagy *et al.* (2021), found a high regression coefficients between the vegetation indices and the wheat yield ( $R^2 = 0.757$ , RMSE =  $0.357 \text{ t ha}^{-1}$ ). The best time for wheat yield prediction with Landsat 8-NDVI was found to be the beginning of full biomass period from the 138<sup>th</sup> to 167<sup>th</sup> day after sowing (18 May to 16 June), which it is the same period that we found in our study. Gop and Savenkov (2016), found that The correlation was significant between the yield of spring wheat and the NDVI ( $R^2 = 0.859$ ). The study demonstrated that the NDVI was shown to be responsible for 85% of the variation in the yield of spring wheat. The approximate average increase in the grain yields of spring wheat was about  $6.7 \text{ t ha}^{-1}$ , with an increase of 0.1 in NDVI value. Tuğaç *et al.* (2022) found that the highest correlation between NDVI and yield was during the flowering period ( $R^2 = 0.63$ ). They also found that the best prediction performance was achieved with the MLP model for MODIS, with a root mean square error (RMSE) ranging from  $0.23\text{-}0.65 \text{ t ha}^{-1}$ . According to Mashaba *et al.* (2017), the relationship between NDVI and wheat yield was significant with an  $R^2$  value of 0.73 and RMSE of  $0.41 \text{ t ha}^{-1}$ . In

Latvia, Vannoppen *et al.* (2020) found that the linear regression model fit had a good estimate of the model parameter, with an adjusted  $R^2$  of 0.71. Pismennaya *et al.* (2021), investigated the correlation between MODIS-NDVI and winter wheat yield in the arid zone of the Central Pre-Caucasus region, using data from 2017-2020. Their findings revealed a very strong positive correlation ( $R^2 = 0.78$ ) between winter wheat yield and NDVI. Moreover, they reported an average increase of  $0.20 \text{ t ha}^{-1}$  in wheat grain yield for every 0.1 increase in NDVI value. In central Europe, Panek and Gozdowski (2020), found a strong relationship between cereal-grain yield and MODIS-NDVI in spring (April), three to four months before the harvest. The increase in the NDVI in early spring by 0.1 unit increases the grain yield of cereals by about 1.1 to  $2.6 \text{ t ha}^{-1}$ .

This fluctuation in results between different studies is due to That NDVI measures the potential yield and does not account for any subsequent crop developments that occur after the forecast date. Factors such as drought, diseases, or pest outbreaks occurring after the forecast date may lead to overestimations of crop yield. Additionally, satellite images are susceptible to various atmospheric effects, including clouds and volcanic eruptions, which can compromise data quality and subsequently affect the developed crop-yield models. Further research is necessary to validate the equations under different

weather scenarios and to enhance the relationship's strength by incorporating auxiliary data

## CONCLUSION

This study has successfully demonstrated the effective utilization of MODIS-NDVI for predicting cereal crop yield (wheat and barley) in the semi-arid regions of Algeria, providing reliable forecasts two to three months before harvest. A robust correlation between cereals-grain yield and NDVI was observed during early spring (specifically on March 13<sup>th</sup>). From the forecasting model that was developed based on twenty training years a 0.1 unit increase in mean NDVI during April corresponded to a cereals-grain yield increase ranging from 0.659 to 0.746 t ha<sup>-1</sup>. The root mean square error (RMSE) for the two crop cereals ranged from 0.01 t ha<sup>-1</sup> to 0.276 t ha<sup>-1</sup>. These findings highlight the utility of MODIS satellite data in enhancing the accuracy of regional-level cereal-grain yield prediction in Algeria, particularly during the early spring period. This enables improved planning of trade and food policies, which heavily rely on cereals-grain production.

## REFERENCES

- Adeniyi, O. D., Szabo, A., Tamás, J., & Nagy, A. (2020). Wheat Yield Forecasting Based on Landsat NDVI and SAVI Time Series. *Preprints, 2020*, 2020070065. <https://doi.org/10.20944/preprints202007.0065.v1>
- Adhab, M., & Alkuwaiti, N. A. (2022). Germiniviruses occurrence in the middle east and their impact on agriculture in Iraq. In R. K. Gaur, P. Sharma & H. Czosnek (Eds.), *Germinivirus: Detection, Diagnosis and Management* (pp. 171-185) Cambridge, US: Academic Press. <https://doi.org/10.1016/B978-0-323-90587-9.00021-3>
- Al-Ani, R. A., Adhab, M. A., El-Muadhidi, M. A., & Al-Fahad, M. A. (2011). Induced systemic resistance and promotion of wheat and barley plants growth by biotic and non-biotic agents against barley yellow dwarf virus. *African Journal of Biotechnology, 10*(56), 12078-12084.
- Algérie Eco. (2022). Cereals: Algeria imported 10.6 million tonnes during the 2021/2022 campaign. Retrieved from <https://www.algerie-eco.com/2022/07/13/cereales-lalgerie-a-importe-106-millions-de-tonnes-durant-la-campagne-2021-2021>
- Ammar, M. (2014). *Organisation de la chaîne logistique dans la filière céréales en Algérie: état des lieux et perspectives*. Master of Sciences, CIHEAM-AMM.
- Anaz, A., Kadhim, N., Sadoon, O., Alwan, G., & Adhab, M. (2023). Sustainable Utilization of Machine-Vision-Technique-Based Algorithm in Objective Evaluation of Confocal Microscope Images. *Sustainability, 15*(4), 3726. <https://doi.org/10.3390/su15043726>
- Atzberger, C., Vuolo, F., Klisch, A., Rembold, F., Meroni, M., Marcio, P. M., & Formaggio, A. (2016). Agriculture. In P. S. Thenkabail (Eds.), *Remote Sensing Handbook* (pp. 71-103) Florida, US: CRC Press.
- Becker-Reshef, I., Justice, C., Barker, B., Humber, M., Rembold, F., Bonifacio, R., Zappacosta, M., Budde, M., Magadzire, T., Shitot, e C., Pound, J., Constantino, A., Nakalembe, C., Mwangi, K., Sobue, S., Newby, T., Whitcraft, A., Jarvis, I., & Verdin, J. (2020). Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM Crop Monitor for Early Warning. *Remote Sensing of Environment, 237*, 111553. <https://doi.org/10.1016/j.rse.2019.111553>
- Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment, 114*(6), 1312-1323. <https://doi.org/10.1016/j.rse.2010.01.010>
- Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., & Peng, B. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology, 274*, 144-459. <https://doi.org/10.1016/j.agrformet.2019.03.010>
- Chennafi, H., Bouzerzour, H., Aidaoui, A., & Saci, A. (2006). Yield response of durum wheat (*Triticum durum* Desf.) cultivar Waha to deficit irrigation under semi arid growth conditions. *Asian Journal of Plant Sciences, 5*(5), 854-860. <https://doi.org/10.3923/ajps.2006.854.860>
- Climate Data. (2022). *Climate Setif (Algeria)*. Retrieved from <https://fr.climate-data.org/afrique/algerie/setif/setif-3595>
- Dorosh, P., & Salam, A. (2006). Wheat markets and price stabilization in Pakistan: An analysis of policy options. *The Pakistan Development Review, 47*(1), 71-87.
- Fang, H., Liang, S., Hoogenboom, G., Teasdale, J., & Cavigelli, M. (2008). Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model. *International Journal of Remote Sensing, 29*(10), 3011-3032. <https://doi.org/10.1080/01431160701408386>
- Fang, J., Piao, S., He, J., & Ma, W. (2004). Increasing terrestrial vegetation activity in China, 1982-1999. *Science in China Series C: Life Sciences, 47*, 229-240. <https://doi.org/10.1007/BF03182768>
- FAO. (2018). *The State of Food Security and Nutrition in the World 2018: Building climate resilience for food security and nutrition*. Rome, Italy: FAO.
- Geng, L., Mengdi, L., Zhang, G., & Ye, L. (2022). Barley: a potential cereal for producing healthy and functional foods. *Food Quality and Safety, 6*, fyac012. <https://doi.org/10.1093/fqsafe/fyac012>
- GIMMS. (2021). *Global Agricultural Monitoring System*. Retrieved January 12, 2021 from <https://glam1.gsfc.nasa.gov>
- Google Earth Pro. (2023). Google Earth Pro [Computer Software]. <https://earth.google.com/intl/earth/versions/#download-pro>
- Gop, N. V., & Savenkov, O. A. (2019). Relationships between the NDVI, Yield of Spring Wheat, and Properties of the Plow Horizon of Eluviated Clay-Illuvial Chernozems and Dark Gray Soils. *Eurasian Soil Science, 52*, 339-347. <https://doi.org/10.1134/S1064229319030050>
- Huang, J., Wang, X., Li, X., Tian, H., & Pan, Z. (2013). Remotely sensed rice yield prediction using multi-temporal NDVI data derived from NOAA's AVHRR. *PLoS One, 8*(8), e70816. <https://doi.org/10.1371/journal.pone.0070816>
- Khalaf, L. K., Adhab, M., Aguirre-Rojas, L. M., & Timm, A. E. (2023). Occurrences of wheat curl mite aceria tosicHELLa keifer 1969 (eriphyidae) and the associated viruses, (WSMV, HPWMOV, TriMV) in IRAQ. *Iraqi Journal of Agricultural Sciences, 54*(3), 837-849. <https://doi.org/10.36103/ijas.v54i3.1767>
- Khalaf, L., Chuang, W.-P., Aguirre-Rojas, L. M., Klein, P., & Smith, C. M. (2019). Differences in *Aceria tosicHELLa* population responses to wheat resistance genes and wheat virus transmission. *Arthropod-Plant Interactions, 13*, 807-818. <https://doi.org/10.1007/s11829-019-09717-9>
- Kouadio, L., Newlands, N. K., Davidson, A., Zhang, Y., & Chipanshi, A. (2014). Assessing the performance of MODIS NDVI and EVI for seasonal crop yield forecasting at the ecodistrict scale. *Remote Sensing, 6*(10), 10193-10214. <https://doi.org/10.3390/rs61010193>
- Li, C., Qi, J., Yang, L., Wang, S., Yang, W., Zhu, G., Zou, S., & Zhang, F. (2014). Regional vegetation dynamics and its response to climate change—a case study in the Tao River Basin in Northwestern China. *Environmental Research Letters, 9*(12), 125003. <https://doi.org/10.1088/1748-9326/9/12/125003>
- Liu, H., Zhang, X., Xu, Y., Ma, F., Zhang, J., Cao, Y., Li, L., & An, D. (2020). Identification and validation of quantitative trait loci for kernel traits in common wheat (*Triticum aestivum* L.). *BMC Plant Biology, 20*, 529. <https://doi.org/10.1186/s12870-020-02661-4>
- Lykhovyd, P. V. (2020). Sweet corn yield simulation using normalized difference vegetation index and leaf area index. *Journal of Ecological Engineering, 21*(3), 228-236. <https://doi.org/10.12911/22998993/118274>
- Mashaba, Z., Chirima, G., Botai, J. O., Combrinck, L., Munghezulu C., & Dube, E. (2017). Forecasting winter wheat yields using MODIS NDVI data for the Central Free State region. *South African Journal of Science, 113*(11/12), 1-6. <https://doi.org/10.17159/sajs.2017/20160201>
- Mekhlouf, A., Dehbi, F., Hanachi, A., & Harbi, M. (2012). Réponses de blé dur aux basses températures en relation avec la capacité de production. *Agriculture, 3*(1), 13-23.
- Mkhabela, M., Bullock, P., Gervais, M., Finlay, G., & Sapirstein, H. (2010). Assessing indicators of agricultural drought impacts on spring wheat yield and quality on the Canadian Prairies. *Agricultural and Forest Meteorology, 150*(3), 399-410. <https://doi.org/10.1016/j.agrformet.2010.01.001>
- Mulianga, B., Bégué, A., Simoes, M., & Todoroff, P. (2013). Forecasting regional sugarcane yield based on time integral and spatial

- aggregation of MODIS NDVI. *Remote Sensing*, 5(5), 2184-2199. <https://doi.org/10.3390/rs5052184>
- Nagy, A., Fehér, J., & Tamás, J. (2018). Wheat and maize yield forecasting for the Tisza river catchment using MODIS NDVI time series and reported crop statistics. *Computers and Electronics in Agriculture*, 151, 41-49. <https://doi.org/10.1016/j.compag.2018.05.035>
- Nagy, A., Szabó, A., Adeniyi, O. D., & Tamás J. (2021). Wheat Yield Forecasting for the Tisza River Catchment Using Landsat 8 NDVI and SAVI Time Series and Reported Crop Statistics. *Agronomy*, 11(4), 652. <https://doi.org/10.3390/agronomy11040652>
- Panek, E., & Gozdowski, D. (2021). Relationship between MODIS Derived NDVI and Yield of Cereals for Selected European Countries. *Agronomy*, 11(2), 340. <https://doi.org/10.3390/agronomy11020340>
- Phiri, D., Simwanda, M., Salekin, S., Nyirenda V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: A review. *Remote Sensing*, 12(14), 2291. <https://doi.org/10.3390/rs12142291>
- Pismennaya, E. V., Azarova1, M. Y., Golosnoy, E.V., Odintsov, S.V., & Kipa, L.V. (2021). Relationship between NDVI index obtained from MODIS and winter wheat yield. *Earth and Environmental Science*, 848, 012110. <https://doi.org/10.1088/1755-1315/848/1/012110>
- Rouabhi, A. (2017). Spatiotemporal characterization of the annual Rainfall in Setif region-Algeria. *Revue Agriculture*, 4(1), 31-38.
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, 351, 309-317.
- Sabins, F. F. Jr. (1987). Remote sensing – principles and interpretation. (2<sup>nd</sup> ed.). New York, UK: W. H. Freeman and Company.
- Tuğaç, M. G., Özbayoğlu, A. M., Torunlar, H., & Karakurt, E. (2022). Wheat Yield Prediction with Machine Learning based on MODIS and Landsat NDVI Data at Field Scale. *International Journal of Environment and Geoinformatics*, 9(4), 172-184. <https://doi.org/10.30897/ijegeo.1128985>
- USGS. (2021). Global Croplands. Retrieved January 12, 2021 from <https://croplands.org/app/map?lat=0&lng=0&zoom=2>
- Vannoppen, A., Gobin, A., Kotova, L., Top, S., Cruz, L., Viksna, A., Aniskevich, S., Bobylev, L., Buntemeyer, L., Caluwaerts, S., Troch, R. D., Gnatiuk, N., Hamdi, R., Remedio, A. R., Sakalli, A., Vyver, H. V. D., Schaeybroeck, B. V., & Termonia, P. (2020). Wheat Yield Estimation from NDVI and Regional Climate Models in Latvia. *Remote Sensing*, 12(14), 2206. <https://doi.org/10.3390/rs12142206>
- Vozhehova, R., Maliarchuk, M., Biliaieva, I., Lykhovyd, P. V., Maliarchuk, A., & Tomnytskyi, A., (2020). Spring row crops productivity prediction using normalized difference vegetation index. *Journal of Ecological Engineering*, 21(6), 176-182. <https://doi.org/10.12911/22998993/123473>
- Wang, Y., Xu, X., Huang, L., Yang, G., Fan, L., Wei, P., & Chen, G. (2019). An Improved CASA Model for Estimating winter wheat yield from remote sensing images. *Remote Sensing*, 11(9), 1088. <https://doi.org/10.3390/rs11091088>
- Yin, Z., & Williams, T. H. L. (1997). Obtaining spatial and temporal vegetation data from Landsat MSS and AVHRR/NOAA satellite mages for a hydrologic model, *Photogrammetric Engineering and Remote Sensing*, 63(1), 69-77.
- Zhang, P-P, Zhou, X.-X., Wang, Z.-X., Mao, W., Li, W.-X., Yun, F., Guo, W.-S., & Tan, C.-W. (2020). Using HJ-CCD image and PLS algorithm to estimate the yield of field-grown winter wheat. *Scientific Reports*, 10, 5173. <https://doi.org/10.1038/s41598-020-62125-5>