

Université de Toulouse

MASTER 2 GÉOMATIQUE

« **Sciences Géomatiques en environneMent et Aménagement** » (**SIGMA**)

<http://sigma.univ-toulouse.fr>

MÉMOIRE DE RECHERCHE

**Cartographie du maïs irrigué et non irrigué dans le Loiret
à partir d'images Sentinel-2**

DAHINGER Virginie



**Centre d'Etudes Spatiales
de la BIOSphère**



Maître de stage : DEMAREZ Valérie

Tuteur-enseignant : SHEEREN David

Septembre 2019

Résumé

En tant que première puissance agricole (en valeur) et important producteur de maïs au niveau européen, la France tient une place centrale sur le marché agricole. Pour estimer les productions de maïs, les services déconcentrés de l'État distinguent le maïs irrigué du non irrigué à l'échelle départementale. Ce mémoire de Master étudie le potentiel des images optiques (Sentinel-2) pour la détection automatique des pratiques culturales (irrigué/non irrigué) pour une même culture (maïs) en cours et en fin de saison. L'étude se concentre sur le département du Loiret avec des données collectées en 2017 et 2018. La classification supervisée des pixels de maïs est réalisée par la chaîne de traitement *iota*² selon l'algorithme *Random Forest* à partir d'informations *in-situ* utilisées pour l'apprentissage du modèle. L'originalité de la méthode provient de l'utilisation d'indices spectraux (NDVI, NDWI, NDRE), traités en cumuls saisonniers et en cumuls mensuels, pour extraire des séries temporelles les informations les plus pertinentes et limiter le volume de données à traiter. En 2018, l'analyse en fin de saison (mai à octobre inclus) permet d'obtenir 89% (*Overall Accuracy*) des pixels de maïs correctement classés. L'analyse en cours de saison présente des résultats satisfaisants à partir de la fin du mois de juillet (87%). En 2017, les résultats peu concluants (63%) obtenus en fin de saison peuvent s'expliquer par les fortes précipitations durant l'été et par la nébulosité. La combinaison de données optiques et radar (Sentinel-1) est une piste d'amélioration pour pallier ce manque de données.

Mots-clés : *cultures irriguées ; cartographie des cultures en cours de saison ; indices spectraux ; séries temporelles d'images satellitaires ; Sentinel-2 ; Random Forest.*

Abstract

As the leading agricultural (in value) power and important maize producer in Europe, France holds a central position in the agricultural market. In order to estimate maize production, the decentralized government departments distinguish irrigated and non-irrigated maize at the departmental level. This Master's thesis examines the potential of optical images (Sentinel-2) for the automatic detection of cultural practices (irrigated / non-irrigated) for the same crop (corn) in-season and at the end of the season. The study focuses on the Loiret department with data collected in 2017 and 2018. The supervised classification of maize pixels is performed by the *iota*² processing chain according with the *Random Forest* algorithm using *in-situ* informations for learning the model. The originality of the method comes from the use of spectral indices (NDVI, NDWI, NDRE), processed in seasonal and monthly cumulations, to extract the relevant informations from the time series and to limit the volume of processed data. In 2018, the analysis at the end of the season (May to October included) allows to obtain 89% (*Overall Accuracy*) correctly classified corn pixels. The in-season analysis shows satisfactory results from the end of July (87%). In 2017, the unsatisfying results (63%) obtained at the end of the season can be explained by the heavy precipitation of the summer and by cloudiness. The combination of optical and radar data (Sentinel-1) is an improvement path to overcome this lack of data.

Keywords: *irrigated crops; in-season crop mapping; spectral indices; satellite image time series; Sentinel-2; Random Forest.*

Remerciements

Je remercie particulièrement ma maîtresse de stage Valérie Demarez pour sa confiance, son suivi et ses précieuses relectures, qui m'ont permis de mener à bien ce stage. Je remercie également Yann Pageot pour ses conseils techniques et son aide informatique.

Merci au CNES et à son programme TOSCA qui a financé mon stage dans le cadre d'appel aux laboratoires scientifiques pour mener des recherches.

Ce travail est aussi le fruit d'une collaboration avec la DRAAF Centre-Val de Loire qui a sollicité le CESBIO et mis à disposition les données terrain, indispensables à la construction du modèle. Merci à Audrey Oddos et Gérard Guillaume pour l'intérêt porté à mon travail et leur disponibilité pour répondre à mes questions.

Par ailleurs, je tiens à remercier Anne-Laure Gibelin et Franck Souverain pour les données météorologiques qui m'ont permis de mieux appréhender le territoire et d'analyser les résultats.

Table des matières

Introduction.....	1
Partie 1 : Application des méthodes de télédétection pour l'estimation des surfaces irriguées sur le département du Loiret	5
Section I - Contexte et présentation du stage.....	5
A. Présentation de la structure d'accueil.....	5
B. Présentation de la structure partenaire	7
C. Contexte et objectifs du stage.....	8
Section II - Présentation du territoire d'étude : le Loiret	8
A. Caractéristiques géographiques	8
B. Caractéristiques climatiques	10
C. Pratiques agricoles et irrigation.....	12
Section III - Utilisation des images satellitaires en agriculture.....	15
A. Les données de télédétection et l'agriculture.....	15
B. Les opportunités associées au programme <i>Copernicus</i>	15
C. Les méthodes de classification	16
D. Association des données optiques et radar	17
Partie 2 : Classification et cartographie des parcelles de maïs irrigué et non irrigué	19
Section I - Description des données	19
A. Données satellitaires	19
B. Données de référence	22
C. Données sur la composition du sol.....	22
Section II - Présentation des outils.....	23
A. La chaîne de traitement iota^2	23
B. Le serveur	26
Section III - Détail de la méthode	27
A. Échantillonnage et validation	27
B. Interpolation linéaire	28
C. Période d'observation	28
D. Indices spectraux et apport de la temporalité	29
E. Normalisation des données GlobalSoilMap	38
F. Configuration des classifications	38
G. Évaluation des performances des classifications	38
H. Simplification des résultats.....	39
Section IV - Résultats & Analyse	40
A. Evaluation de la méthode d'échantillonnage et de la période d'observation	40
B. Evaluation du classifieur	41

C. Apport des cumuls saisonniers et mensuels	42
D. Apport des données GlobalSoilMap.....	43
E. Apport des séries temporelles.....	44
F. Traitement en cours de saison	45
G. Simplification et validation des résultats	46
Conclusion & Perspectives	49
Bibliographie.....	51
Annexes	53

Table des figures

Figure 1 : Localisation du territoire d'étude.....	8
Figure 2 : Carte du relief du département du Loiret	9
Figure 3 : Carte d'occupation du sol du Loiret en 2018.....	9
Figure 4 : Diagramme ombrothermique du Loiret en 2017	11
Figure 5 : Diagramme ombrothermique du Loiret en 2018	11
Figure 6 : Cartes des arrêtés de restriction d'eau du Loiret en 2017 et 2018 (www.propluvia.developpement-durable.gouv.fr).....	12
Figure 7 : Volumes d'eau prélevés pour l'irrigation en 2016 sur le Loiret	13
Figure 8 : Part de la SAU irriguée par rapport à la SAU en région Centre (Agreste Recensement agricole 2010).....	14
Figure 9 : Evolution des superficies irriguées par culture en nappe de Beauce (Agreste - Recensement agricole 2010)	14
Figure 10 : Différences entres les classifieurs implémentés par la bibliothèque d'apprentissage statistique <i>Scikit-learn</i>	16
Figure 11 : Résolution spatiale par longueurs d'ondes capturées par Sentinel-2 (ESA)	19
Figure 12 : Carte des emprises des tuiles Sentinel-2 sur le territoire d'étude	20
Figure 13 : Images Sentinel-2 disponibles en 2017 et 2018 pour couvrir le territoire d'étude	20
Figure 14 : Carte de la répartition des données de référence en 2017 et 2018	22
Figure 15 : Carte GSM de prédiction de la texture du sol (g/kg) (Mulder et al. 2016).....	23
Figure 16 : Etapes de traitement $iota^2$ (Inglada, 2017)	25
Figure 17 : Profils temporels des NDVI moyens des parcelles de maïs issues des données de référence .	29
Figure 18 : Séries temporelles des NDVI du maïs irrigué et non irrigué en 2017 et 2018	31
Figure 19 : Séries temporelles des NDWI du maïs irrigué et non irrigué en 2017 et 2018.....	32
Figure 20 : Séries temporelles des indices de brillance du maïs irrigué et non irrigué en 2017 et 2018....	32
Figure 21 : Séries temporelles des NDRE-1 du maïs irrigué et non irrigué en 2017 et 2018	33
Figure 22 : Séries temporelles des NDRE-2 du maïs irrigué et non irrigué en 2017 et 2018	33
Figure 23 : Séries temporelles des NDRE-3 du maïs irrigué et non irrigué en 2017 et 2018	34
Figure 24 : Séries temporelles des cumuls de NDVI du maïs irrigué et non irrigué en 2017 et 2018	35
Figure 25 : Séries temporelles des cumuls de NDWI du maïs irrigué et non irrigué en 2017 et 2018.....	35
Figure 26 : Séries temporelles des cumuls d'indice de brillance du maïs irrigué et non irrigué en 2017 et 2018.....	35
Figure 27 : Séries temporelles des cumuls de NDRE-1 du maïs irrigué et non irrigué en 2017 et 2018.....	36
Figure 28 : Séries temporelles des cumuls de NDRE-2 du maïs irrigué et non irrigué en 2017 et 2018.....	36
Figure 29 : Séries temporelles des cumuls de NDRE-3 du maïs irrigué et non irrigué en 2017 et 2018.....	36
Figure 30 : Carte de comparaison du cumul de NDVI traité en cumul saisonnier et en cumul mensuel en 2018.....	37
Figure 31 : Comparaison des scores des classifications en 2017 et 2018 selon la méthode d'échantillonnage et la période d'observation.....	41
Figure 32 : Comparaison des scores des classifications en 2018 selon les cumuls saisonniers des indices spectraux	42
Figure 33 : Temps de traitement pour une classification avec un cumul saisonnier	42
Figure 34 : Comparaison des scores des classifications en 2018 selon les cumuls mensuels des indices spectraux	43
Figure 35 : Comparaison des scores des classifications en 2018 avec ou sans les données GSM	43
Figure 36 : Comparaison des scores des classifications en 2018 avec ou sans la série temporelle.....	44
Figure 37 : Temps de traitement pour une classification avec une série temporelle de 6 mois	44
Figure 38 : Comparaison mois par mois des scores des classifications en 2018.....	46
Figure 39 : Cartographie du maïs irrigué et non irrigué du Loiret en 2018 (configuration "stack")	48

Table des tableaux

Tableau 1 : Nébulosité en 2017 et 2018.....	21
Tableau 2 : Synthèse des tables attributaires des données de référence en 2017 et 2018	27
Tableau 3 : Matrice de confusion de la configuration "stack" avec la série temporelle.....	45
Tableau 4 : Matrice de confusion de la configuration "indexM" avec tous les indices traités en cumuls mensuels.....	45
Tableau 5 : Statistiques agricoles sur les productions du maïs en 2018 (Agreste)	47
Tableau 6 : Surfaces obtenues par culture et validation des résultats	47

INTRODUCTION

Les scénarii climatiques du GIEC (Groupement d'experts intergouvernemental sur l'évolution du climat) prévoient dans de nombreuses régions du globe une raréfaction de la ressource en eau due à l'augmentation des épisodes de sécheresse. Les cinq grands risques identifiés sont relatifs à l'eau, à l'agriculture et aux moyens de subsistance (GIEC, 2015). En France, selon le rapport de 2018 du Conseil Général de l'Alimentation, de l'Agriculture et des Espaces Ruraux, seize départements prennent systématiquement des arrêtés de restrictions en été et plus de la moitié ont dû s'y résoudre en 2003, 2005, 2006 et 2011. L'impact de cette raréfaction est accru dans de nombreuses régions par l'intensification des prélèvements en eau pour différents usages dont l'agriculture.

L'agriculture n'est pas possible sans eau et le changement climatique impacte directement le cycle de l'eau et l'agriculture. Les sécheresses et les canicules affectent également la vie biologique des sols, la biodiversité et toutes les fonctions liées qui ont également des répercussions sur l'agriculture. La pratique de l'irrigation peut avoir plusieurs objectifs. Selon les zones géographiques, l'irrigation peut viser l'augmentation des rendements, leur sécurisation ou encore l'amélioration de la qualité des produits. L'eau d'irrigation est ainsi considérée comme un intrant indispensable à la production agricole, au même titre que les fertilisants par exemple.

En tant que première puissance agricole européenne et sixième au plan mondial, la France a comme priorité de valoriser ses productions agricoles pour le développement des filières françaises en vue de créer de la valeur, de la richesse et des emplois. Avec une production estimée par le Ministère de l'Agriculture à 73 milliards d'euros en 2018 et 448 500 agriculteurs en constante diminution, la dimension économique et la compétitivité des productions agricoles à l'international est un enjeu majeur. Selon Jean-François Lépy, directeur de *Soufflet Négoce*, société de transport et d'exportation de céréales, la compétitivité est la principale porte d'entrée vers les marchés à l'export, *"c'est le socle de l'économie des fermes céréalières françaises. Si nous ne sommes pas compétitifs, ce socle commence à se fissurer"* et le marché national en pâtira, déclare-t-il en janvier

2019. Ces forts enjeux économiques mobilisent les services déconcentrés du Ministère de l'Agriculture et de l'Alimentation pour collecter continuellement les données agricoles. Ces informations permettent ensuite de suivre et d'expliquer les évolutions de production, de consommation et de prix dans les différentes filières agricoles, c'est ce qu'on appelle la conjoncture agricole.

L'une des cultures particulièrement suivie est le maïs. L'une de ses particularités est l'efficacité de l'utilisation de l'eau, définie par le rendement par unité d'eau. Selon les régions, l'irrigation du maïs permet de gagner 30 à 50 quintaux par hectare (www.irrifrance.com). La France est le premier producteur de maïs en Europe et le maïs est la deuxième production végétale en France après le blé, elle occupe ainsi 9 % de la surface agricole utile (SAU). La région Centre-Val de Loire, surnommée historiquement « le grenier à grains » de la France, produit 1,5 Mt de maïs par an. Parallèlement, sur le site de l'Agreste, dédié à la publication des statistiques agricoles, la région est décrite comme la première région française quant à la superficie irriguée et arrive en quatrième position pour le volume d'eau consommé pour l'irrigation en 2012. Ainsi, la surface irrigable du Centre-Val de Loire représente plus de 20 % de la surface agricole utilisée et 60 % de la SAU irrigable sont consacrés au maïs irrigué. La même année, les surfaces irriguées sont également nettement supérieures dans le Loiret comparées aux autres départements. La part du maïs irrigué ne cesse d'augmenter, en 2018 elle représente 80 % du maïs grain produit sur le département (Agreste).

Afin de réaliser la conjoncture agricole du maïs à l'échelle du département du Loiret pour le compte du Ministère, la Direction régionale de l'Alimentation, de l'Agriculture et de la Forêt (DRAAF) Centre-Val de Loire distingue le maïs irrigué du maïs non-irrigué. Un rendement moyen appliqué en fonction de la pratique agricole permet ensuite d'estimer la production totale de maïs du département. Dans ce contexte, la DRAAF Centre-Val de Loire a sollicité le Centre d'Etudes Spatiales de la Biosphère (CESBIO) pour réaliser une cartographie du maïs en fin et en cours de saison. La cohérence des résultats est évaluée en comparant les surfaces obtenues pour chaque culture avec les données de l'Agreste. Le cas échéant, l'étude justifiera de compléter les méthodes d'estimation de la DRAAF par des techniques de télédétection.

En tant que laboratoire de recherche spécialisé dans l'étude du fonctionnement des surfaces continentales avec comme outil privilégié la télédétection, le CESBIO est impliqué dans plusieurs programmes de recherche sur la cartographie de l'occupation du sol et des surfaces irriguées. L'utilisation des techniques de classification des images de télédétection, optiques et radar, basées sur les spécificités des réponses spectrales des surfaces dans les différentes longueurs d'onde est une méthode efficace et aujourd'hui répandue. Le programme européen de monitoring *Copernicus* avec les constellations de satellites *Sentinel*, dédiées à l'observation et la surveillance de la Terre, offre de nouvelles opportunités grâce aux images prétraitées, à haute résolution, spatiale et temporelle, disponibles en libre accès.

L'objectif du stage consiste à faire la distinction entre des pratiques culturales (irrigué / non irrigué) pour une même culture (maïs) dans une zone tempérée (Loiret). Ce travail s'inscrit dans la continuité des travaux menés en Occitanie sur la cartographie du maïs irrigué et non irrigué où l'utilisation d'images optiques (Landsat-8) associées aux images radar (Sentinel-1) améliore la classification de l'irrigué (Demarez et al., 2019). Pour atteindre cet objectif, la méthode consiste à évaluer le potentiel des images Sentinel-2 pour distinguer les différences entre les deux pratiques. Ce choix s'appuie sur l'hypothèse selon laquelle la croissance du maïs est liée à son apport en eau. Les contrastes moins nets observés entre les pratiques en zone tempérée qu'en zone semi-aride rendent la distinction moins évidente. La durée du stage ne permettant pas d'investiguer l'apport des données Sentinel-1, une attention particulière est portée sur l'apport de primitives pour améliorer la distinction.

Pour répondre à cette problématique, une première partie est consacrée à la contextualisation de l'étude et à l'apport général de la télédétection en agriculture. Puis, une seconde partie plus spécifique présentera l'étude visant à cartographier les parcelles de maïs irrigué et non-irrigué du département du Loiret en 2017 et 2018.

PARTIE 1 : APPLICATION DES METHODES DE TELEDETECTION POUR L'ESTIMATION DES SURFACES IRRIGUEES SUR LE DEPARTEMENT DU LOIRET

La première partie du mémoire permet de comprendre les enjeux liés au sujet de recherche et à la réflexion autour de la problématique. Ainsi, le développement proposé porte premièrement sur le **contexte et la présentation du stage** (Section I), deuxièmement sur la **présentation du territoire d'étude** (Section II), et troisièmement sur l'**utilisation des images satellitaires en agriculture** (Section III).

Section I - CONTEXTE ET PRESENTATION DU STAGE

Le travail effectué lors du stage de six mois est le résultat d'une collaboration entre un laboratoire scientifique, la structure d'accueil, et un partenaire institutionnel. Dans ces conditions, les objectifs du stage sont définis pour satisfaire les deux parties.

- *Présentation de la structure d'accueil*
- *Présentation de la structure partenaire*
- *Contexte et objectifs du stage*

A. PRESENTATION DE LA STRUCTURE D'ACCUEIL

1. LE CESBIO UNE UNITE MIXTE DE RECHERCHE

Le Centre d'Etudes Spatiales de la BIOSphère (CESBIO) est une Unité Mixte de Recherche issue de l'association de plusieurs centres de recherche. Le CESBIO est sous la tutelle :

- du Centre National d'Etudes Spatiales (CNES),
- du Centre National de la Recherche Scientifique (CNRS),
- de l'Institut de Recherche pour le Développement (IRD),
- de l'Institut National de la Recherche Agronomique (INRA),
- et de l'Université Paul Sabatier (Toulouse III) UPS.

Sous la Direction de Laurent Polidori depuis 2016, le CESBIO regroupe environ 130 personnes dont 65 permanents (enseignants-chercheurs, ingénieurs et techniciens de laboratoire) qui étudient la dynamique et le fonctionnement de la biosphère continentale et ses interactions avec le climat et les activités humaines.

Les travaux menés par le CESBIO s'intègrent plus largement dans le pôle "Univers, Planète, Espace, Environnement" de l'Université Toulouse III - Paul Sabatier, organisé autour de l'Observatoire Midi



Pyrénées (OMP). L'OMP fédère ainsi les laboratoires des sciences de l'univers, de la planète et de l'environnement, autour des missions de recherche, d'observation, d'enseignement et de coopération internationale entre les différents Observatoires des Sciences de l'Univers (OSU). C'est donc naturellement que l'implantation principale du CESBIO se situe à Toulouse sur le campus de l'OMP, au 13 avenue du colonel Roche. Une implantation secondaire se trouve également à Auch, dans les locaux de l'IUT.

2. LES MISSIONS DU CESBIO

Le CESBIO a pour mission de contribuer au progrès des connaissances sur le fonctionnement des surfaces continentales et leurs interactions avec l'Homme, en s'appuyant sur des données satellitaires. Pour cela, plusieurs domaines de recherche sont mobilisés comme la physique, les sciences de la vie ou encore l'écologie, permettant ainsi une articulation entre le domaine spatial et l'étude de l'environnement.

Pour répondre à ces enjeux, les missions du CESBIO se structurent selon trois activités principales : l'observation et la mesure, la modélisation et enfin la valorisation.

Observation et mesure : contribution du CESBIO au développement et à l'utilisation des systèmes spatiaux d'observation de la biosphère. Conception de nouvelles missions spatiales, traitement d'images et l'élaboration de produits thématiques avec une spécialité marquée autour des séries temporelles.

Modélisation : développement de modèles explicatifs ou prévisionnels du fonctionnement des surfaces continentales sous la pression de changements climatiques et anthropiques à différentes échelles spatiales et temporelles.

Valorisation : liens avec les activités de formation, développement de partenariats avec les acteurs de la gestion des territoires et collaboration avec les industriels.

3. LE PROJET TOSCA PARCELLE

Le stage s'inscrit dans le programme TOSCA (Terre, Océanographie, Surface Continentale et Atmosphère), financé par le CNES, qui soutient de nombreux projets portés par des laboratoires scientifiques.

Le projet TOSCA PARCELLE qui en dépend, a pour objectif d'encourager les efforts de mutualisation et de capitalisation sur la chaîne de traitement *iota*² développée pour analyser automatiquement l'occupation des sols. Cet outil, élaboré par le Centre d'Expertise Scientifique (CES) « Occupations des Sols » (OSO), permet de lever des problèmes techniques liées au traitement et à l'analyse des données par un grand nombre d'images de télédétection (e.g. : une tuile S2 peut représenter sur une année jusqu'à 500 Gigaoctets). C'est la raison pour laquelle nous avons choisi d'utiliser cet outil dans mon stage.

Ce stage participe également aux activités menées au sein des CES « OSO » et « Irrigation » du Pôle Theia animés respectivement par J.Inglada et V.Demarez. Le pôle Theia a été créé en 2012 par neuf institutions publiques françaises, impliquées dans l'observation de la terre et des sciences de l'environnement. Cette

organisation a pour objectif de mettre à disposition de la communauté scientifique et des politiques publiques une infrastructure de données spatiales pour offrir un ensemble de prestations techniques.

B. PRESENTATION DE LA STRUCTURE PARTENAIRE

1. LES MISSIONS GENERALES DE LA DRAAF

Le processus de décentralisation de l'État a conduit à transférer une partie des compétences relatives à la mise en œuvre et au suivi des politiques nationales et communautaires vers des entités locales. Placées sous l'autorité du Préfet de région, les Directions Régionales de l'Alimentation, de l'Agriculture et de la Forêt, communément appelées DRAAF (ou DAAF en outre-mer) sont de cette manière des services déconcentrés du Ministère de l'Agriculture et de l'Alimentation. Les DRAAF accompagnent le développement rural et l'aménagement des territoires.

Les DRAAF assurent un rôle de coordination, de programmation, de contrôle et de circulation de l'information dans plusieurs domaines liés à l'agriculture et l'alimentation comme la politique sociale agricole, les statistiques agricoles, l'élevage, la filière bois, la protection des végétaux ou encore l'enseignement agricole.

2. SUIVI DES PRODUCTIONS AGRICOLES

Dans le cadre de ces missions de statistiques agricoles, le Service Régional de l'Information Statistique et Economique (SRISE) de la DRAAF Centre-Val de Loire réalise le suivi de la production de maïs. Le SRISE, est en charge de la conjoncture végétale régionale pour le compte du Service de la Statistique et de la Prospective qui correspond au service statistique du Ministère de l'Agriculture.

Les SRISE réalisent des enquêtes pour établir des prévisions puis des estimations des surfaces, des rendements et des productions des principales grandes cultures, tout au long de la campagne agricole (*Situation mensuelle grandes cultures, voire Annexe n°1*). Les estimations sont établies pour chaque département, puis agrégées par le Service Central pour obtenir une estimation nationale. Le suivi par les SRISE garantit l'impartialité dans la construction de ces données.

Une fois agrégée au niveau national, ces estimations sont publiées sur le site Agreste (« Infos Rapides Grandes Cultures ») et accompagnées des données de conjoncture (céréales, oléagineux, protéagineux, betteraves, pommes de terre, maïs fourrage et surfaces de jachères agronomiques). Le maïs est la seule culture qui fait l'objet d'une distinction entre le maïs irrigué et le maïs non-irrigué, car les rendements et donc la production varient sensiblement. Les premières estimations de surfaces sont faites après les semis d'hiver (estimations de décembre et février) et pratiquement en même temps que les semis de printemps (avril et mai), puis perfectionnées les mois suivants. Les premières estimations de rendements sont faites un à deux mois avant la récolte pour les céréales à paille et le colza, deux à trois mois pour les cultures d'été (maïs, betteraves, tournesol). Chaque mois, les estimations sont mises à jour en fonction des informations reçues.

C. CONTEXTE ET OBJECTIFS DU STAGE

Le stage de six mois réalisé au CESBIO est le résultat d'un partenariat entre le laboratoire et la DRAAF Centre-Val de Loire qui met à notre disposition des données de référence (vérités terrain) concernant le caractère irrigué ou non des parcelles de maïs.

Cette collaboration a deux principaux objectifs, le premier est d'accompagner la DRAAF Centre-Val de Loire dans l'évaluation des superficies de maïs irrigué et de maïs non-irrigué du département du Loiret à partir des données de télédétection. Les résultats obtenus au cours du stage s'intégreront dans la thèse de Yann Pageot sur l'apport de Sentinel-1 et Sentinel-2 pour la détection et le suivi des cultures irriguées sur le bassin de l'Adour-Garonne en cours et en fin de saison. La thèse et le stage sont dirigés par V.Demarez, Maître de conférences à l'Université Toulouse III - Paul Sabatier.

Pour atteindre ces objectifs, l'utilisation des images satellitaires à haute résolution, spatiale et temporelle, Sentinel-2 permettront de réaliser une cartographie des surfaces de maïs irrigué et non-irrigué pour les années étudiées (2017 et 2018). Le stage a pour ambition d'établir une méthode de cartographie robuste, transposable et transférable. C'est la raison pour laquelle notre choix s'est porté sur l'utilisation de la chaîne de traitement *iota*² (*Diagramme de Gantt, voire annexe n°2*).

Section II - PRESENTATION DU TERRITOIRE D'ETUDE : LE LOIRET

Le territoire d'étude s'étend sur l'intégralité du département du Loiret qui appartient à la région Centre-Val de Loire (**Figure 1**). Afin de réaliser une cartographie des parcelles irriguées du Loiret, une analyse des spécificités du territoire comme les caractéristiques géographiques, le climat ou les pratiques d'irrigation, permet d'établir des hypothèses et de faciliter la compréhension des résultats.

- *Caractéristiques géographiques*
- *Caractéristiques climatiques*
- *Pratiques agricoles et d'irrigation*

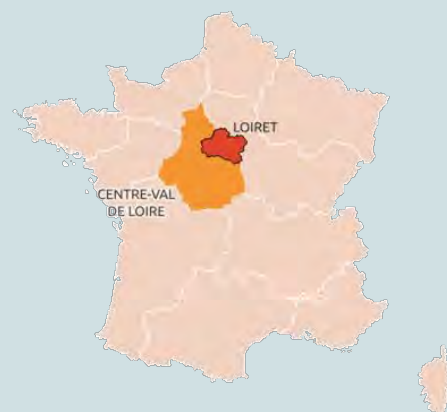


Figure 1 : Localisation du territoire d'étude

A. CARACTERISTIQUES GEOGRAPHIQUES

Situé au sud du bassin parisien, le Loiret est un département de la région Centre-Val de Loire qui est traversé par la Loire qui lui confère son nom. Le Loiret s'étend sur près de 7 000 km² avec une population de plus de 650 000 habitants en 2016. Avec une altitude voisine de 100 mètres, une minimale de 67 m et une maximale de 281 m (IGN, 2014), le Loiret est une vaste plaine (**Figure 2**).

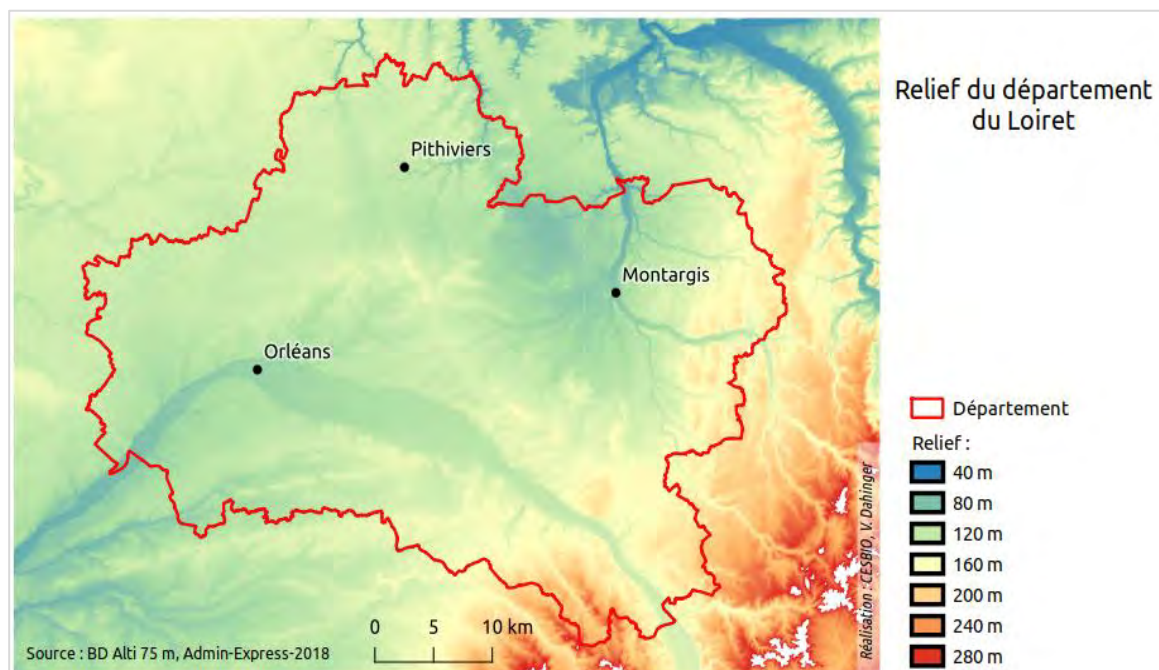


Figure 2 : Carte du relief du département du Loiret

Le territoire d'étude se situe au carrefour de plusieurs régions naturelles. Le Val de Loire situé de part et d'autre de la Loire, la Sologne, au Sud, composée d'une vaste forêt marécageuse, la forêt d'Orléans au centre, le Gâtinais relativement vallonné au Nord-est, la Beauce réputée pour ses grandes plaines céréalières à l'Ouest et au Sud-est le Berry et la Puisaye.

Ces grandes régions naturelles influencent l'occupation du sol (Figure 3). De manière générale, le département est largement dominé par les territoires agricoles qui occupent près des deux tiers du département. Les forêts et les milieux naturels représentent quant à eux plus d'un tiers du territoire et les espaces artificialisés seulement 6 % (Corine Land Cover 2012).

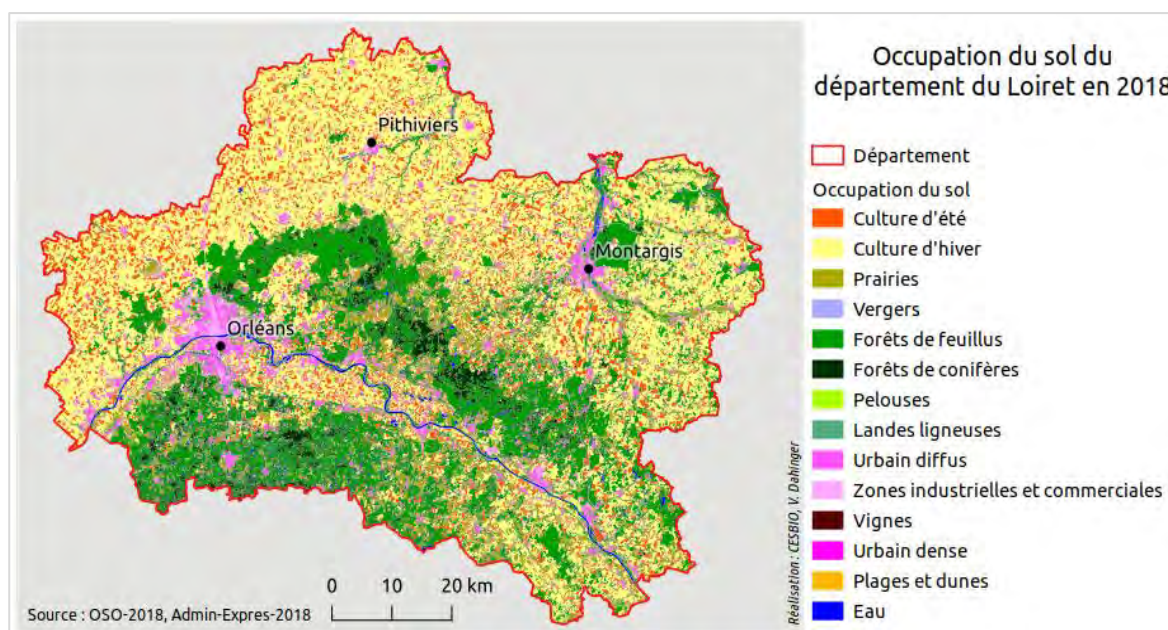


Figure 3 : Carte d'occupation du sol du Loiret en 2018

B. CARACTERISTIQUES CLIMATIQUES

Le département du Loiret est sous influence océanique avec un impact sur son climat. On note toutefois des différences, avec notamment des températures légèrement plus froides en hiver et plus chaudes en été, des précipitations un peu moins importantes et des vents plus faibles.

1. METEOROLOGIE

Selon une analyse des météorologistes de Météo France, réalisée à partir de quatre stations dans la région Centre-Val de Loire depuis les années 1960, les températures moyennes ont augmenté de 0,3°C par décennie. Ces réchauffements sont d'autant plus marqués en été avec +0,4°C par décennie. Le cumul des précipitations annuelles est quant à lui très variable d'une année à l'autre. Les météorologistes constatent également une accentuation du phénomène ces vingt-cinq dernières années.

Afin d'analyser les conditions climatiques spécifiques aux deux années étudiées au cours du stage, les données météorologiques de 2017 et 2018, de trois stations de Météo-France ont été observées (*Notebook du traitement et analyse des données météorologiques du Loiret, voire annexe n°3*). Ces stations ont la particularité de se situer dans des villes relativement éloignées l'une de l'autre : Amilly (proche de Phitiviers), Orléans et Phitiviers-le-viel (**Figure 2 et 3**).

En 2017, la température moyenne sur l'année est de 12°C avec un cumul total des précipitations de 603 mm. Les valeurs extrêmes ont été atteintes au mois de janvier avec -10,6°C pour la minimale et 36,8°C pour la maximale en mai. Les volumes des précipitations ont été particulièrement importants en août et en septembre avec des cumuls mensuels maximums respectifs de 80 et 82 mm d'eau. Au contraire, les mois les plus secs sont avril et octobre avec des cumuls mensuels minimums de 20 et 18 mm d'eau.

Une période de sécheresse est déterminée en fonction de l'indice de Gaussen (Bagnouls et Gaussen, 1952) qui identifie les mois secs. Les mois secs sont ceux qui vérifient l'équation $P < 2T$, où P est le cumul de précipitations mensuels (mm par mois) et T est la moyenne de température mensuelle (°C). Les **Figures 4 et 5** sont des diagrammes ombrothermiques (Gaussen, 1956) qui représentent graphiquement l'indice de Gaussen. Les courbes de températures sont tracées à partir de l'axe des ordonnées à gauche, les cumuls des températures représentés en bâton sont quant à eux tracés à partir de l'axe des ordonnées à droite. La relation entre les deux axes vérifie l'équation $P = 2T$.

A partir de la **Figure 4** on constate deux périodes de sécheresse en avril et en octobre avec des carences en pluie par rapport aux températures.

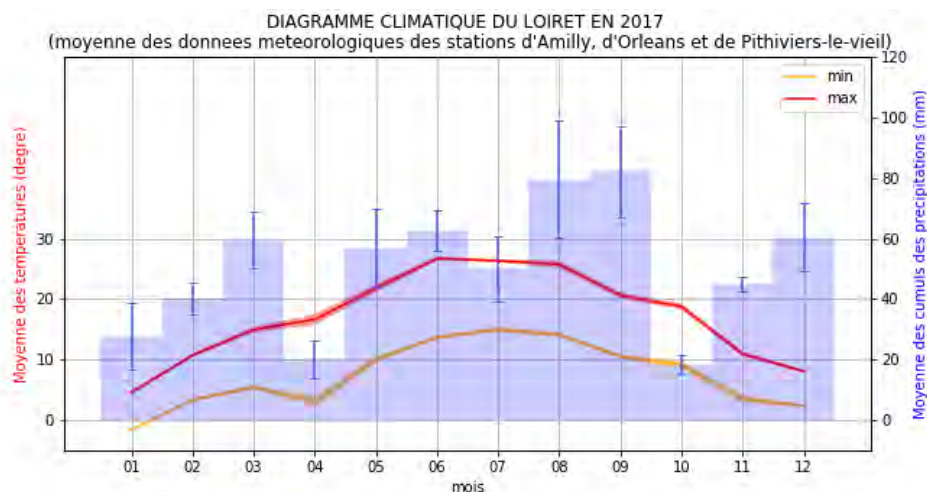


Figure 4 : Diagramme ombrothermique du Loiret en 2017

En 2018, la température moyenne sur l'année est également de 12°C avec un cumul total des précipitations de 637 mm. Les données moyennées sont pratiquement identiques à l'année précédente. En observant plus attentivement, un pic de froid à -12,6°C est observé en février et 37,5°C pour la maximale atteinte sur deux mois consécutifs en juillet et en août. Les volumes des précipitations varient particulièrement par rapport à 2017, avec des cumuls mensuels maximums en janvier et en mai, jusqu'à 98 et 75 mm respectivement. Au contraire, les mois les plus secs sont août, septembre et octobre avec des cumuls mensuels minimums de 12, 12 et 29 mm d'eau.

A partir de la Figure 5 on constate une longue période de sécheresse pendant l'été et le début de l'automne avec des carences en pluie et des températures très chaudes en été.

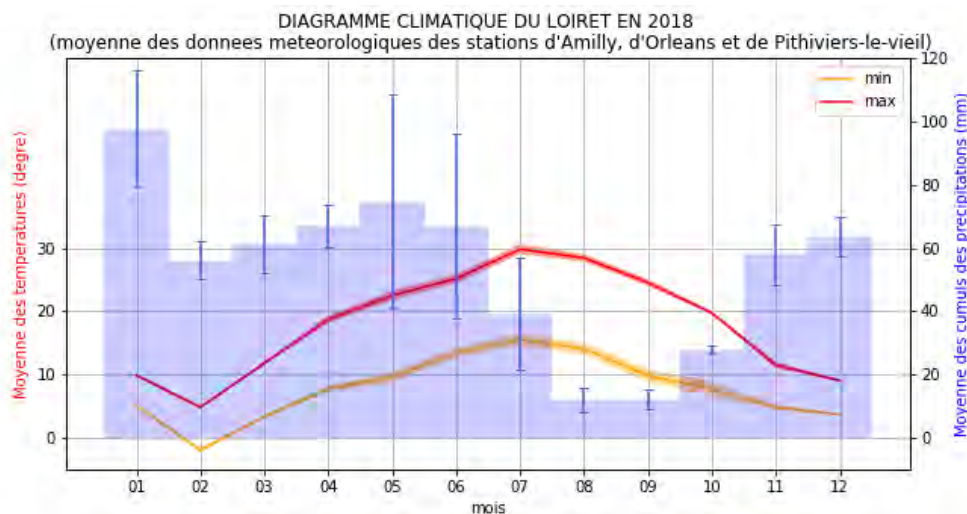


Figure 5 : Diagramme ombrothermique du Loiret en 2018

2. SECHERESSE

En 2017, pour la première fois un arrêté anti-sécheresse a été signé au mois de novembre pour 68 communes du Loiret. C'est le résultat de faibles précipitations pendant l'année et d'une faible recharge des nappes phréatiques.

En 2018, l'été a été chaud, le déficit pluviométrique combiné aux températures élevées a accentué la sécheresse des sols superficiels. Le 12 octobre le Préfet du Loiret reconnaît le « cas de force majeure » pour le département.

Selon les prévisions *Climat HD* de Météo France, l'assèchement des sols a tendance à être de plus en plus marqué au cours du XXI^e siècle en toute saison. La situation actuelle du Loiret, placé en zone rouge dès le mois de juillet 2019, confirme ces prévisions.

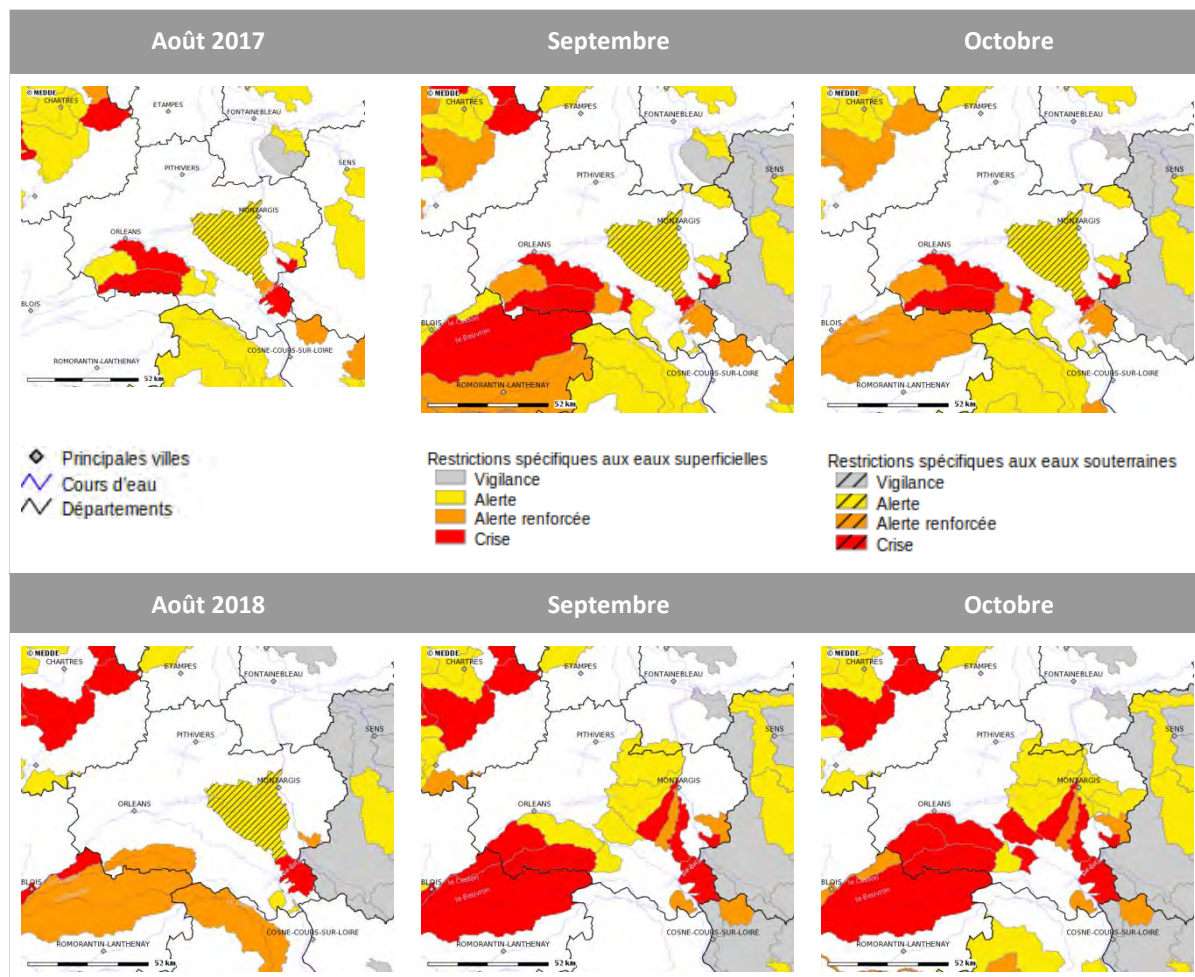


Figure 6 : Cartes des arrêtés de restriction d'eau du Loiret en 2017 et 2018 (www.propluvia.developpement-durable.gouv.fr)

C. PRATIQUES AGRICOLES ET IRRIGATION

Selon l'Agreste dans le cadre du recensement agricole 2010, avec plus de 471 000 hectares de surface potentiellement irrigable, la région Centre se range en première position en France, bien avant l'Aquitaine (302 000 ha) et le Midi-Pyrénées (290 000 ha). Elle arrive en quatrième position sur les volumes d'eaux consommés pour l'irrigation.

1. SUPERFICIE IRRIGUEE ET METHODE D'IRRIGATION

Le cœur de Beauce est particulièrement bien équipé avec un taux de surface irrigable qui dépasse les trois-quarts de la SAU (221 000 ha), soit 70 % des 318 000 ha irrigués dans la région Centre en 2010.

L'irrigation est particulièrement présente dans le département du Loiret avec 72 % de SAU qui est équipée avec une superficie moyenne des exploitations irrigables de 119 ha.

En 2010, 55 % des exploitations ont irrigué en Beauce contre 48 % en 2000. Cette hausse s'explique par une disparition plus importante des exploitations ne pratiquant pas l'irrigation. Le taux de disparition des exploitations irriguées est de 11 % alors qu'il est de 32 % pour les autres.

Selon les données sur les prélèvements quantitatifs en eau (Eaufrance, www.eaufrance.fr), le département du Loiret a prélevé 98 474 059 m³ d'eau pour l'irrigation en 2016 (Figure 7).

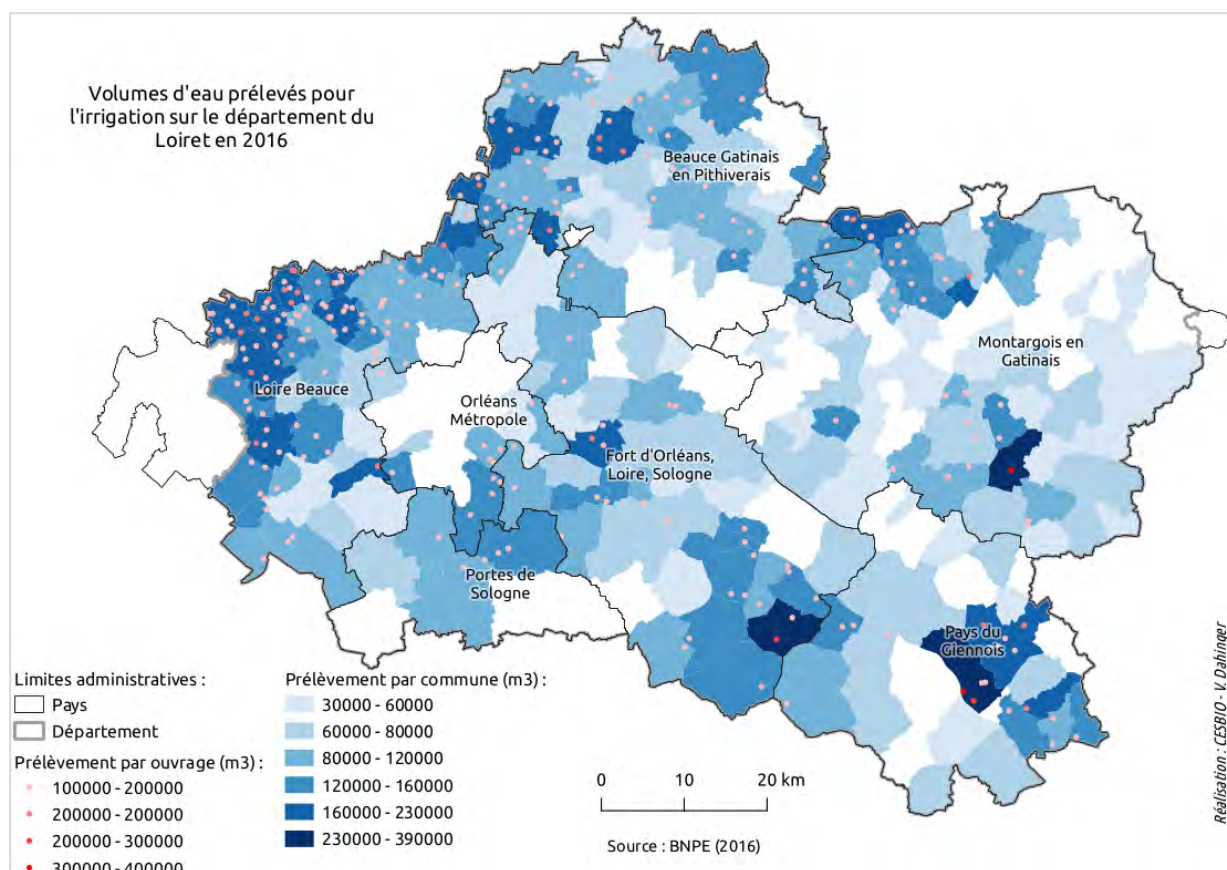


Figure 7 : Volumes d'eau prélevés pour l'irrigation en 2016 sur le Loiret

La méthode d'irrigation très largement dominante en région Centre est l'aspersion dans 99 % des cas. Le goutte-à-goutte est surtout présent dans les vergers d'Indre-et-Loire. La lame d'eau apportée sur les cultures de la région Centre est relativement modeste avec seulement 90 mm par hectare comparée à la moyenne française de 165 mm en 2010. La position septentrionale de la région explique en partie ce constat. L'irrigation en région Centre est utilisée en appoint pour pallier les déficits pluviométriques et assurer la qualité et les quantités produites. En 2010, près des trois quarts de l'eau d'irrigation utilisée en région Centre sont pompés dans les nappes phréatiques par des puits et des forages. A l'échelle du Loiret en 2016, 93,4 % de l'eau prélevée pour l'irrigation provient des eaux souterraines.

2. CULTURES IRRIGUEES ET EVOLUTIONS DES PRATIQUES

En Beauce, l'accès à l'eau a permis de développer les cultures à forte valeur ajoutée, comme la betterave avec une progression de sa superficie de 16% en 10 ans, mais également les pommes de terre dont les surfaces cultivées ont plus que doublé sur la même période. En 2010 et en région Centre, plus de 90% des cultures sous serres, des légumes de plein champs, des pommes de terre et des betteraves sont irrigués (**Figure 8**). L'irrigation sur blé s'est également développée passant de 5% de l'assolement en 2000 à 51% en 2010. Quant aux parcelles de maïs en grain, 60% sont irriguées. L'irrigation est, en région Centre, le facteur de production essentiel pour le développement du potentiel de production agricole.

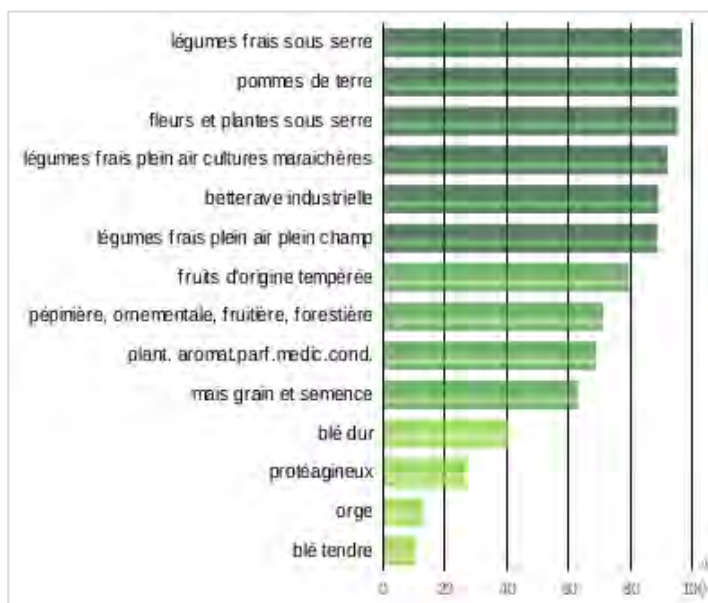


Figure 8 : Part de la SAU irriguée par rapport à la SAU en région Centre (Agreste - Recensement agricole 2010)

Entre 2000 et 2010, en Beauce, on observe une diminution de 30% des surfaces irriguées de maïs grain et semence qui s'explique par une politique agricole et des différentiels de prix par rapport au blé, moins favorables au maïs que par le passé (**Figure 9**).

Le niveau d'irrigation est très variable selon les conditions climatiques des années (précipitations et températures). Les étés particulièrement secs de ces dernières années peuvent amener à une irrigation plus importante pour maintenir les niveaux de production agricole.

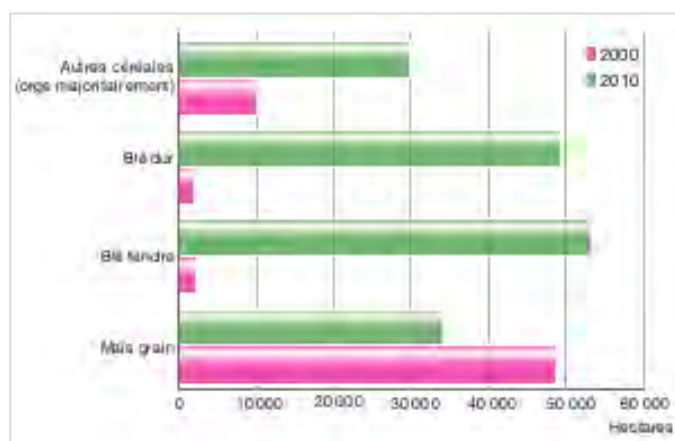


Figure 9 : Evolution des superficies irriguées par culture en nappe de Beauce (Agreste - Recensement agricole 2010)

Section III - UTILISATION DES IMAGES SATELLITAIRES EN AGRICULTURE

Les années 1970 marquent l'essor de la télédétection pour l'étude de la surface de la Terre à partir de satellites et ouvrent l'époque de la télédétection moderne. L'augmentation de la résolution spatiale et la diversification des capteurs permettent de visualiser ce que l'œil humain ne peut pas. Parallèlement, la diffusion accélérée et l'augmentation de la puissance des ordinateurs contribuent à développer de nouvelles méthodes d'analyse des données.

- *Les données de télédétection et l'agriculture*
- *Les opportunités associées au programme Copernicus*
- *Les méthodes de classification*
- *Association des données optiques et radar*

A. LES DONNEES DE TELEDETECTION ET L'AGRICULTURE

L'acquisition à distance d'informations sur les objets terrestres, permise par les propriétés des ondes électromagnétiques émises ou réfléchies par ces objets, offre un large champ d'application aux données de télédétection. L'analyse des signatures spectrales et les techniques de seuillage permettent d'identifier les objets du sol. Pour les objets complexes qui ne disposent pas d'une signature spectrale caractéristique, l'analyse conjointe de plusieurs bandes ou des indicateurs de synthèse peuvent permettre de distinguer ces objets. Les indices spectraux sont des résultats d'opérations entre bandes spectrales utilisés pour extraire de l'information. Dès le début des années soixante-dix, l'indice de végétation *Normalized Difference Vegetation Index* (Rouse 1973, Tucker 1979) est utilisé pour l'étude des couverts végétaux. Il est aujourd'hui l'indice spectral le plus populaire et notamment intégré dans des modèles de prédictions basés sur des critères agrométéorologiques. L'usage du cumul des valeurs de NDVI, observé pendant la saison de culture est utilisé dans plusieurs études visant à estimer la production de biomasse (Geffard 1994). Ainsi, on constate que le rendement en blé au Maroc est hautement corrélé au NDVI cumulé (Balaghi, 2006). Le cumul de NDVI est aussi employé pour l'analyse des boisements et reflète bien la souffrance des peuplements résineux du Sud du Massif Central et la probabilité de dépérissement (Cheret et al., 2017). L'utilisation des cumuls de NDVI, ou d'autres indices spectraux, n'était jusqu'à ce jour pas évalué pour la classification d'images acquises par télédétection.

B. LES OPPORTUNITES ASSOCIEES AU PROGRAMME COPERNICUS

Les séries temporelles utilisées pour les classifications sont capturées par les satellites d'observation de la Terre tels que LANDSAT ou SPOT. Ces derniers fournissent des images acquises à haute résolution spatiale (10 à 30 m), dans le visible et l'infrarouge, mais avec une répétitivité temporelle insuffisante pour suivre la dynamique des cultures (une image par mois en moyenne). Le « programme européen de surveillance de la Terre », nommé *Copernicus*, offre de nouvelles perspectives pour la communauté scientifique. Les capteurs Sentinel-1 et Sentinel-2 fournissent des images radar et optiques respectivement, à une haute résolution temporelle (1 à 5 jours) et spatiale (10 à 60m). Le capteur des

Sentinel-2 a été spécialement conçu pour observer la végétation terrestre. Il mesure les quantités d'énergies dans 13 bandes spectrales, allant du visible au moyen infrarouge. Ses trois bandes en végétation proche infrarouge (*vegetation red edge*) permettent d'évaluer la vigueur d'une plante jusqu'à une résolution de 10 m. De plus, les données Sentinel sont distribuées, en accès libre et gratuit (www.peps.cnes.fr, www.esa-sen2agri.org ou www.theia.cnes.fr), et soumises à différents prétraitements (ortho-rectifiées, corrections géométriques et radiométriques) également *open-source*. La grande variété d'images produites favorise l'émergence de méthodes et d'outils de traitement pour répondre aux divers enjeux thématiques.

C. LES METHODES DE CLASSIFICATION

En classification numérique on distingue deux grandes approches : la classification supervisée et non supervisée. La classification supervisée est un processus dit inductif, où le classifieur considère un ensemble d'exemples labellisés, la cible à apprendre est donc connue. En revanche, la classification non supervisée ne dispose d'aucune information a priori sur les données. L'objectif du classifieur est de regrouper les objets dans des classes en se basant sur leur ressemblance.

Les méthodes de classification d'images de télédétection sont traditionnellement basées sur l'information spectrale associée à chaque pixel. Les méthodes orientées objet considèrent que les informations sémantiques nécessaires à l'interprétation d'une image ne sont pas représentées par les pixels individuels mais dans des groupes de pixels associés, des images-objets.

L'apprentissage automatique réalisé par le classifieur est basé sur un algorithme choisi en fonction des données à traiter. Selon le classifieur choisi, les résultats varient et les critères comme la précision, la durée d'apprentissage, la linéarité ou le nombre de paramètres sont important à étudier au préalable.

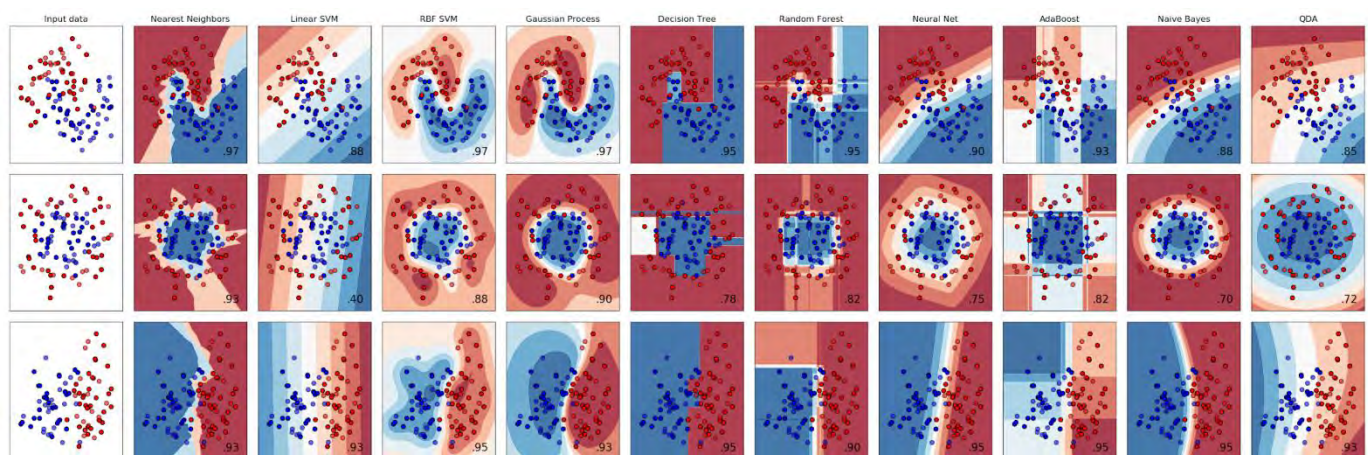


Figure 10 : Différences entre les classifieurs implémentés par la bibliothèque d'apprentissage statistique Scikit-learn

« L'arbre de décision » est un des algorithmes de *Machine learning*. A partir d'observations étiquetées, qui constituent les racines, l'arbre de décision classe des observations futures en parcourant une série de branches. L'intersection de branches s'appelle un nœud et représente une règle qui permet de diviser au mieux les observations de départ. Lors du classement, les observations parcourent l'arbre en fonction des règles qu'elles respectent ou pas. L'algorithme des « Forêts aléatoires », ou *Random Forest*, introduit

par Leo Breiman en 2001, se fonde sur les arbres de décisions. Des variables sont extraites des observations et permettent d'entraîner un arbre de décision sur le jeu de données. Cette opération est réalisée n fois pour obtenir n arbres qui proposent une classe, la feuille. Le modèle prédit finalement la classe de l'objet selon le vote majoritaire des arbres (Liaw et Wiener, 2002). Cet algorithme fait partie des méthodes les plus précises et les plus efficaces. Parmi ses avantages, l'interprétabilité facilitée par l'ensemble des règles auquel la donnée a répondu est un des aspects de son succès. Le modèle étant non paramétrique, par opposition au modèle paramétrique (distribution normale, une classe est caractérisée par des paramètres statistiques), l'algorithme est aussi résistant aux valeurs aberrantes. Une donnée aberrante au sein du jeu d'apprentissage contamine la feuille et non pas le modèle et a donc un faible impact pour les autres. Ces caractéristiques en font un classifieur très puissant, même si le jeu d'entraînement est faible, avec un temps de traitement adapté au *Big data* (Guenier et Poggi, 2016).

L'omniprésence et la complexité du *Machine learning* en traitement d'images motivent le développement de logiciels visant à faciliter les classifications. La chaîne de traitement *iota*², développée par le CESBIO, permet par exemple d'automatiser la production des cartes d'occupation du sol en France métropolitaine (Inglada, 2017).

D. ASSOCIATION DES DONNEES OPTIQUES ET RADAR

A l'échelle mondiale, de nombreux résultats de classifications effectuées à partir de séries temporelles multi-spectrales ont été publiées, notamment sur la cartographie de l'occupation et de l'usage du sol, ou les surfaces irriguées (Goetz et al., 2004 ; Dheeravath et al., 2010). L'utilisation des données radar pour la détection des surfaces irriguées présente l'avantage de ne pas être sensible à la couverture nuageuse. De plus, les radars à synthèse d'ouverture sont sensibles à différents paramètres des surfaces continentales (sol, végétation). La complémentarité des images optiques et radar offre un potentiel important pour améliorer les résultats de classification. En Chine, la combinaison des données (Sentinel-1 et Landsat-8) a permis d'augmenter de 10 points la part du blé d'hiver correctement détecté (Zhou et al., 2017). La surveillance des variations spatio-temporelles à l'aide d'observations Sentinel-1 et -2, en Inde, a été utilisée pour classer des parcelles de riz, de maïs et d'autres cultures irriguées (Ferrant et al., 2017). En zone tempérée, la détection des pratiques d'irrigation est encore peu développée, et le traitement en cours de saison en est à ses balbutiements. De récents travaux montrent que l'association des données optiques et radar améliore la classification précoce des cultures irriguées dans le sud-ouest de la France (Demarez et al., 2019).

Pour conclure la première partie, le stage a pour objectif de distinguer les pratiques d'irrigation sur les parcelles de maïs du Loiret pour répondre aux besoins de la DRAAF. Situé en zone tempérée, le territoire d'étude est de plus en plus touché par des épisodes de sécheresse. La forte corrélation entre le rendement du maïs et la quantité d'eau disponible pour la plante, motive les agriculteurs à se tourner vers l'irrigation de leur culture. A partir des informations de longueurs d'ondes (visible et infrarouge) contenue dans les images satellitaires, la classification supervisée des pixels de maïs permet d'identifier la pratique (irrigué / non irrigué). Malgré la complémentarité des données Sentinel-1 (radar) et Sentinel-2 (optique), l'étude se concentre uniquement sur l'analyse des séries temporelles optiques pour explorer l'apport d'une déclinaison d'indices spectraux traités en cumul.

PARTIE 2 : CLASSIFICATION ET CARTOGRAPHIE DES PARCELLES DE MAÏS IRRIGUE ET NON IRRIGUE

Cette partie est consacrée à la description du travail de classification et de cartographie des parcelles de maïs irrigué et non irrigué. Les quatre sections suivantes fourniront pour commencer une **description des données** (Section I), puis **une présentation des outils** (Section II), suivie du **détail de la méthode** (Section III), pour terminer sur la **présentation des résultats et leur analyse** (Section IV).

Section I - DESCRIPTION DES DONNEES

Pour cartographier l'occupation du sol ou des surfaces irriguées, dans notre cas le maïs, le travail de classification d'images est couramment utilisé. L'approche de classification supervisée retenue pour l'étude nécessite différentes données pour alimenter le modèle.

- *Données satellitaires*
- *Données de référence*
- *Données sur la composition du sol*

A. DONNEES SATELLITAIRES

Pour cette étude les données Sentinel-2 utilisées sont les produits distribués par Theia de niveau 2A (www.theia.cnes.fr). Ces produits sont corrigés des effets atmosphériques et accompagnés de masques des nuages et de leurs ombres. Le traitement des données est réalisé avec la chaîne MAJA (Lonjou et al., 2016), par le centre de traitement Muscate du CNES et pour le compte du centre de données Theia. Les réflectances de surface sont fournies sous la forme d'un fichier GeoTiff pour chaque bande, à une résolution de 10m, 20m ou 60m selon les bandes spectrales (Figure 11).

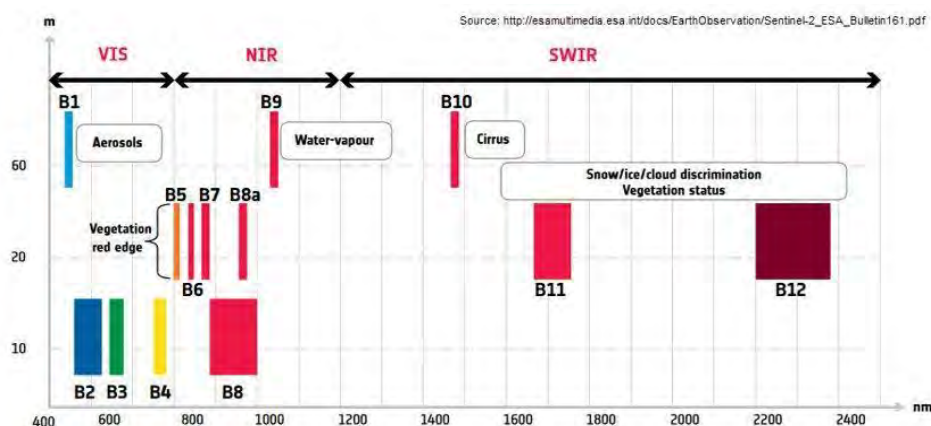


Figure 11 : Résolution spatiale par longueurs d'ondes capturées par Sentinel-2 (ESA)

Pour couvrir entièrement le territoire d'étude, quatre tuiles Sentinel-2 sont nécessaires, les tuiles 31UCP, 31UDP, 31TCN et 31TDN (**Figure 12**). Chaque tuile est de format 110 X 110km², avec une superposition de 10km des tuiles adjacentes. Les deux premiers numéros du nom de la tuile (dans notre cas 31) correspondent au numéro de fuseau de la zone UTM (au total 60 fuseaux de 8 degrés en longitude). Chaque zone est divisée en bandes latitudinales de 6 degrés, représentée par la première lettre qui s'accroît du Pôle Sud au Pôle Nord. Ainsi, les deux premiers chiffres et la première lettre font référence à une grille de cases UTM. Chaque case est ensuite découpée en tuiles. La deuxième lettre désigne la position de la tuile sur l'axe Ouest-Est. Et la troisième, sur l'axe Sud-Nord.

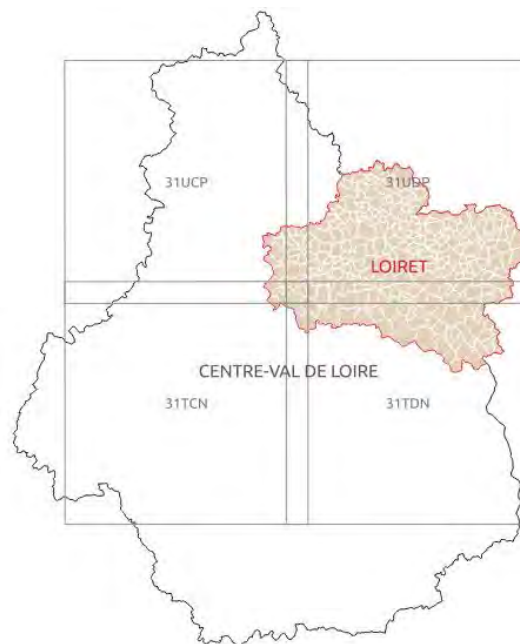


Figure 12 : Carte des emprises des tuiles Sentinel-2 sur le territoire d'étude

La **Figure 13** représente le nombre de dates disponibles sur le serveur du CNES, par tuile et par mois.

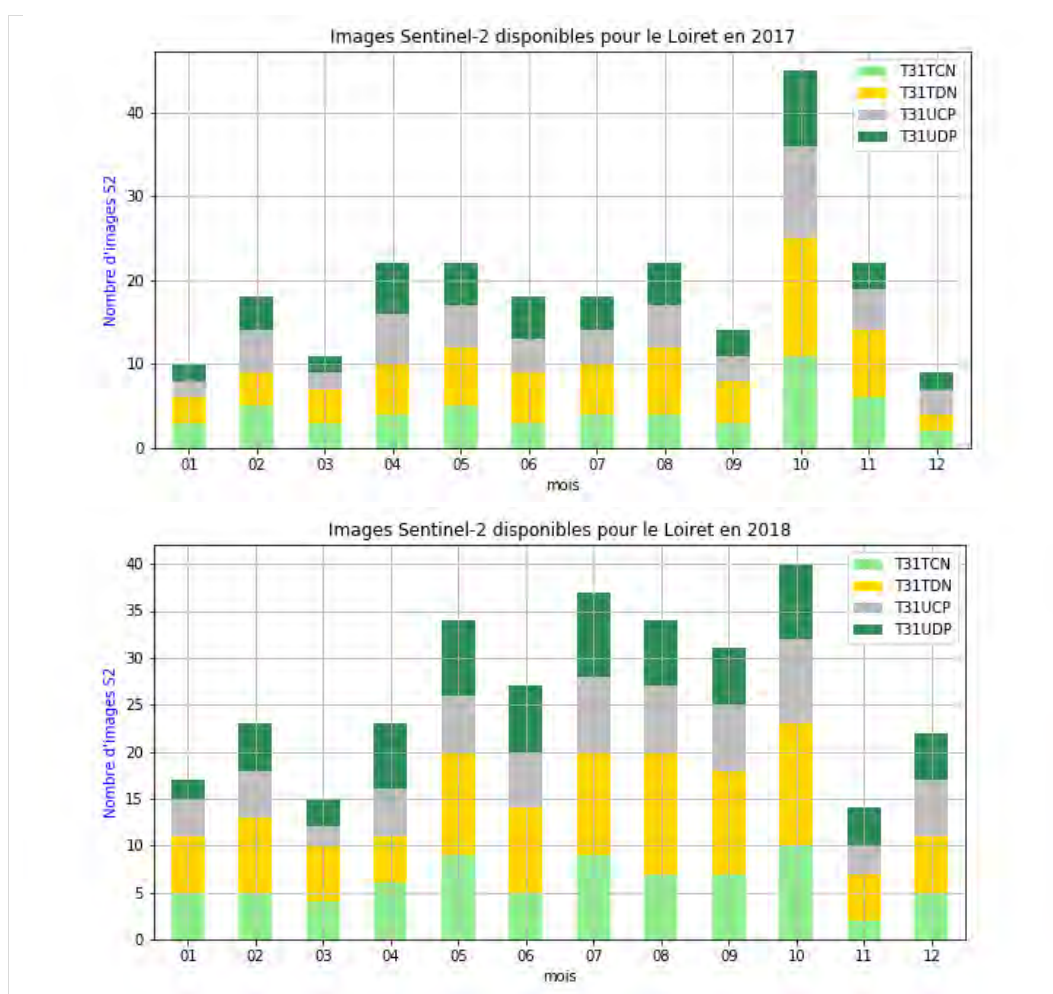


Figure 13 : Images Sentinel-2 disponibles en 2017 et 2018 pour couvrir le territoire d'étude

Les tests préliminaires visant à ajuster la configuration de la classification sur les quatre tuiles, ont été réalisés sur la tuile 31UDP uniquement, choisie pour sa couverture de près de 75 % du territoire d'étude,

pour diminuer les temps de traitement, la charge de travail et les volumes de données. Le **Tableau 1** regroupe les miniatures des images Sentinel-2 en 2017 et 2018 (www.theia.cnes.fr). On remarque une couverture nuageuse importante au cours de l'année 2017, à l'exception des mois d'hiver et du mois de novembre. Au contraire, l'année 2018 est globalement moins affectée par les nuages, avec une couverture faible voire absente de mai à octobre.

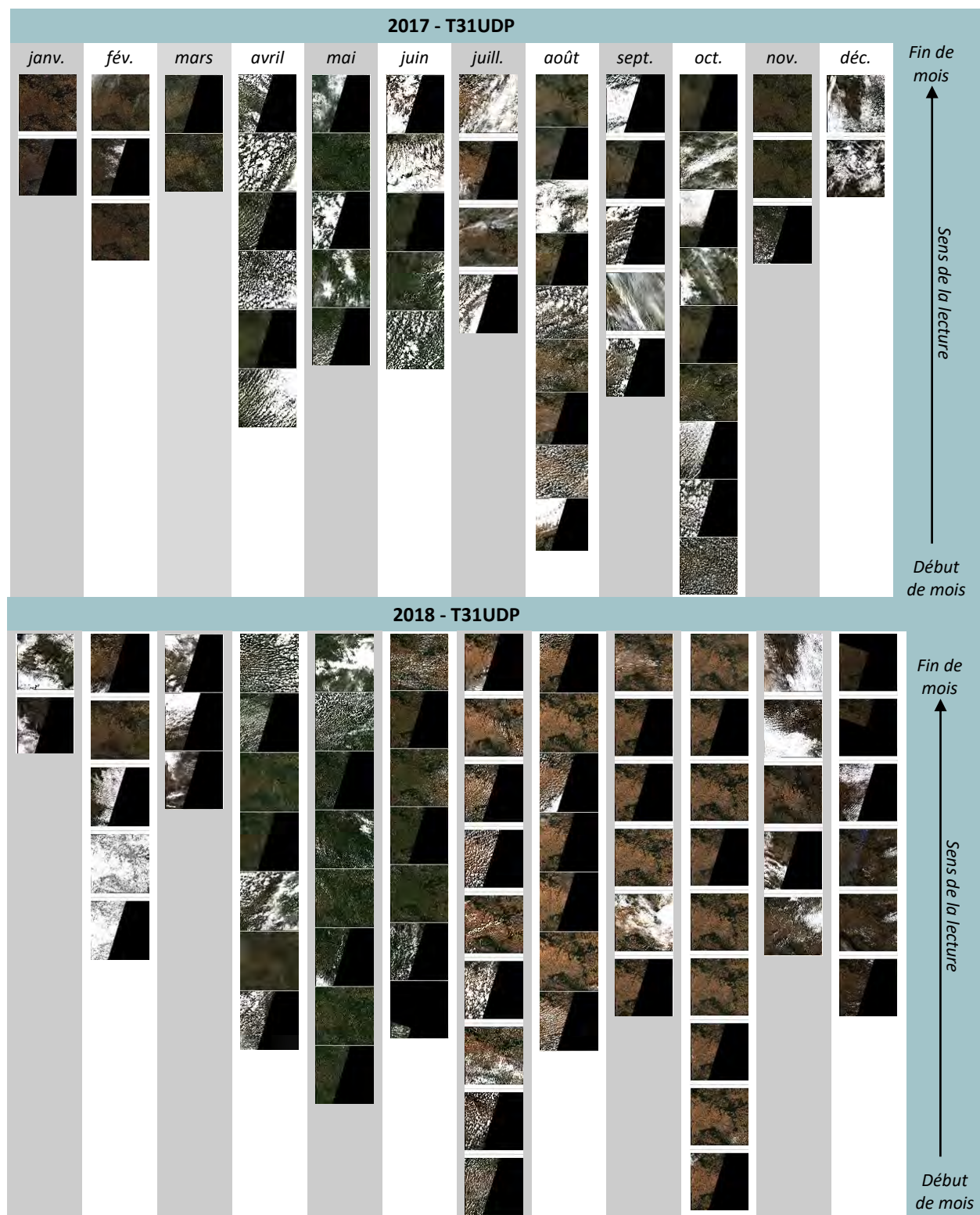


Tableau 1 : Nébulosité en 2017 et 2018

B. DONNEES DE REFERENCE

En tant que partenaire de l'étude, la DRAAF Centre-Val de Loire a mis à disposition les données de référence pour 2017 et 2018, collectées lors de campagnes de terrain.

En 2017, la campagne réalisée le 3 août a été axée sur l'observation visuelle de l'irrigation ou non du maïs, avec comme principaux critères la présence du matériel d'arrosage, ou traces de sa présence, et l'état physique du maïs. Ces données ont aussi été complétées par des données recueillies par le Service Eau Environnement Forêt de la Direction Départemental des Territoires (DDT) du Loiret, lors des contrôles terrain dits « environnement » portant sur le bilan azoté.

En 2018, la campagne du 2 août a suivi la même méthodologie que l'année précédente. Les données ont pu être complétées par celles de la DDT.

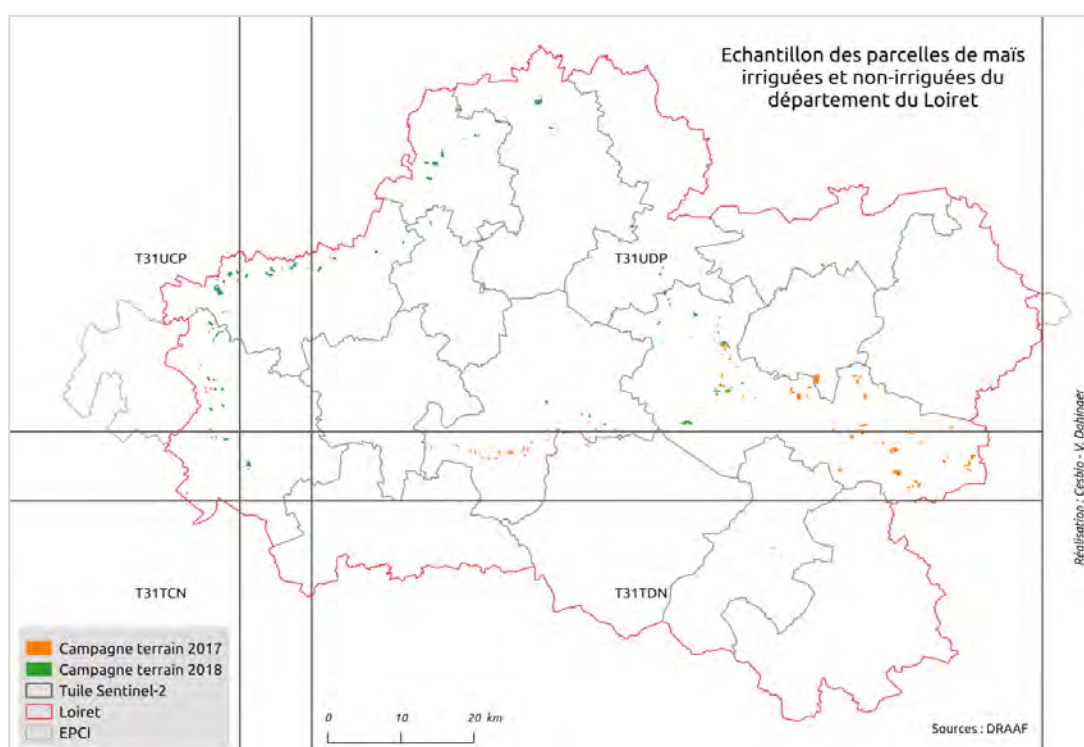


Figure 14 : Carte de la répartition des données de référence en 2017 et 2018

C. DONNEES SUR LA COMPOSITION DU SOL

Le programme *GlobalSoilMap* (GSM, www.globalsoilmap.net) vise à produire une base de données numérique des propriétés du sol, harmonisée et mondiale, à une résolution de 100 m. Lancé en 2006 à l'initiative de l'Union Internationale de Science du Sol, il est porté localement par un consortium d'organismes spécialisés. Pour atteindre ces objectifs, le programme GSM s'appuie sur l'émergence des approches de cartographie de sols à base de modélisation statistique. Les propriétés sont délivrées sous la forme de valeurs moyennes assorties d'intervalles de confiance pour rendre compte de l'incertitude et renseignent au minimum sur :

- la teneur en carbone organique,
- la granulométrie,

- le pH,
- la profondeur exploitable par les racines.

A partir des propriétés précédentes, la masse volumique et la réserve en eau utile en sont déduites.

En France, l'INRA délivre aux nœuds d'une grille de 90 m de résolution, des estimations quantifiées de propriétés pérennes comme la composition des sols en pourcentage (argile, limon, sable, éléments grossiers), la teneur en carbone organique, le pH ou encore la réserve en eau utile des sols. Ces données sont fournies pour six intervalles de profondeurs (0-5 cm, 5-15 cm, 15-30 cm, 30-60 cm, 60-100 cm, 100-200 cm), avec une estimation des incertitudes.

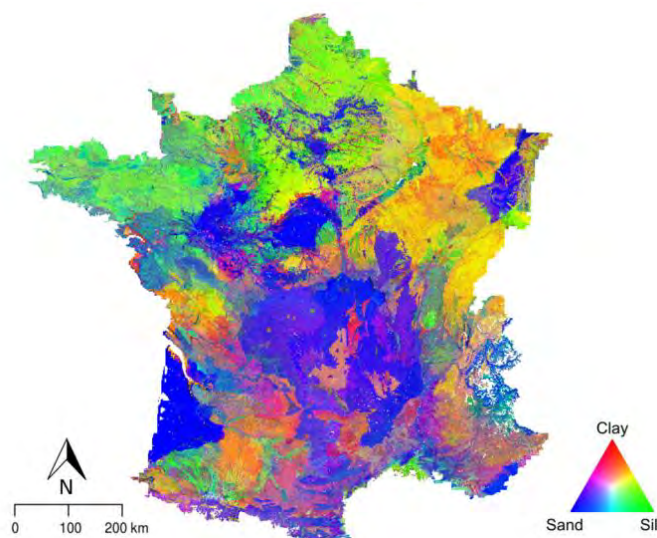


Figure 15 : Carte GSM de prédiction de la texture du sol (g/kg) (Mulder et al. 2016)

Section II - PRESENTATION DES OUTILS

Face à l'important volume de données à traiter, l'utilisation d'outils dédiés au stockage, à l'analyse et à l'exploitation des données permet de faciliter les traitements.

- La chaîne de traitement *iota*²
- Le serveur

A. LA CHAÎNE DE TRAITEMENT IOTA²

« Infrastructure pour l'Occupation des sols par Traitement Automatique Incorporant les Orfeo Toolbox Applications » est le nom de la chaîne de traitement *iota*² développée par le CES OSO pour la production des cartes d'occupation du sol. Cet outil *open-source* conçu pour gérer de gros volumes de données, sur une zone multi-tuiles, des séries temporelles et multi-capteurs, est en perpétuelle mutation pour élargir ses possibilités et répondre au mieux aux besoins.

1. LES INPUTS

La méthode de classification appliquée par *iota*² repose sur la procédure de classification supervisée des pixels à partir de descripteurs spatio-spectro-temporel. Son originalité principale est de pouvoir l'utiliser pour un large territoire d'étude, de traiter de grands volumes de données et de produire des cartes de classification de manière automatique. Pour cela, l'utilisateur doit fournir à *iota*² des inputs, regroupés en trois grandes catégories :

- **Données de référence :**

Les données dites de référence correspondent aux réalités terrain. Ces données représentent un échantillon de parcelles de maïs irrigué et de maïs non irrigué utilisées pour l'entraînement du classifieur et la validation du modèle. Elles sont regroupées dans une base de données géographique au format *shapefile*. Chaque entité détient une géométrie valide et un label (type *integer*), par exemple le label « irrigué » et le label « non irrigué » (label à partir de 1 et non 0). Le système de projection n'est pas contraint, toutefois la projection cartographique conforme de la surface de la Terre comme la projection Transverse universelle de Mercator (UTM) n'est pas acceptée par *iota*².

- **Données satellitaires :**

*iota*² peut traiter les images issues d'un ou de plusieurs capteurs simultanément. Les séries temporelles doivent être des produits de niveau 2A, c'est-à-dire ortho-rectifiées, exprimées en réflectance de surface, accompagnées de masques nuages, ombres de nuages, neige et eau.

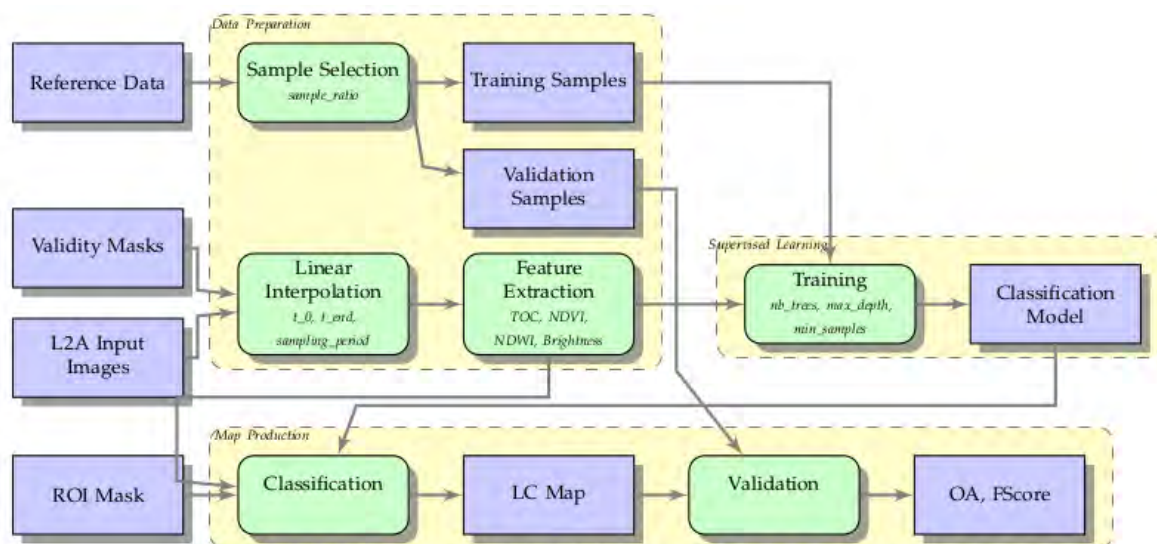
- **Données auxiliaires :**

Les données auxiliaires ou primitives sont des inputs optionnels fournis sous la forme de raster. Le raster doit respecter l'emprise d'une tuile d'image satellitaire et comporter les mêmes caractéristiques que la classification finale réalisée par *iota*², détaillées dans le fichier de configuration (système de référencement, résolution spatiale).

2. LA METHODE DE CLASSIFICATION DE *IOTA*²

La qualité principale du logiciel est sa rapide prise en main et la possibilité de changer les paramètres de classification sans intervention sur le code-source de l'algorithme. Comme son nom l'indique, *iota*² utilise les applications *Orfeo Toolbox* pour réaliser les traitements.

La **Figure 16** présente les grandes étapes de la classification et sont expliquées ci-après.

Figure 16 : Etapes de traitement *iota*² (Inglada, 2017)

a. ÉCHANTILLONNAGE

La méthode d'échantillonnage consiste tout d'abord à sélectionner un jeu d'entraînement en fonction du label des polygones. Il s'agit de séparer aléatoirement les données de référence pour consacrer une partie à l'apprentissage, l'autre à la validation. La séparation est réalisée en fonction d'un ratio déterminé par l'utilisateur. De manière générale on utilise entre 50 % et 70 % pour l'apprentissage, choisi selon le volume des données de référence. Ensuite, des pixels sont sélectionnés au sein des polygones puis dédiés à l'entraînement du classifieur. Plusieurs méthodes de sélection sont disponibles et l'utilisateur peut choisir de sélectionner le même nombre n d'échantillons dans chaque classe, de sélectionner un nombre donné d'échantillon toutes classes confondues pour maintenir la proportion initiale ou d'attribuer un nombre spécifique par classe. Pour limiter le biais lié à la sélection aléatoire des polygones, *iota*² offre la possibilité de choisir le nombre de tirages aléatoires. Chaque tirage abouti à un modèle, à une classification et à une validation.

b. INTERPOLATION LINEAIRE

L'interpolation linéaire consiste à interpoler les valeurs des pixels considérés comme invalides en raison d'une éventuelle couverture nuageuse à une date t . Pour cela, *iota*² se base sur la valeur de ce pixel à la date $t-1$ et à $t+1$, à condition qu'il n'y ait pas de couverture nuageuse, et attribue une valeur moyenne de réflectance au pixel de la date t . Ce processus de *gapfilling* est un traitement indispensable à la classification multi-tuiles. En effet, même si les périodes de revisite des satellites sont constantes, les dates et les périodes d'acquisition pour les tuiles voisines peuvent être différentes. En réalisant une interpolation linéaire systématique, les dates d'acquisition sont identiques pour toutes les tuiles. L'utilisateur peut préciser l'intervalle de temps souhaité entre chaque image.

c. EXTRACTION DE DONNEES

*iota*² réalise l'extraction de données pour l'entraînement et la classification. Les séries temporelles interpolées sont également utilisées pour calculer des indices spectraux pour chaque pixel et à chaque date d'acquisition. Par défaut, l'indice de végétation par différence normalisé (NDVI), l'indice de stress hydrique (NDWI) et l'indice de brillance (IB) sont calculés. En sauvegardant les données

extraites par *iota*² (*Code d'extraction et de sauvegarde des données interpolées, voir Annexe n°4*), les données (dates interpolées et indices) peuvent se substituer aux images satellitaires fournies comme input, en les intégrant en tant que données auxiliaires. Cette méthode permet d'optimiser les temps de traitement, notamment lors de classification multi-tuiles.

d. ENTRAÎNEMENT

Le jeu d'entraînement généré au début du traitement est utilisé pour apprendre au classifieur à différencier les labels. Dans notre cas, le classifieur doit séparer les pixels en deux classes : le maïs irrigué et le non irrigué. Plusieurs méthodes d'apprentissage supervisées sont disponibles et l'utilisateur peut choisir entre la méthode des k plus proches voisins (KNN), une classification naïve bayésienne, ou encore les machines à vecteurs de support (SVM). Pour cette étude, la méthode *Random Forest* a été sélectionnée en raison de sa robustesse aux erreurs des données de référence (Inglada et al. 2017). Un des inconvénients de ce classifieur est qu'il détermine lui-même le nombre d'arbres décisionnels.

e. CLASSIFICATION

Une fois le modèle généré, la classification est réalisée sur l'emprise totale des tuiles. Chaque pixel est alors attribué à un label.

f. VALIDATION

Au terme de la classification, l'étape de validation permet d'évaluer les résultats obtenus. Les données de validation sont utilisées pour créer une matrice de confusion. Le label attribué à chaque pixel du jeu de validation est comparé à celui obtenu par le classifieur. Un ensemble d'indices de performance est calculé à partir de la matrice de confusion et présenté plus en détail dans la partie III.G.

B. LE SERVEUR

Les volumes de données générés lors de classifications à partir de séries temporelles multi-spectrales et multi-tuiles dépassent les capacités humaines d'analyse, et même celles des outils informatiques classique. A titre indicatif, une classification avec *iota*² sur une série temporelle d'un an et quatre tuiles Sentinel-2, traite environ 1,5 To de données satellitaires et génère environ 130 Go de données (intermédiaires et résultats). Le *Big Data* nécessite donc de puissantes machines informatiques avec une capacité de stockage et de calcul particulièrement importante.

L'accès à l'un des serveurs du CESBIO (Cesbiocalc1) a permis de stocker les données utilisées comme inputs, d'installer *iota*² avec ses dépendances et de lancer les traitements.

Caractéristiques du Cesbiocalc1 :

- Mémoire vive : 400 Go
- Hébergeur : *Promox Virtuel Environnement* solution de virtualisation libre, basée sur l'hyperviseur Linux KVM
- Modèle : émulation de Intel® Xeon® Silver 4114 CPU fréquence de 2,20 GHz
- Capacité de stockage : 11 To

Section III - DETAIL DE LA METHODE

L'outil *iota*² offre la possibilité d'ajuster la configuration de la classification pour s'adapter au thème de recherche et aux données. Le détail de la méthode de classification ci-dessous justifie les choix du paramétrage de la chaîne de traitement, le travail préliminaire effectué sur les données et la démarche d'optimisation des résultats.

- Échantillonnage et validation
- Interpolation linéaire
- Période d'observation
- Indices spectraux et apport de la temporalité
- Normalisation des données *GlobalSoilMap*
- Configuration des classifications
- Evaluation des performances de classifications
- Simplification des résultats

A. ÉCHANTILLONNAGE ET VALIDATION

Les données de référence sont relativement homogènes en considérant les surfaces sur les deux années malgré des tournées différentes. On constate cependant un décalage significatif du nombre de polygones de maïs non irrigué entre 2017 et 2018. La répartition entre les labels est quant à elle fortement déséquilibrée en terme de surface pour les deux années. En effet, on compte quatre fois plus de pixels irrigués que de non irrigués, ce qui représente un biais de surreprésentation et un risque de sur-apprentissage pour le classifieur. Parallèlement, on note pour 2018 que les nombres de polygones pour chaque culture ne sont pas très éloignés, soit 60 parcelles de maïs non irrigué et 71 parcelles de maïs irrigué. On peut d'ores et déjà conclure que les parcelles de maïs irrigué sont nettement plus grandes que les parcelles de maïs non irrigué.

Code culture	Pratique	Surface en ha		Nombre de pixels <i>estimé en fonction de la surface pour une résolution 10m</i>		Nombre de polygones	
		2017	2018	2017	2018	2017	2018
MIE <i>maïs ensilage</i>	non irrigué	11.43	7.97	1 143	797	4	2
	irrigué	21.03	17.25	2 103	1 725	3	2
MIS <i>maïs grain</i>	non irrigué	250.26	242.99	25 026	24 299	36	58
	irrigué	781.02	757.72	78 102	75 772	74	69
TOTAL maïs non irrigué		261.69	250.96	26 169	25 096	40	60
TOTAL maïs irrigué		802.05	774.97	80 205	77 497	77	71

Tableau 2 : Synthèse des tables attributaires des données de référence en 2017 et 2018

Lorsque deux habitats très différents se côtoient, on peut constater un développement des végétaux situés en limite de parcelle différent du reste de la culture. On appelle ce phénomène l'effet bordure. Pour s'affranchir de ce biais, les données de référence font l'objet d'une érosion de 10 mètres sur les polygones. Ce géotraitement permet aussi de limiter les éventuels décalages entre la numérisation des parcelles dans le Registre parcellaire graphique (RPG) et la réalité observée sur les images satellitaires. Une fois l'érosion réalisée, les parcelles d'une superficie inférieure à 0,1 ha sont supprimées.

Les tests, présentés dans la section IV consacrée à l'analyse des résultats, visant à comparer les méthodes d'échantillonnage de *iota*² permettent de mesurer le biais de surreprésentation de la culture irriguée. Deux méthodes d'échantillonnage sont testées :

- Tous les pixels du jeu d'apprentissage disponibles sont utilisés pour l'entraînement,
- Le jeu d'entraînement est fixé à 10 000 pixels par culture pour équilibrer les classes et lever le biais de sur-apprentissage.

B. INTERPOLATION LINEAIRE

L'interpolation linéaire réalisée par *iota*² permet d'interpoler le signal optique mesuré (réflectance) entre deux images sans nuages. Ainsi, l'intégralité des images Sentinel-2 de niveau 2A acquises sur notre zone est utilisée pour les quatre tuiles en 2017 et 2018.

Dans le cadre de l'étude, une interpolation linéaire de 10 jours est choisie pour garantir trois images par mois (*Extrait des descriptifs des bandes avec libellés des images interpolées, voir Annexe n°5*).

C. PERIODE D'OBSERVATION

1. ANNEE ENTIERE VERSUS SAISON DE CULTURE

Une méthode de classification basée sur l'utilisation d'images sur une année entière présente l'avantage d'être reproductible entre les années et les territoires car elle ne nécessite pas de connaître les cycles culturels (*Cycle de croissance et besoin du maïs, voir Annexe n°6*). D'un autre côté, le traitement des images acquises uniquement durant le cycle cultural permet de limiter le volume et les temps de traitement des données. L'analyse des profils temporels de NDVI en 2017 et en 2018 a permis de déterminer la période de croissance du maïs sur le département du Loiret qui a ensuite été utilisée pour les classifications.

Pour les deux années, les valeurs moyennes des NDVI suivent la même tendance. Globalement, le maïs est en phase de croissance à partir du mois de mai avec un maximum obtenu en août pour l'année 2017 et en juillet pour l'année 2018. Après ce pic de l'été, l'activité chlorophyllienne décroît, le maïs entre en phase de sénescence jusqu'en octobre-novembre. Ensuite, de nouvelles variations sont observées, ce qui peut être expliqué par la repousse de végétaux (**Figure 17**).

Nous avons donc fixé la durée de la période culturale de mai (début) à octobre (fin).

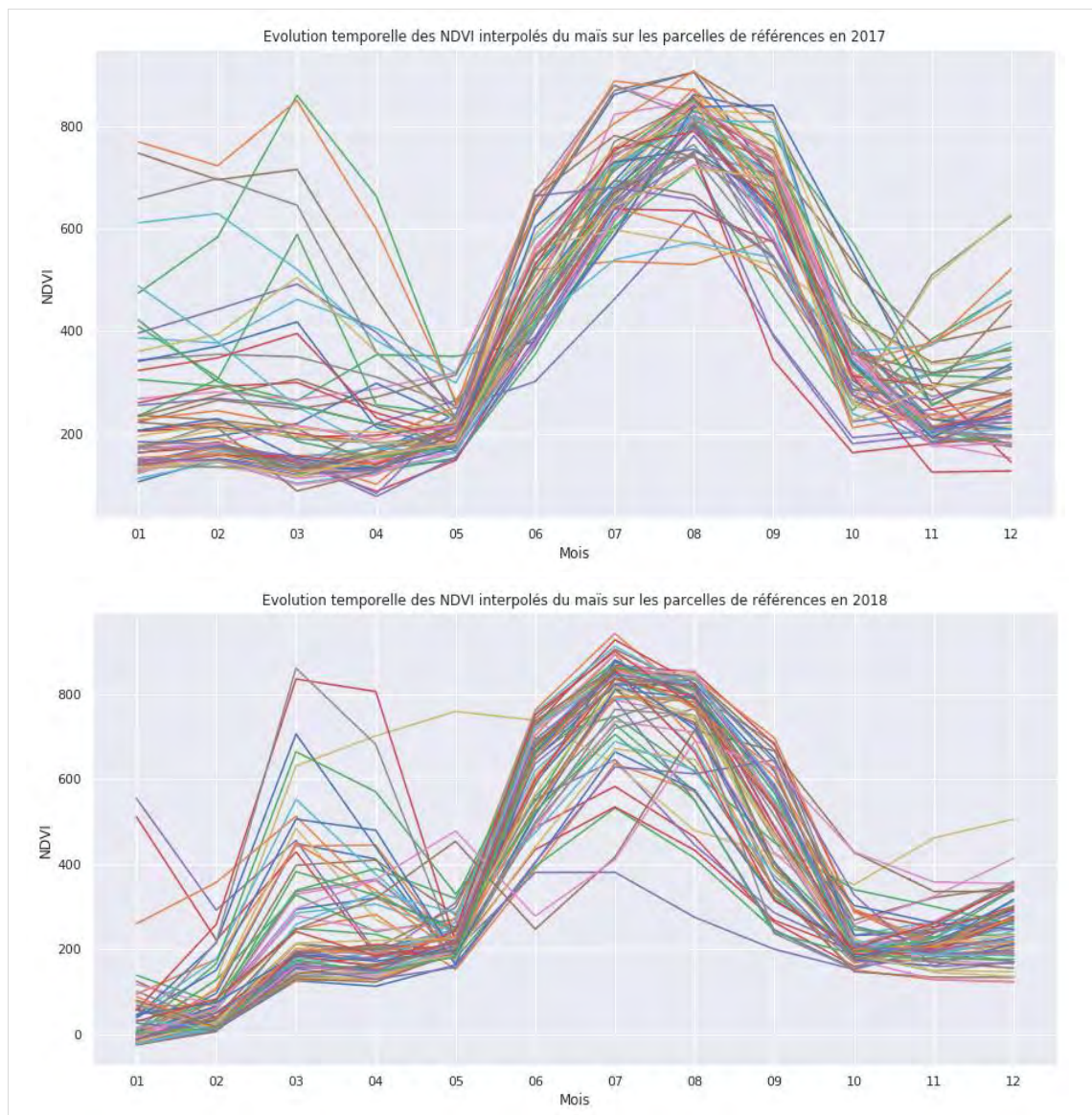


Figure 17 : Profils temporels des NDVI moyens des parcelles de maïs issues des données de référence

2. CLASSIFICATION EN COURS DE SAISON

La DRAAF a besoin de fournir une estimation de la production du maïs à l'échelle du département le plus tôt possible dans la saison. Ainsi la distinction entre le maïs irrigué et non irrigué doit être la plus précoce. Nous avons donc réalisé des classifications en cours de saison afin de déterminer à partir de quelle période cette distinction était possible.

D. INDICES SPECTRAUX ET APPORT DE LA TEMPORALITE

De nombreux travaux ont montré que l'utilisation d'indices spectraux tels que les NDVI, NDWI et IB permettait d'améliorer la classification des cultures. Dans cette étude, nous avons ajouté dans le processus de classification les cumuls temporels de ces indices et évalué leur apport.

1. DEFINITIONS

L'analyse des caractéristiques des indices spectraux a permis d'identifier les indices les plus adaptés à notre problématique.

INDICE DE VEGETATION

L'indice le plus connu et le plus utilisé est *l'indice de végétation par différence normalisé ou indice de Tucker*. Il s'agit d'un indicateur de l'état de santé, de la vigueur et de la quantité de végétation. Son expression est la suivante :

$$\rho_{PIR} : \text{réflectance dans la bande proche infrarouge} \quad NDVI = \frac{\rho_{PIR} - \rho_R}{\rho_{PIR} + \rho_R}$$

$$\rho_R : \text{réflectance dans la bande rouge}$$

Les valeurs du NDVI sont comprises entre -1 et +1, les valeurs négatives correspondant aux surfaces autres que les couverts végétaux, comme la neige, l'eau ou les nuages, pour lesquelles la réflectance dans le rouge est supérieure à celle du proche infrarouge. Pour les sols nus, les réflectances étant à peu près du même ordre de grandeur dans le rouge et le proche infrarouge, le NDVI présente des valeurs proches de 0. Les formations végétales quant à elles, ont des valeurs de NDVI positives, généralement comprises entre 0,1 et 0,7, les valeurs les plus élevées correspondant aux couverts les plus denses.

INDICE DE STRESS HYDRIQUE

Les indices de stress hydrique, comme le MSI (Hunt and Rock, 1989) ou le NDWI (Gao, 1996) sont calculés à partir des bandes spectrales du moyen infrarouge et du proche-infrarouge. Ces indices varient en fonction de la teneur en eau des feuilles. Ils permettent de déceler lorsque les végétaux sont en état de stress hydrique et sont par conséquent très utiles pour le suivi de la végétation en zone sèche. Son expression est la suivante :

$$\rho_{MIR} : \text{réflectance dans la bande infrarouge moyen} \quad NDWI = \frac{\rho_{PIR} - \rho_{MIR}}{\rho_{PIR} + \rho_{MIR}}$$

INDICE DE BRILLANCE

L'indice de brillance permet de distinguer les surfaces minérales des couvertures végétales. Les sols minéralisés et le bâti sont des surfaces réfléchissantes, contrairement à l'eau et aux sols très humides qui sont des surfaces plus absorbantes. La végétation obtient quant à elle une valeur moyenne comme indice de brillance. Son expression est la suivante :

$$IB = \sqrt{R * R + PIR * PIR}$$

INDICE DE DIFFERENCE DANS L'INFRAROUGE NORMALISE

L'indice de différence dans l'infrarouge normalisé (NDRE) est relativement similaire au NDVI avec la spécificité de se baser sur le red-edge, zone située entre le rouge et le proche-infrarouge, particulièrement sensible aux variations de teneurs en pigment foliaires notamment chlorophylliens. Le NDRE est donc potentiellement un bon indicateur de santé des végétaux. Son expression est la suivante:

$$NDRE = \frac{(NIR - RE)}{(NIR + RE)}$$

Le capteur Sentinel-2 offre la possibilité de calculer trois NDRE en utilisant la bande B8A associé aux bandes B5, B6 ou B7, donnant respectivement les indices NDRE-1, NDRE-2 et NDRE-3.

2. PROFILS TEMPORELS

La méthode que nous avons choisi de développer est basée sur la distinction entre les modalités de croissance des cultures, supposées être différentes entre les cultures irriguées et non irriguées. Cette différence est particulièrement notable pour les années sèches où le maïs irrigué connaît un meilleur développement que le maïs non irrigué. La comparaison des profils temporels moyens des indices spectraux du maïs irrigué et du non irrigué sur les deux années d'études, permet de sélectionner les indices spectraux les plus pertinents. Les profils temporels présentés ci-dessous sont tracés à partir des images interpolées générées par *iota*².

En observant les moyennes des NDVI du maïs à partir de la **Figure 18**, on constate deux phénomènes particuliers. Le premier concerne la comparaison entre les cultures. En 2017, le NDVI ne permet pas de distinguer le maïs irrigué du non irrigué car les profils sont quasiment identiques. Au contraire, on note un écart plus marqué entre les cultures en 2018. La différence entre les deux années est probablement liée aux conditions météorologiques et à la période de sécheresse de 2018 qui engendre probablement un apport en eau d'irrigation plus important, responsable des écarts entre les cultures. Le deuxième phénomène observé concerne la comparaison interannuelle des profils. En 2017, les valeurs de NDVI augmentent progressivement jusqu'au début du mois d'août, un léger plateau de 10 jours, puis les valeurs décroissent rapidement jusqu'à la mi-octobre. En 2018, la croissance se termine plus tôt, début juillet pour le maïs irrigué et mi-juillet pour le maïs non-irrigué. Le léger plateau devient très important pour l'irrigué et s'étend jusqu'à début septembre, alors qu'il disparaît entièrement pour le non irrigué. Ensuite, les valeurs décroissent plus ou moins fortement. Deux hypothèses se posent pour expliquer les comportements. L'une est liée à la différence de pratique, le recours à l'irrigation permet de maintenir les cultures plus vertes et plus longtemps, ce qui peut expliquer le plateau en 2018. L'autre hypothèse porte sur la qualité des images satellitaires. En effet, le manque de données claires en 2017 (Section I) ne permet pas de suivre précisément la dynamique des cultures malgré l'interpolation linéaire qui lisse les variations.

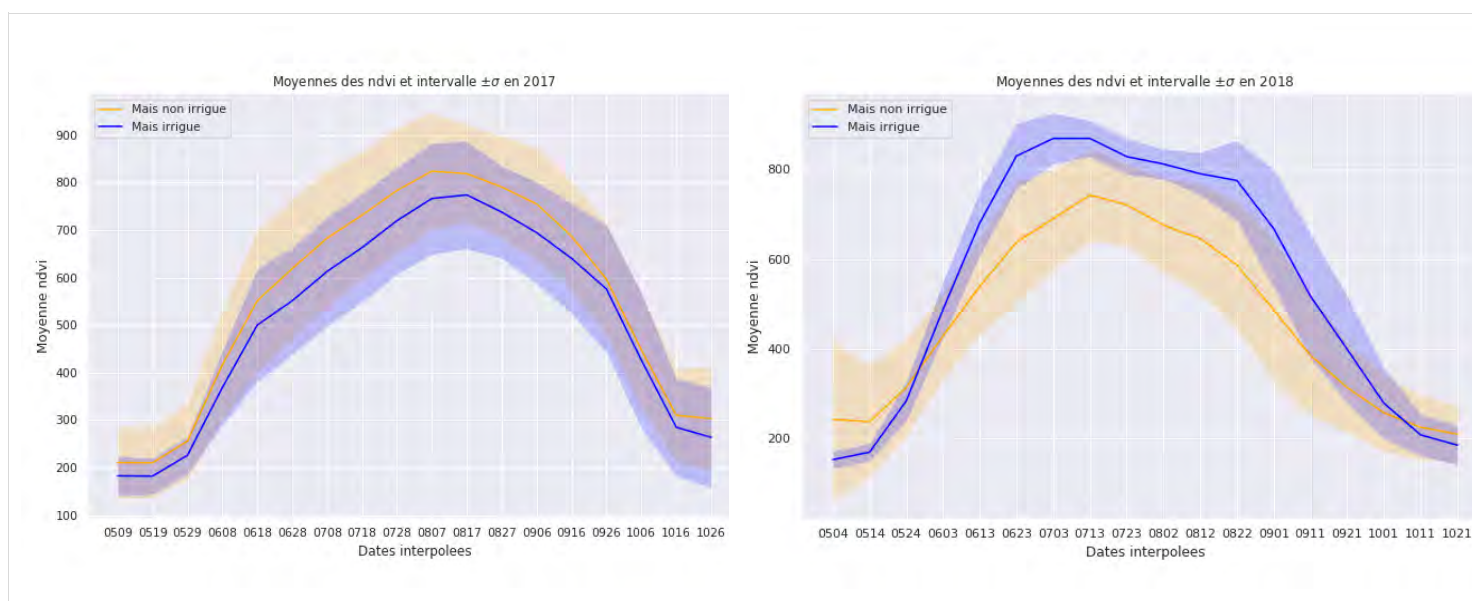


Figure 18 : Séries temporelles des NDVI du maïs irrigué et non irrigué en 2017 et 2018

En 2017, la **Figure 19** montre globalement que la teneur en eau des feuilles du maïs ne varie pas en fonction des pratiques d'irrigation. Contrairement à l'année précédente, 2018 est une année où le stress

hydrique du maïs est beaucoup plus important lorsque les parcelles ne sont pas irriguées, en comparaison aux parcelles irriguées. La comparaison interannuelle des profils de NDWI amène à se questionner sur les comportements d'irrigation (volume, dates) et à l'impact de l'irrigation sur la phénologie du maïs par rapport à un apport en eau naturel par la pluie.

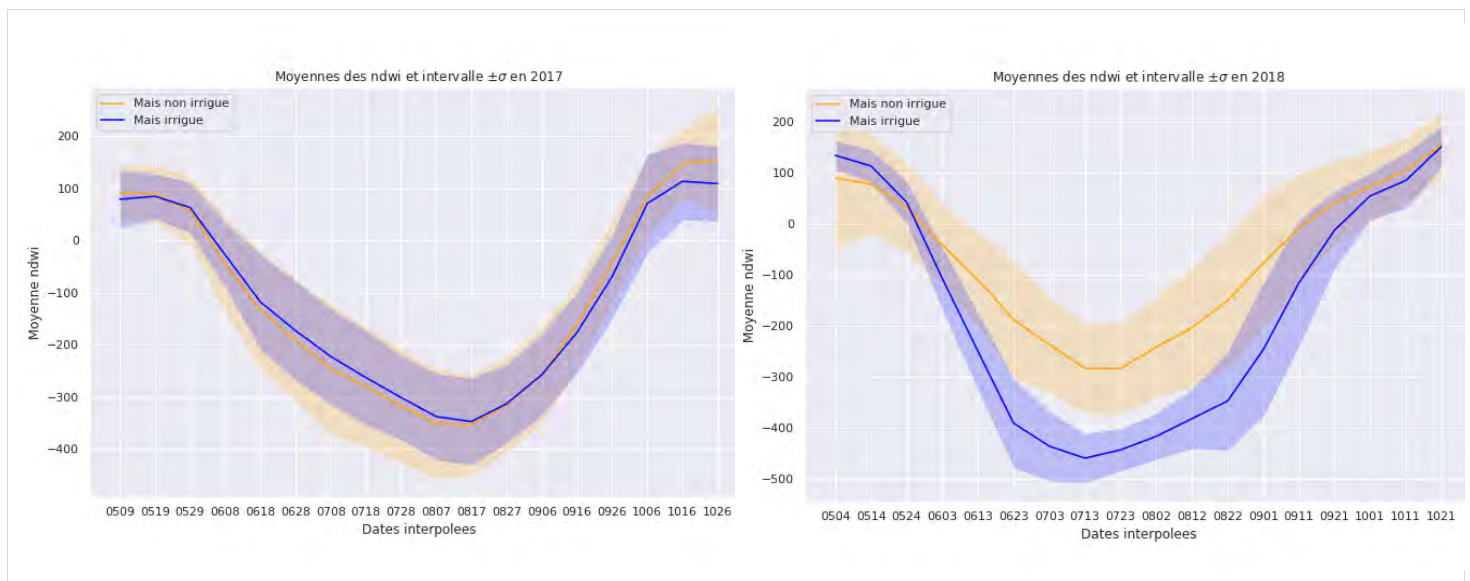


Figure 19 : Séries temporelles des NDWI du maïs irrigué et non irrigué en 2017 et 2018

En comparant les deux années, l'IB ne suit pas la même tendance. En 2017, on observe à partir de la [Figure 20](#) une forte décroissance de l'indice, avec des valeurs très légèrement supérieures pour le maïs non irrigué en cœur de saison. En 2018, la valeur de l'indice est plus stable avec une augmentation pendant les mois de l'été et une nette supériorité du maïs irrigué. Un pic pour le maïs irrigué uniquement s'observe en toute fin de saison. Pour les deux années, la décroissance progressive de l'IB peut s'expliquer par la sénescence du maïs. Les hausses observées en fin de saison sont sûrement dues aux repousses au sol après la récolte, avec un développement plus rapide sur les parcelles irriguées.

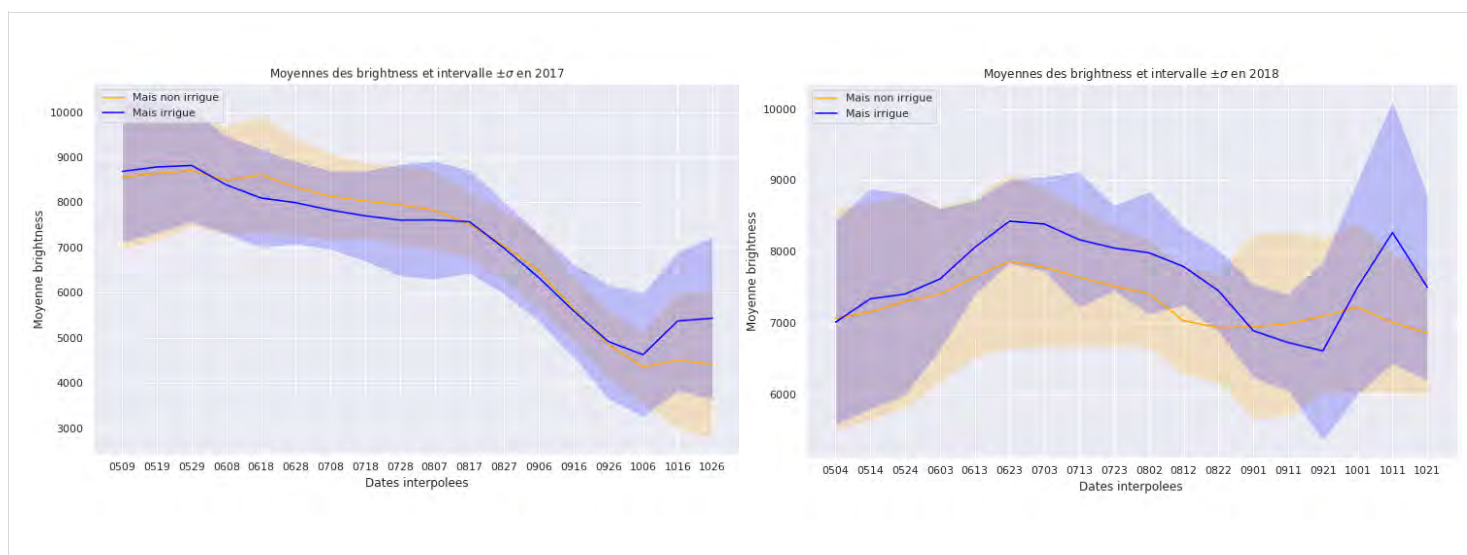


Figure 20 : Séries temporelles des indices de brillance du maïs irrigué et non irrigué en 2017 et 2018

Pour rappel, le NDRE-1 est calculé à partir des bandes 5 et 8A Sentinel-2. Son profil temporel est quasiment identique à celui du NDVI en 2017 et en 2018, son analyse est donc similaire ([Figure 21](#)).

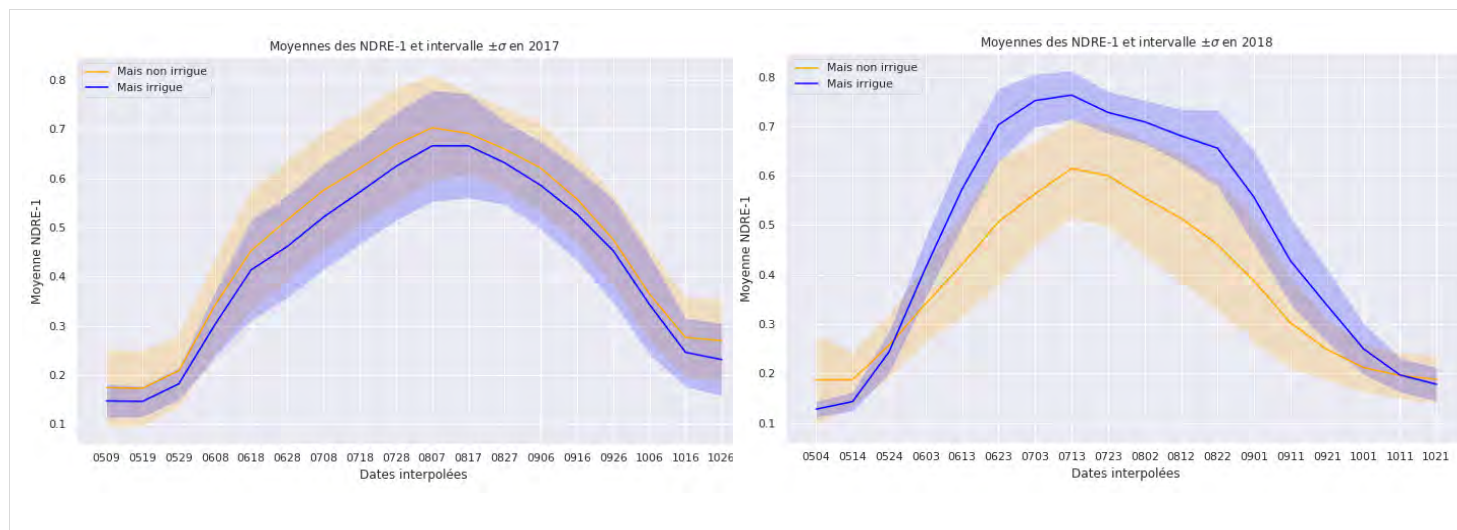


Figure 21 : Séries temporelles des NDRE-1 du maïs irrigué et non irrigué en 2017 et 2018

La tendance des profils temporels observée en 2017 se poursuit avec le NDRE-2 (bandes 6 et 8A) avec une homogénéité des courbes du maïs irrigué et du non irrigué sur la [Figure 22](#). En revanche, en 2018 l'indice permet de très bien distinguer les deux cultures, avec une progression des valeurs beaucoup plus importante pour le maïs irrigué, une valeur maximale obtenue mi-juillet et une décroissance progressive jusqu'à la fin de la saison.

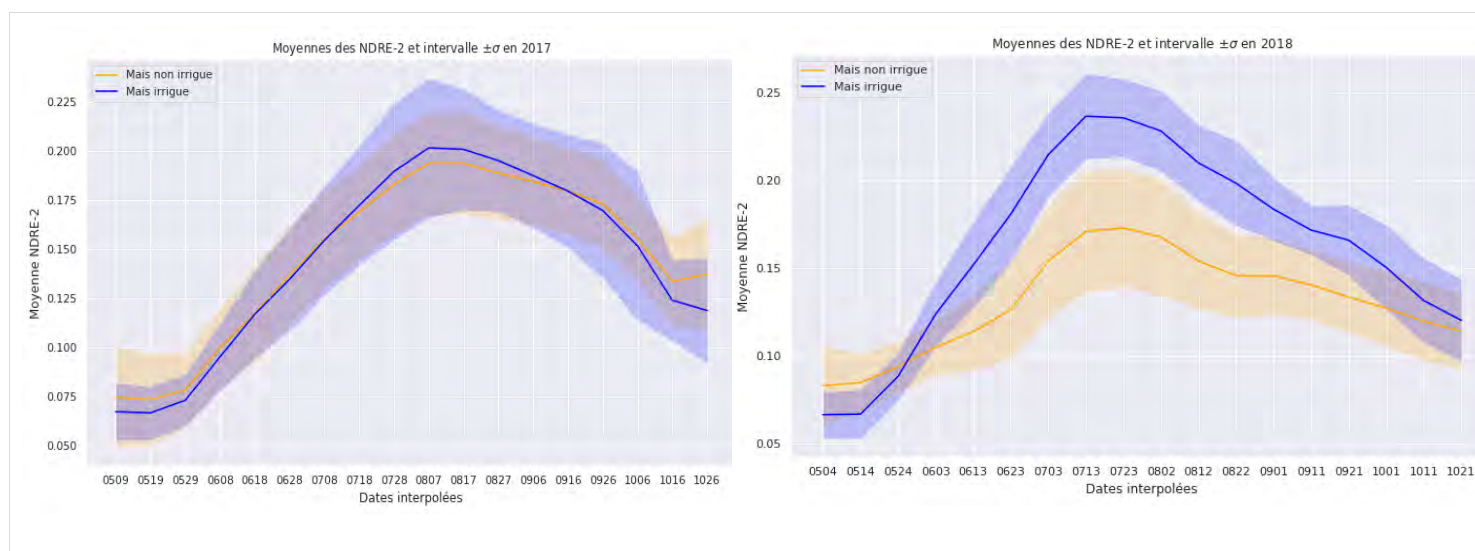


Figure 22 : Séries temporelles des NDRE-2 du maïs irrigué et non irrigué en 2017 et 2018

La [Figure 23](#) représentent les profils temporels du NDRE-3 (bandes 7 et 8A) en 2017 et 2018. On souligne une tendance complètement différente par rapport aux NDRE-1 et NDRE-2. Malgré certaines particularités des profils selon les cultures en 2018, la distinction n'est pas nette.

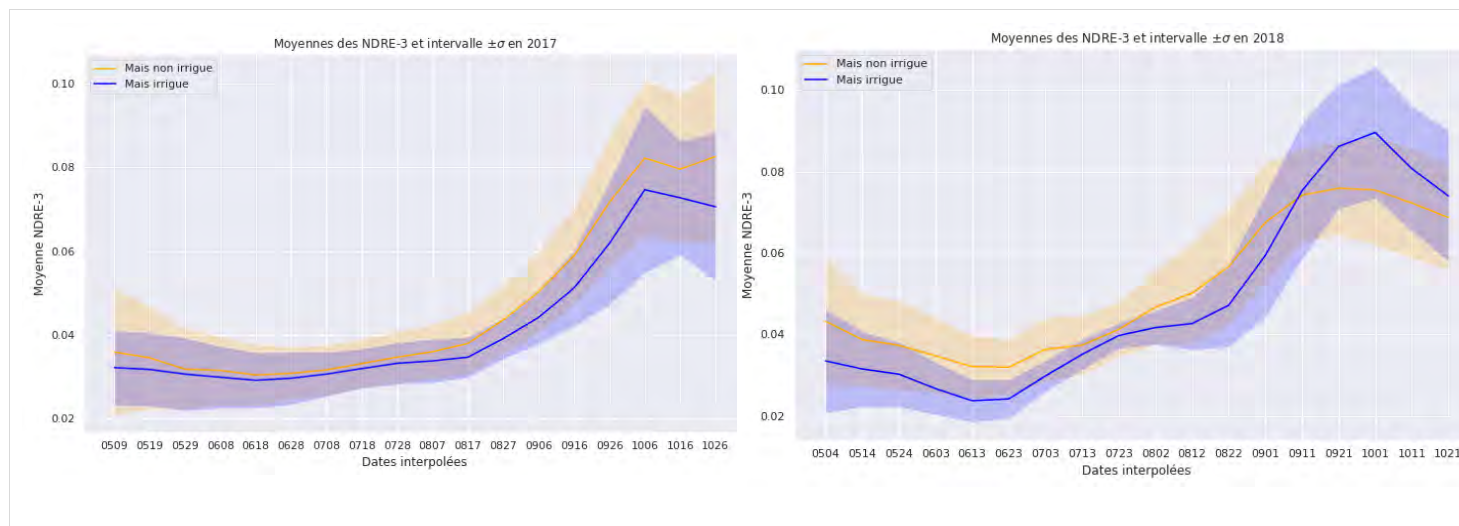


Figure 23 : Séries temporelles des NDRE-3 du maïs irrigué et non irrigué en 2017 et 2018

L'analyse des profils temporels permet d'en dégager deux conclusions. La première porte sur l'année 2017 pour laquelle on constate, quelque soit l'indice, une très forte similarité entre le maïs irrigué et le maïs non irrigué. La tendance est cohérente avec l'observation des données météorologiques et du diagramme de Gaussen, qui ne montraient pas de sécheresse significative durant l'été. La seconde se base sur la comparaison des profils en 2018 qui permet d'identifier la pertinence des indices. Les indices qui permettent une discrétisation nette sont le NDWI et le NDRE-2, le NDVI et le NDRE-1 présentent également des résultats encourageant. Les moyennes de l'indice de brillance et du NDRE-3 des deux cultures ont des profils temporels distincts, en revanche les écart-types importants constatés peuvent être source de confusion.

3. CUMULS TEMPORELS DES INDICES

Nous avons également analysé l'évolution dans le temps des cumuls temporels de ces indices avant de les intégrer dans le processus de classification (*Notebook du traitement, analyse et représentation graphique des indices spectraux, voir Annexe n°7*).

A partir des **Figures 24 à 29**, les conclusions tirées précédemment se vérifient à nouveau. Les caractéristiques de l'année 2017 ne sont pas adaptées à la discrétisation du maïs à partir d'images Sentinel-2. De plus, l'IB et le NDRE-3 traités en cumul ne sont pas ou peu pertinent pour distinguer les pratiques.

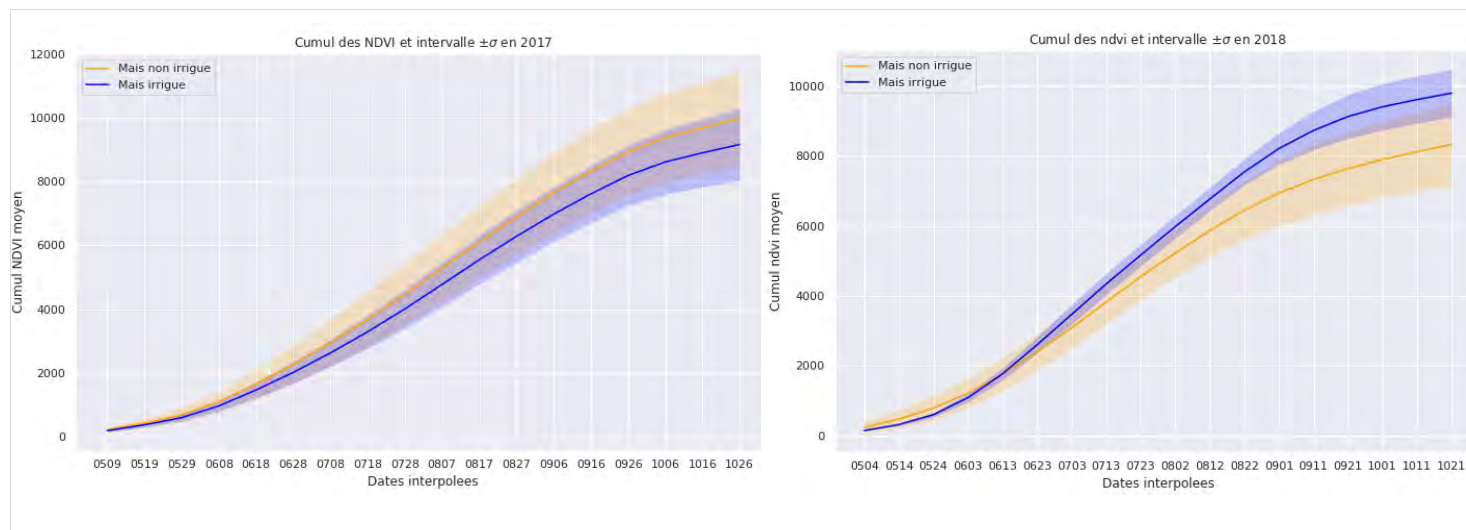


Figure 24 : Séries temporelles des cumuls de NDVI du maïs irrigué et non irrigué en 2017 et 2018

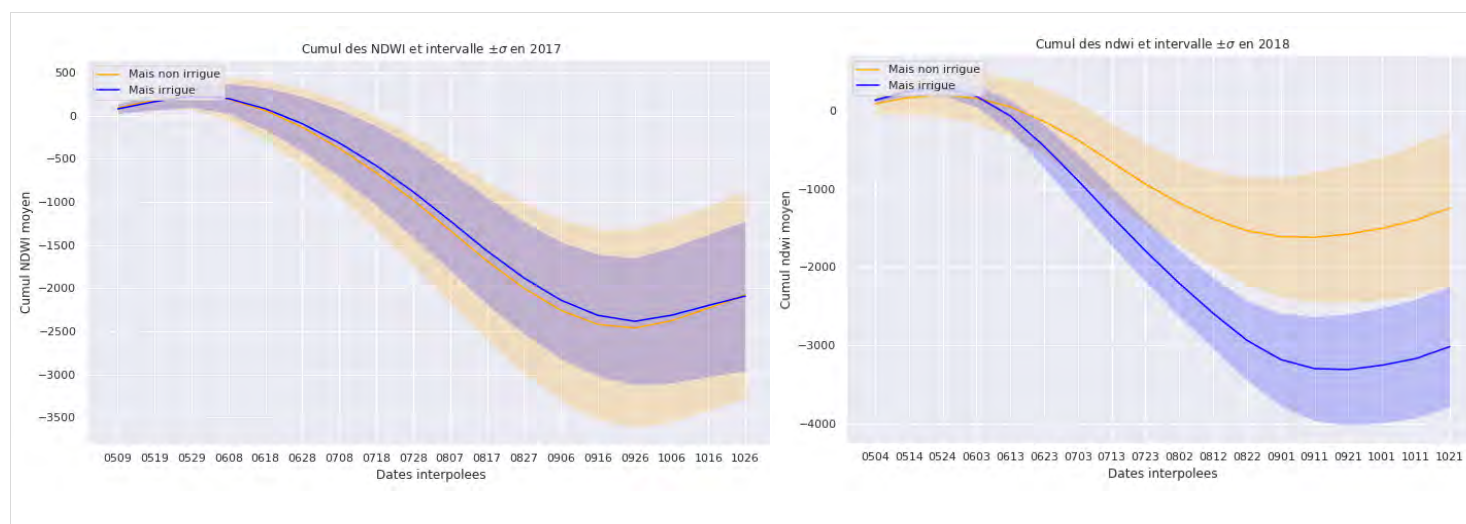


Figure 25 : Séries temporelles des cumuls de NDWI du maïs irrigué et non irrigué en 2017 et 2018

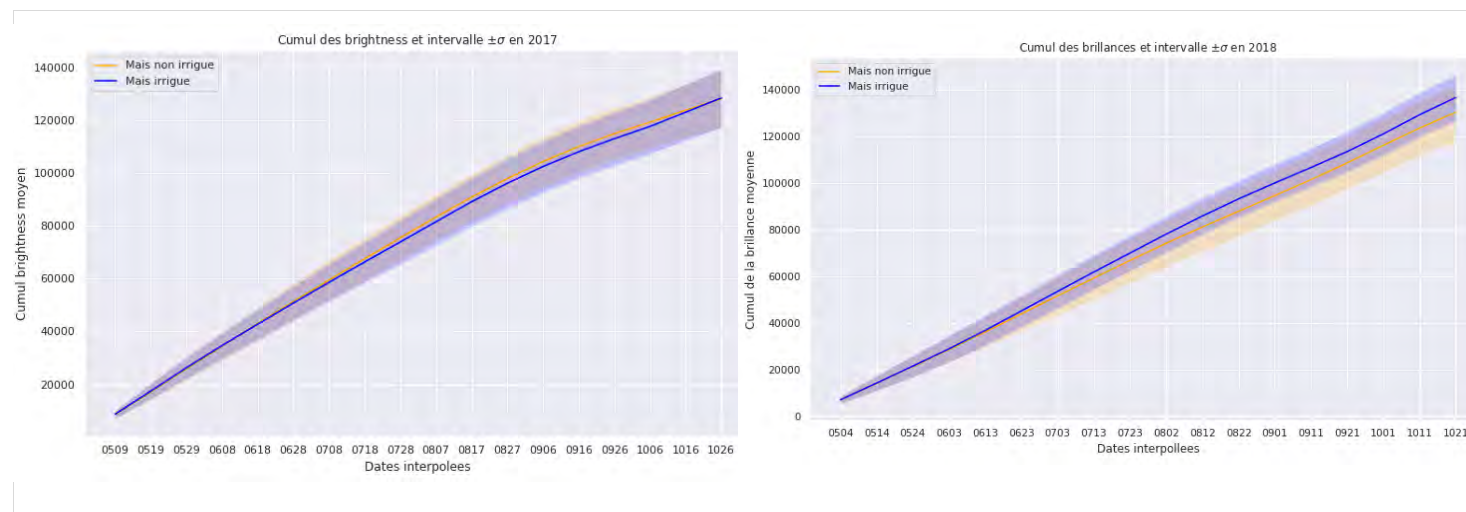


Figure 26 : Séries temporelles des cumuls d'indice de brillance du maïs irrigué et non irrigué en 2017 et 2018

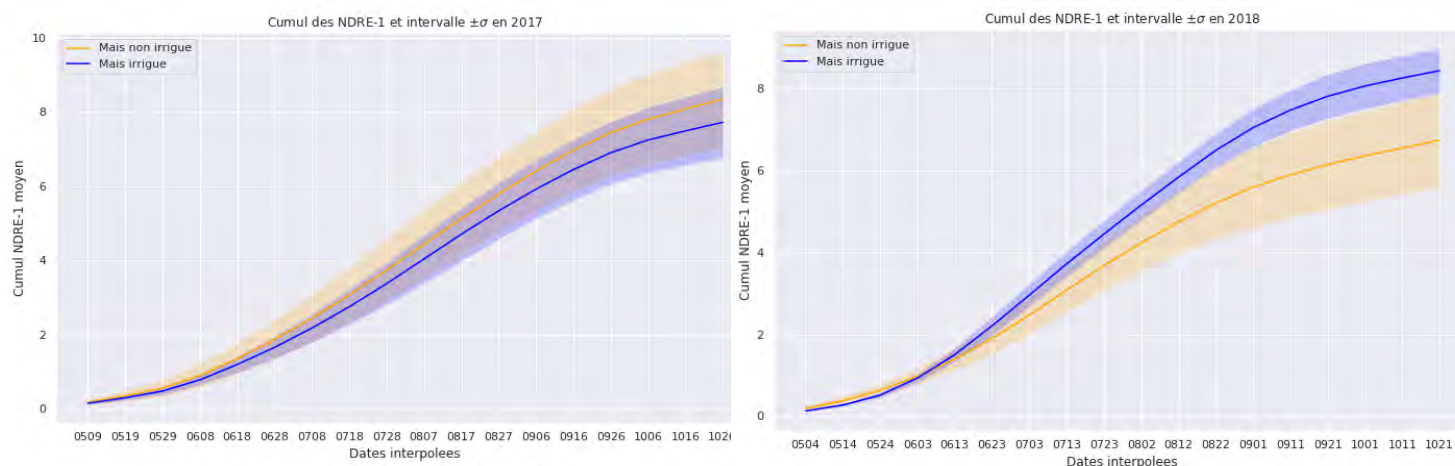


Figure 27 : Séries temporelles des cumuls de NDRE-1 du maïs irrigué et non irrigué en 2017 et 2018

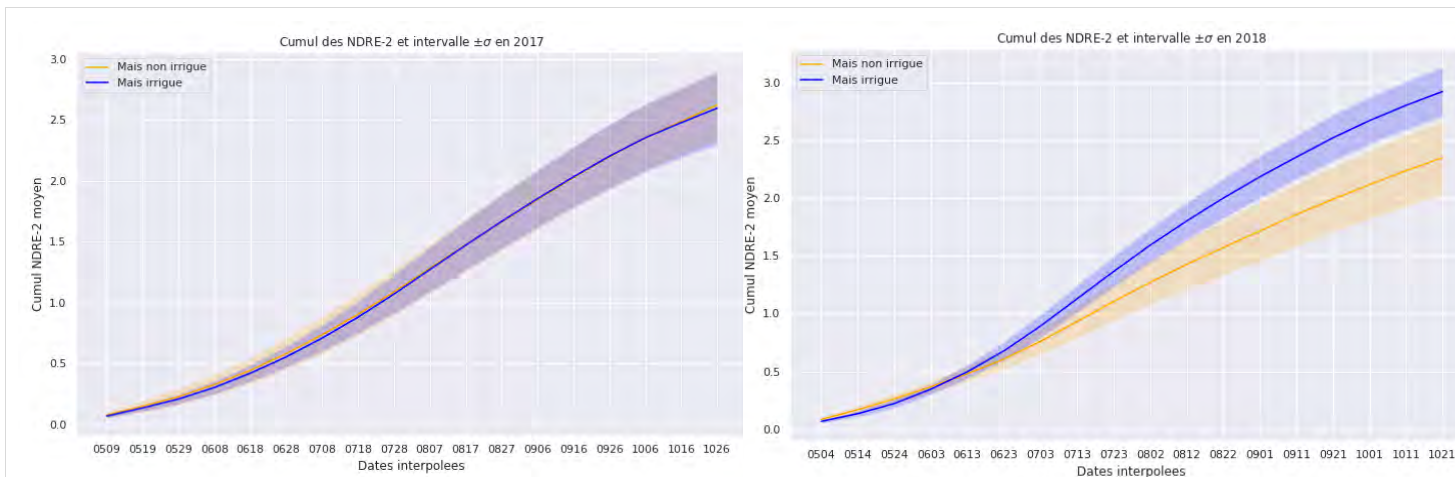


Figure 28 : Séries temporelles des cumuls de NDRE-2 du maïs irrigué et non irrigué en 2017 et 2018

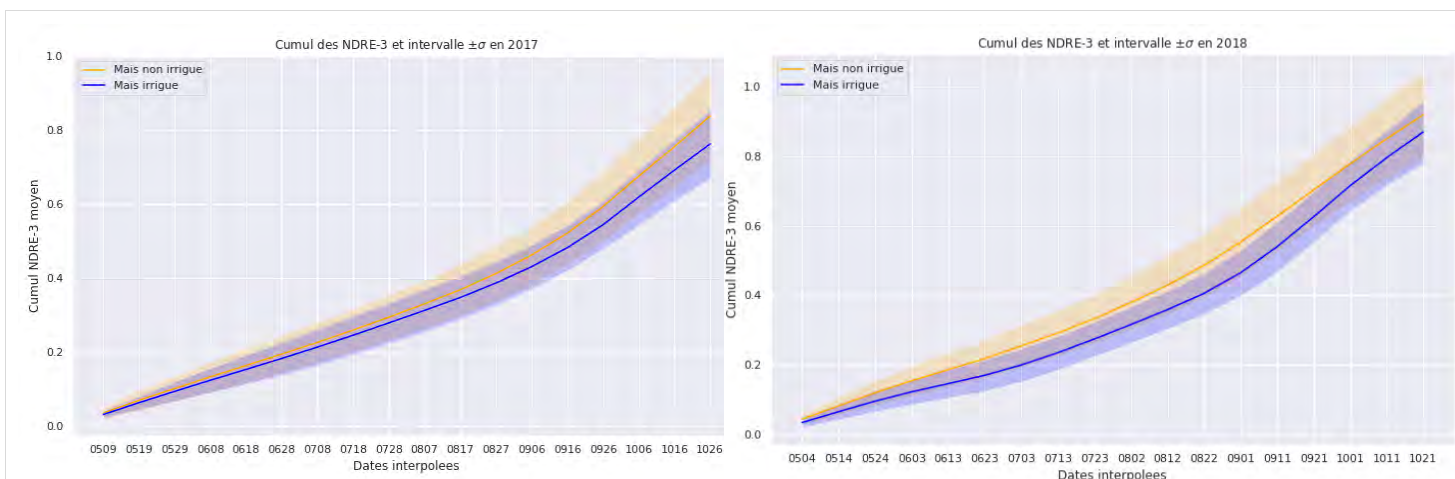


Figure 29 : Séries temporelles des cumuls de NDRE-3 du maïs irrigué et non irrigué en 2017 et 2018

Suite à cette analyse nous avons décidé d'utiliser dans les processus de classification les profils temporels du NDVI, NDWI, NDRE-1 et NDRE-2 en les traitant en cumuls. Deux types de cumuls sont calculés (*Illustration des indices calculés en cumul, voir Annexe n°8*) :

- un **cumul saisonnier** équivalent à la somme de chaque indice décadaire sur toute la saison de culture,
- des **cumuls mensuels** obtenus en sommant mensuellement chaque indice décadaire.

Le développement d'un code en python (*Code du calcul des indices traités en cumul, voir Annexe n°9*) m'a permis de calculer les cumuls mensuels et de saison pour chaque indice sous la forme d'un raster, sur l'emprise de chaque tuile (31UCP, 31UDP, 31TCN et 31TDN). Le traitement est réalisé à partir de la série temporelle interpolée extraite par *iota*². Ensuite, les bandes correspondantes aux NDVI, NDWI, déjà calculés par *iota*², puis les bandes B8A, B5 et B6, pour calculer les NDRE-1 et 2, sont isolées. Enfin, l'application OTB *BandMath* permet de réaliser une somme de toutes les bandes pour les cumuls de saison, et une somme toutes les trois dates pour les cumuls mensuels.

La **Figure 30** est une carte qui compare le NDVI traité en cumul saisonnier et en cumul mensuel. L'observation intra-parcellaire permet d'identifier des nuances plus importantes avec le cumul saisonnier, qui laisse deviner les rangs de culture ou les passages d'engins agricoles. En revanche, l'observation inter-parcellaire souligne des intensités plus marquées avec le cumul mensuel du mois de juin.

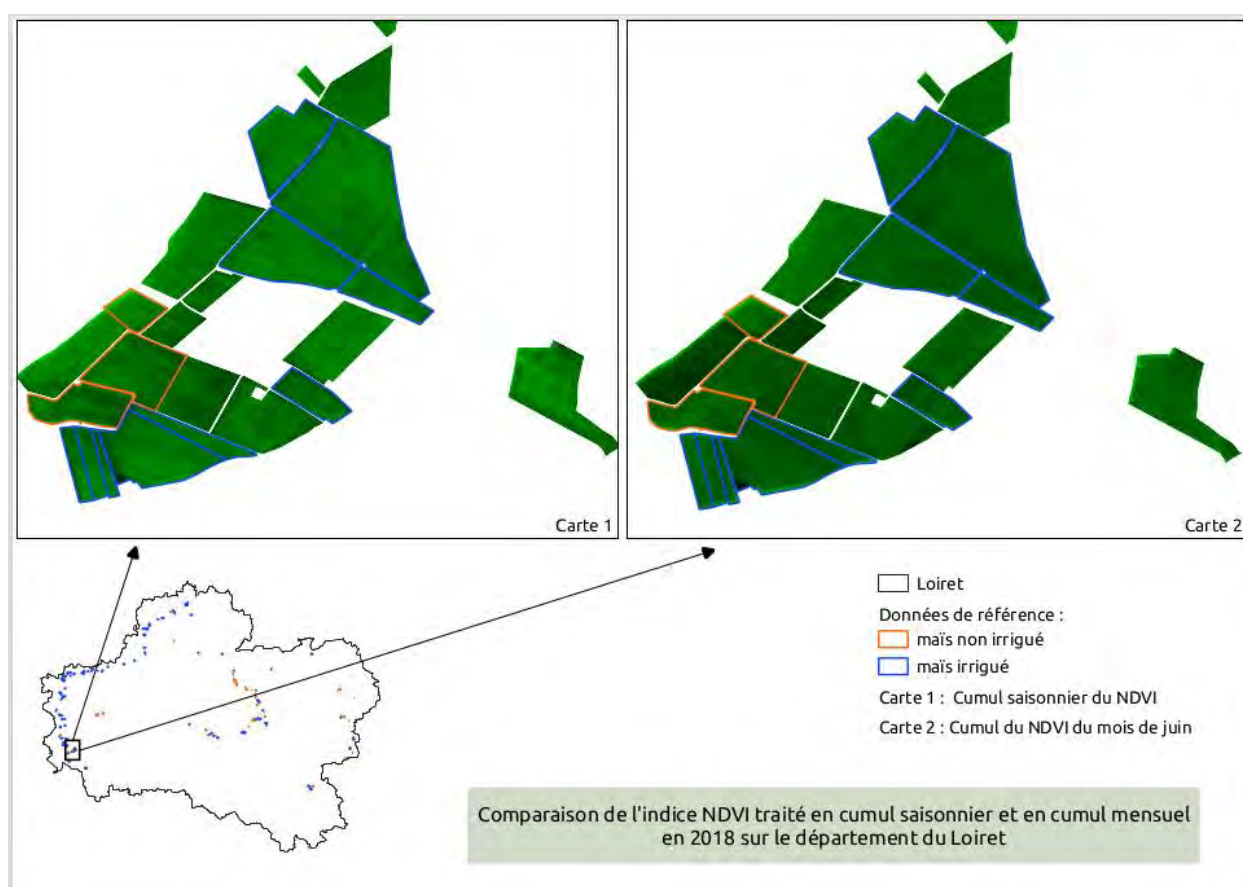


Figure 30 : Carte de comparaison du cumul de NDVI traité en cumul saisonnier et en cumul mensuel en 2018

E. NORMALISATION DES DONNÉES GLOBALSOILMAP

Les données GSM utilisées fournissent des informations sur le territoire français, à une résolution de 90 m sur la teneur en :

- la teneur en argile moyenne sur 10 cm entre 0 et 5 cm de profondeur,
- la teneur en argile moyenne sur 10 cm entre 5 et 15 cm de profondeur,
- la teneur en sable moyenne sur 10 cm entre 0 et 5 cm de profondeur,
- la teneur en sable moyenne sur 10 cm entre 5 et 15 cm de profondeur,
- la teneur en craie moyenne sur 10 cm entre 0 et 5 cm de profondeur,
- la teneur en craie moyenne sur 10 cm entre 5 et 15 cm de profondeur,
- la profondeur des sols.

Chaque raster est découpé selon les emprises des tuiles, et ré-échantillonné à 10 m de résolution. Les sept images sont ensuite concaténées pour simplifier leur utilisation. Les traitements sont réalisés avec les applications OTB *SuperImpose* et *ConcatenateImages*.

F. CONFIGURATION DES CLASSIFICATIONS

Différentes classifications ont été évaluées afin de déterminer la configuration (*Exemple d'un fichier de configuration iota^2 , voir Annexe n° 10*) optimale pour garantir de bons résultats tout en limitant les temps de calcul. Dans un premiers temps, les classifications sont réalisées en fin de saison et permettent d'évaluer la pertinence des indices en cumul saisonnier, puis en cumul mensuel. Ensuite, l'ajout des données *GlobalSoilMap* et des séries temporelles sont utilisés pour améliorer les résultats (*Architecture des données, voir Annexe n°11*). Dans un second temps, les classifications en cours de saison sont testées à partir des configurations les plus prometteuses.

G. ÉVALUATION DES PERFORMANCES DES CLASSIFICATIONS

Les classifications ont été évaluées à partir d'indices calculés par iota^2 , à partir de la matrice de confusion. Les indices de précision, de rappel et F-score sont fournis pour chaque classe, alors que les indices kappa et le coefficient *Overall Accuracy* (OA) pour l'ensemble des classes.

- La **précision** est la fraction des pixels correctement classés au regard de tous les pixels qui ont été attribués à cette classe par le modèle (bien et mal classé) :

$$\text{Précision } i = \frac{\text{Nb de pixels correctement attribués à la classe } i}{\text{Nb de pixels attribués à la classe } i \text{ par le modèle}}$$

- le **rappel** est la fraction des pixels correctement classés au regard de tous les pixels de cette classe selon les données de référence :

$$\text{Rappel } i = \frac{\text{Nb de pixels correctement attribués à la classe } i}{\text{Nb de pixels appartenant à la classe } i \text{ (données de référence)}}$$

- le **F-score** est une mesure qui combine la précision et le rappel :

$$F\text{-score} = \frac{2 \times \text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$$

- le **kappa** exprime la réduction proportionnelle de l'erreur obtenue par une classification, comparée à l'erreur obtenue par une classification au hasard. Un kappa de 0,8 signifie que 80% de la classification ne sont pas dus au hasard.
- l'**Overall Accuracy** mesure la proportion de pixels issus des données terrain, toutes classes confondues, qui ont été correctement classés par le classifieur.

H. SIMPLIFICATION DES RESULTATS

La classification produite est fournie à 10 m de résolution sur l'emprise des tuiles Sentinel-2 fusionnées (*Diagramme de flow des traitements des données, voir Annexe n°12*). Une fois la classification réalisée, plusieurs méthodes de traitement post-classification peuvent être appliquées :

- le vote majoritaire,
- la régularisation (proposé par *iota*²),
- la régularisation et le vote majoritaire.

1. VOTE MAJORITAIRE

Cette méthode consiste à attribuer le label majoritaire présent au sein d'une parcelle. L'utilisation de la librairie *rasterstats* sous python permet de réaliser des statistiques par segment en croisant une image raster (classification) et une couche vecteur (RPG dans notre cas). Ainsi, on obtient pour chaque segment (parcelle RPG) le nombre de pixels par label. Le label majoritaire est ensuite attribué à tous les pixels de la parcelle. Le résultat est une classification sous format vecteur.

2. REGULARISATION

La régularisation (proposée par *iota*²) consiste à dégrader la résolution de l'image de 10 m à 20 m, en fusionnant quatre pixels et en attribuant au pixel de 20 m le label majoritaire. Contrairement au vote majoritaire, la régularisation ne permet pas d'obtenir un label unique par parcelle. En revanche, la régularisation supprime toute information parasite ou dégradée subite par l'image lors de son acquisition ou de son enregistrement. L'altération modifie la valeur de certains pixels répartis

aléatoirement en faisant passer l'intensité à la valeur minimum ou maximum (respectivement 0 et 255), ce qui lui donne le nom de bruit poivre et sel. Le résultat est une classification sous format raster.

3. REGULARISATION ET VOTE MAJORITAIRE

Cette méthode est une combinaison des deux précédentes : une régularisation est d'abord appliquée à la classification suivie d'un vote majoritaire. Le résultat est une classification sous format vecteur.

Ces trois post-traitements ont été évalués en comparant les superficies de maïs irrigué et non irrigué obtenues avec chaque méthode avec celles issues des données de référence et des données Agreste fournies à l'échelle du département.

Section IV - RESULTATS & ANALYSE

La méthode de classification suivie donne lieu à différents résultats. Leur présentation et leur analyse permettent d'évaluer le potentiel des images Sentinel-2 et des indices spectraux pour la cartographie du maïs en fin et en cours de saison.

- Evaluation de la méthode d'échantillonnage et de la période d'observation
- Evaluation du classifieur
- Apport des cumuls saisonniers et mensuels
- Apport des données *GlobalSoilMap*
- Apport des séries temporelles
- Traitement en cours de saison
- Simplification et validation des résultats

A. EVALUATION DE LA METHODE D'ECHANTILLONNAGE ET DE LA PERIODE D'OBSERVATION

Les résultats obtenus montrent d'importants écarts entre les deux années d'études avec des scores nettement supérieurs, quelle que soit la culture, en 2018. On constate aussi des scores plus élevés pour le maïs irrigué. La méthode d'échantillonnage qui consiste à équilibrer le nombre de pixels entre les deux cultures, ainsi que l'utilisation de la période centrée sur la saison, conduisent aux meilleurs résultats.

Les résultats peu satisfaisants obtenus pour l'année 2017 (**Figure 31**) sont cohérents avec l'analyse météorologique réalisée dans la partie 1-II-B. L'absence de sécheresse marquée explique la similarité de croissance des deux cultures et donc une difficulté pour le classifieur à les discriminer. De plus, la couverture nuageuse très importante observée, peut diminuer la fiabilité des données satellitaires malgré l'interpolation linéaire générée par *iota*². En accord avec la DRAAF Centre-Val de Loire, la classification du maïs en 2017 a été écartée pour se concentrer sur l'année 2018 et les classifications en cours de saison.

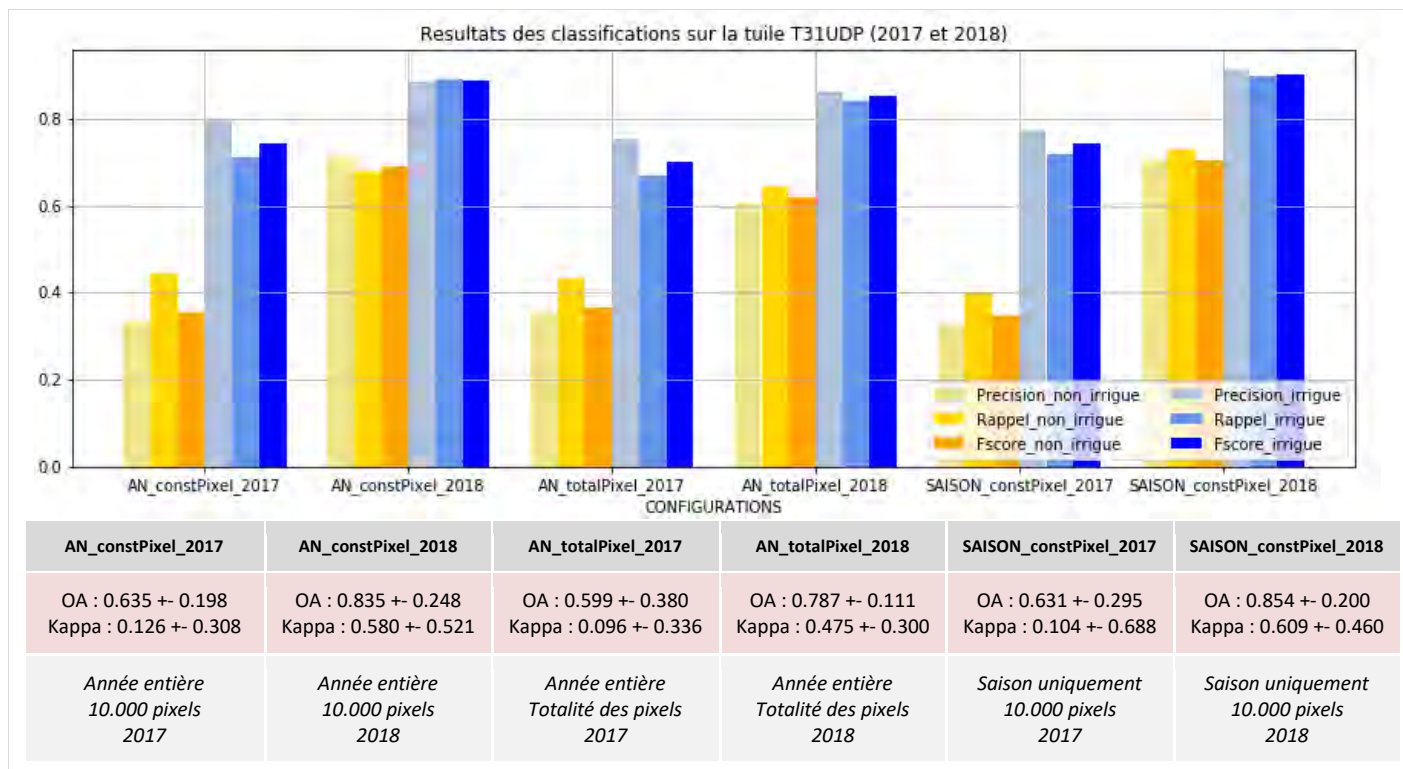


Figure 31 : Comparaison des scores des classifications en 2017 et 2018 selon la méthode d'échantillonnage et la période d'observation

Des tests ont été réalisés pour identifier le nombre de tirages aléatoires optimal. La différence entre les scores issus d'un seul et de cinq tirages aléatoires, avec la même configuration, obtiennent des scores significativement différents. Cette différence devient non significative entre 5 et 10 tirages. Nous avons donc décidé de réaliser cinq tirages aléatoires par configuration.

B. EVALUATION DU CLASSIFIEUR

Comme vu précédemment, le classifieur *Random Forest* détermine le nombre d'arbres décisionnels qui lui paraît optimal. Dans notre cas, le classifieur fixe à dix le nombre d'arbres. Pour évaluer la pertinence de ce paramètre, les scores ont été comparés à ceux obtenus avec le classifieur *Shark Random Forest* avec le nombre d'arbres fixé à cent. Les résultats des deux classifieurs ne présentent pas d'écart significatif, le choix de l'algorithme porte donc sur *Random Forest*.

Ces tests préliminaires nous ont permis de déterminer la configuration pour le déploiement de la méthode sur les quatre tuiles. La configuration choisie est la suivante :

- Période choisie : la saison,
- Nombre de pixels par label du jeu d'apprentissage fixé à 10.000,
- 5 tirages aléatoires,
- Classifieur *Random Forest*
- Cumuls saisonniers et cumuls mensuels.

C. APPORT DES CUMULS SAISONNIERS ET MENSUELS

La pertinence de l'utilisation des cumuls d'indices de saison par rapport aux cumuls mensuels a été analysée en comparant les scores des classifications obtenus avec les différentes primitives.

La **Figure 32** présente les scores des classifications réalisées uniquement avec les cumuls saisonniers à partir de chaque indice utilisés séparément (configurations : ndre1, ndre2, ndvi et ndwi), puis conjointement (configuration : index). Le meilleur résultat est obtenu en utilisant tous les indices (index). Il est toutefois important de noter que l'apprentissage des modèles est réalisé sur une seule (ndre1, ndre2, ndvi, ndwi) ou quatre images (index).

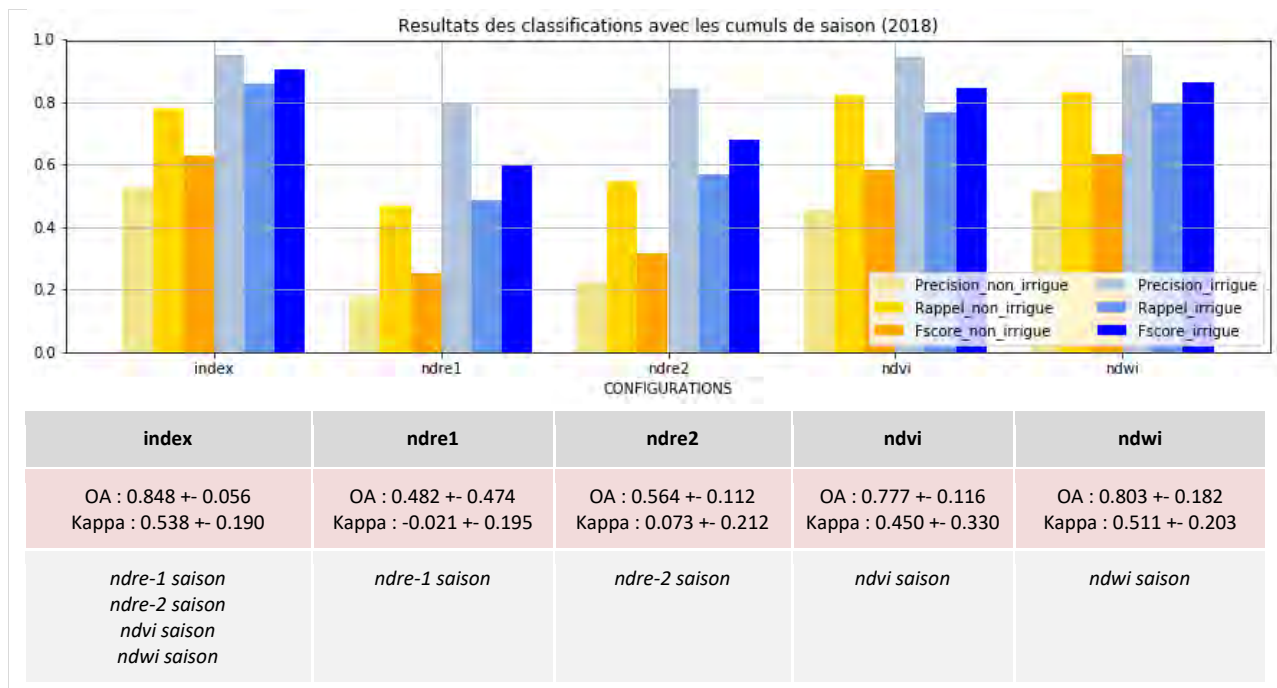


Figure 32 : Comparaison des scores des classifications en 2018 selon les cumuls saisonniers des indices spectraux

L'utilisation de la commande *time* (système d'exploitation *Ubuntu*) au lancement du modèle, permet de mesurer les temps de traitement (**Figure 33**). Une classification réalisée à partir d'une seule image se déroule sur environ 6 heures (temps écoulé). Le temps processus (temps CPU), définit comme le temps passé par un programme sur le processeur, est divisé entre une partie utilisateur et une partie système. Le temps CPU utilisateur, c'est-à-dire le temps passé à exécuter du code en mode utilisateur est de 51 h. Le temps CPU système, soit le temps passé par le noyau en mode système pour le processus, est de 24 minutes. Ce dernier est l'unité de mesure la plus pertinente pour comparer les temps de traitement en fonction du volume de données en *input*. Le modèle génère 26 Go de résultats.

```
real    369m58.197s
user    3080m27.389s
sys     24m15.020s
```

Figure 33 : Temps de traitement pour une classification avec un cumul saisonnier

Le même travail a également été réalisé avec les indices mensuels utilisés individuellement (configurations : ndre1M, ndre2M, ndviM, ndwiM) et conjointement (configuration : indexM). Le

meilleur résultat est à nouveau obtenu avec tous les indices utilisés conjointement (**Figure 34**). En revanche, on obtient un meilleur résultat que la configuration précédente qui n'utilise que le cumul saisonnier. Les cumuls mensuels apportent donc une information supplémentaire. En effet les cumuls mensuels permettent de prendre en compte des différences de croissance observables en cours de saison que le cumul saisonnier finit par gommer.

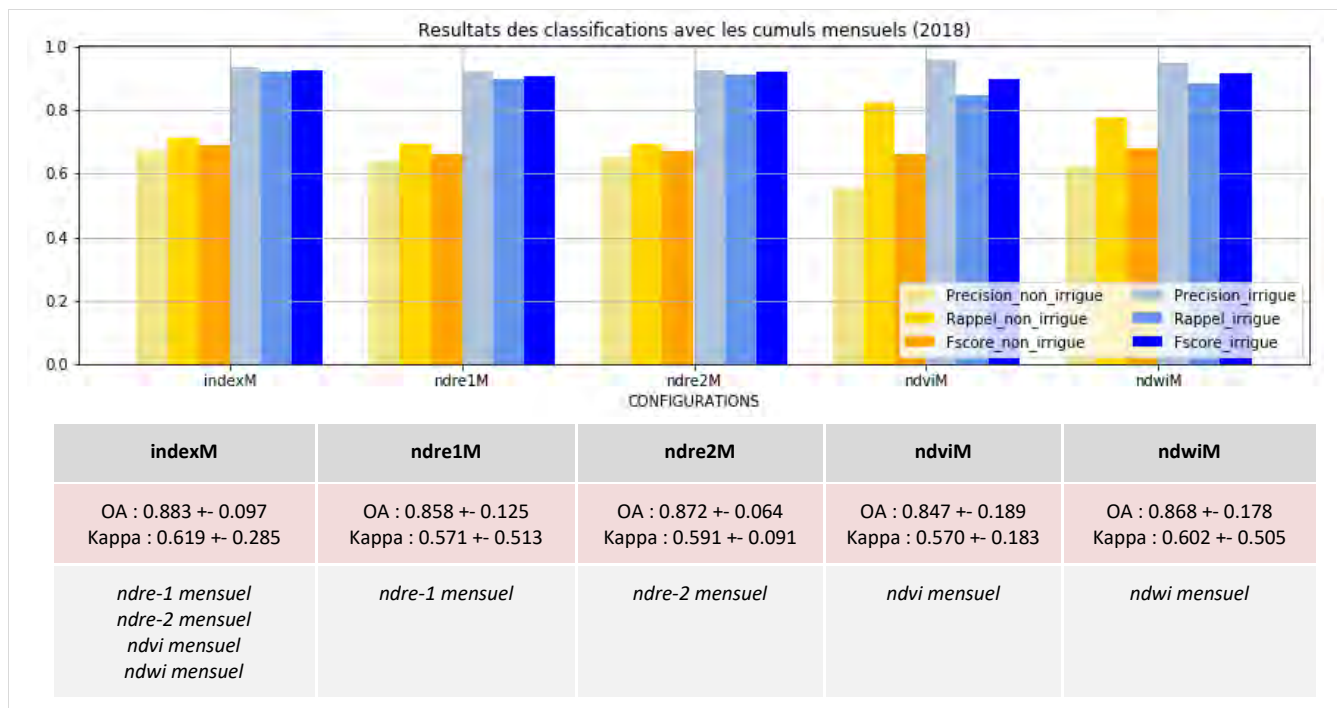


Figure 34 : Comparaison des scores des classifications en 2018 selon les cumuls mensuels des indices spectraux

D. APPORT DES DONNEES GLOBALSOILMAP

Les caractéristiques du sol fournies par les données GSM ne permettent pas de discriminer de manière satisfaisante les pratiques d'irrigation. La **Figure 35** présente les scores obtenus avec trois configurations: les données GSM uniquement (configuration : GSM), tous les indices en cumuls mensuels sans (configuration : indexM) et avec GSM (configuration : indexM_GSM). On constate que les données GSM n'apportent pas d'amélioration aux résultats. Cependant, les informations sur la granulométrie n'ayant pas été utilisées, elles présentent une piste de recherche.

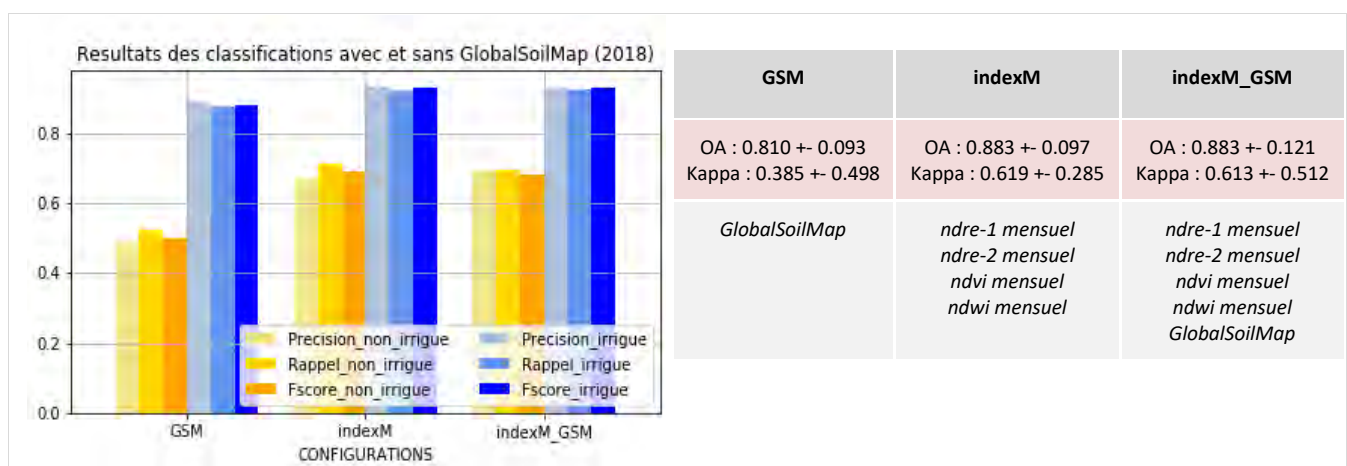


Figure 35 : Comparaison des scores des classifications en 2018 avec ou sans les données GSM

E. APPORT DES SERIES TEMPORELLES

Les informations supplémentaires contenues dans la série temporelle interpolée tous les 10 jours, comme les niveaux de réflectance en fonction de la longueur d'onde, n'apporte pas d'amélioration au modèle. La **Figure 36** présente les scores obtenus avec cinq configurations avec : tous les indices en cumuls mensuels (configuration : indexM), la série temporelle interpolée (configuration : stack), puis la série temporelle interpolée avec les données GSM (configuration : stack_GSM), puis avec tous les indices en cumuls mensuels (configuration : stack_indexM) et une combinaison des trois (configuration : stack_indexM_GSM). L'apport de données supplémentaires ne permet pas d'améliorer les résultats. Les scores obtenus selon les configurations sont sensiblement identiques, à l'exception de celles réalisées avec les données GSM qui dégradent les F-scores.

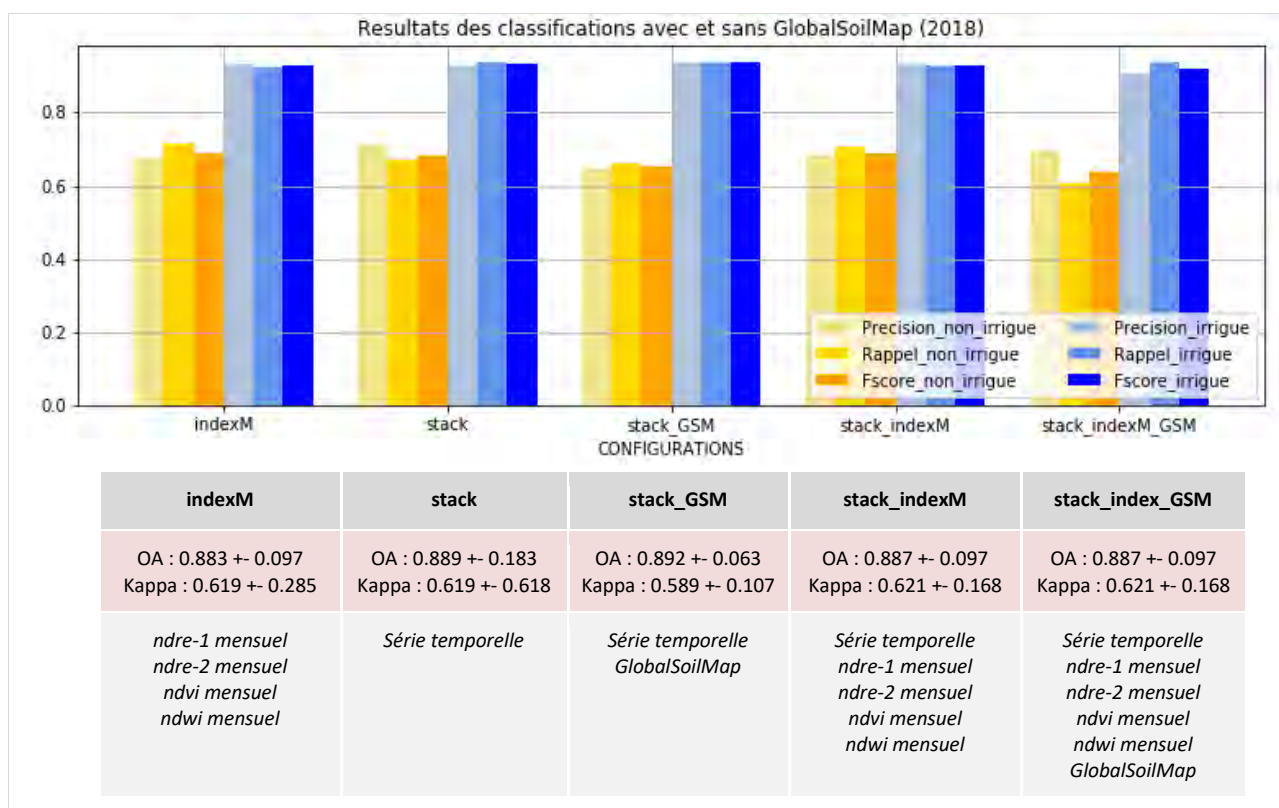


Figure 36 : Comparaison des scores des classifications en 2018 avec ou sans la série temporelle

Par ailleurs, on note une augmentation importante des temps de traitement et des volumes de données générés. La classification réalisée à partir de la série temporelle interpolée (stack) a nécessité 12h (temps réel écoulé) de traitement, 28h de temps CPU utilisateur et 2h de temps CPU système. Le modèle a créé 129 Go de résultats.

```
real 737m42.851s
user 1704m45.256s
sys 169m4.143s
```

Figure 37 : Temps de traitement pour une classification avec une série temporelle de 6 mois

L'observation des F-scores révèle que le maïs non irrigué est moins bien classé que le maïs irrigué. Les matrices de confusion (**Tableau 3 et 4**) permettent d'analyser les confusions entre les classes. Leur

analyse révèle que le classifieur classe une partie (environ $\frac{1}{3}$) des pixels de maïs non-irrigué en maïs irrigué (erreur d'omission de 4 600 pixels pour la configuration stack, et 3 996 pixels pour la configuration indexM). Mais une partie (6%) des pixels de maïs irrigué (erreur de commission = 3 764 pixels → stack, 4 886 pixels → indexM) sont classés en maïs non irrigué. Les erreurs d'omission et de commission étant très proches, au final les proportions entre les deux cultures issues des classifications sont similaires aux proportions observées sur le terrain.

Matrice de confusion Configuration : stack	pixels classés non-irrigué	pixels classés irrigué	Nombre de pixels par culture issu des parcelles de référence	Proportion entre les cultures issues des parcelles de référence
Maïs non- irrigué	8 917.6	4 600.4	13 518.0	18.3 %
Maïs irrigué	3 764.0	56 788.0	60 552.0	81.7 %
Nombre de pixels par culture obtenus par le classifieur	12 681.6	61 388.4	74 070.0	-
Proportion entre les cultures obtenues par le classifieur	17.1 %	82.9 %	-	100 %

Tableau 3 : Matrice de confusion de la configuration "stack" avec la série temporelle

Matrice de confusion Configuration : indexM	pixels classés non-irrigué	pixels classés irrigué	Nombre de pixels par culture issu des parcelles de référence	Proportion entre les cultures issues des parcelles de référence
Maïs non- irrigué	9 872.4	3 996.4	13 868.8	18.2 %
Maïs irrigué	4 886.2	57 244.8	62 131.0	81.8 %
Nombre de pixels par culture obtenus par le classifieur	14 758.6	61 241.2	75 999.8	-
Proportion entre les cultures obtenues par le classifieur	19.4 %	80.6 %	-	100 %

Tableau 4 : Matrice de confusion de la configuration "indexM" avec tous les indices traités en cumuls mensuels

F. TRAITEMENT EN COURS DE SAISON

Les classifications « cours de saison » intègrent les cumuls mensuels de tous les indices. L'analyse des résultats (Figure 38), révèle que la distinction entre les deux cultures peut être faite dès le mois de juillet (7 juillet) avec des scores pratiquement identiques à ceux du mois d'octobre. Pour comprendre la distinction dès le mois de mai, une meilleure connaissance des pratiques d'irrigation est nécessaire. L'irrigation au moment de la levée du maïs est donc une hypothèse à approfondir avec la DRAAF et la Chambre d'Agriculture.

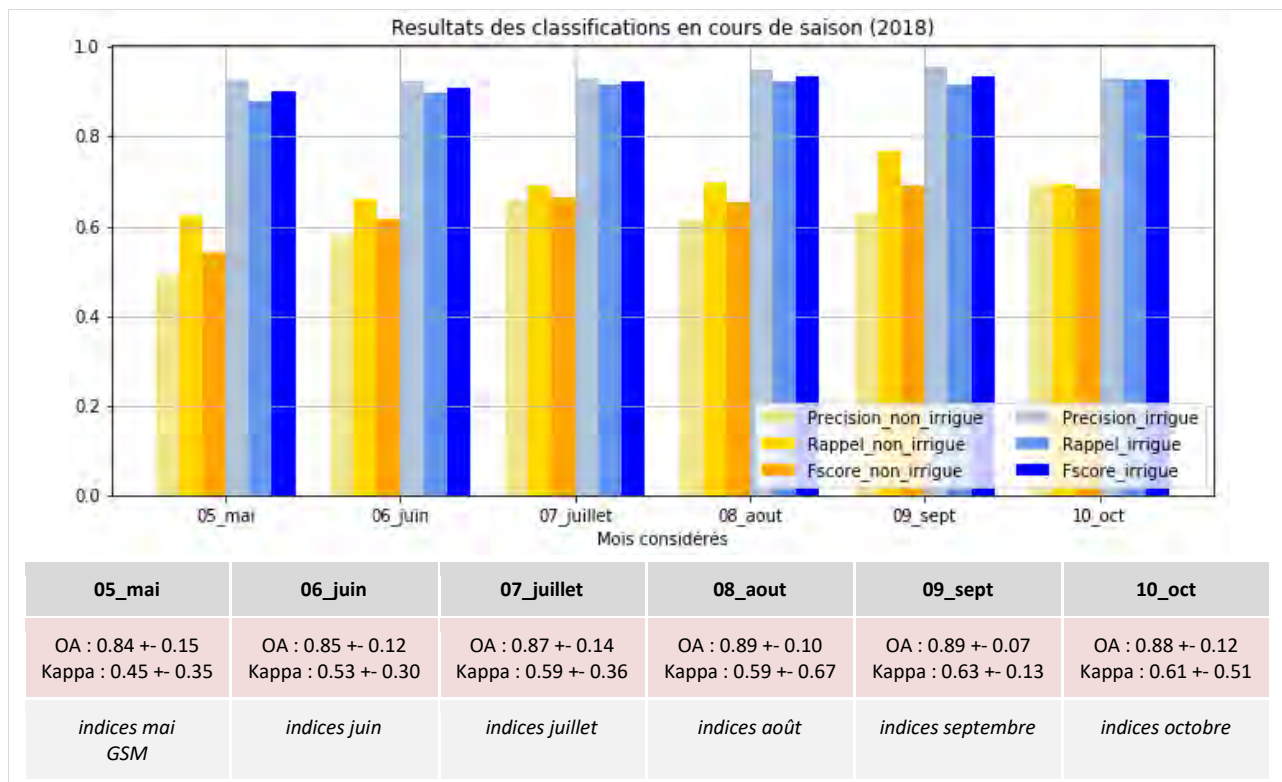


Figure 38 : Comparaison mois par mois des scores des classifications en 2018

G. SIMPLIFICATION ET VALIDATION DES RESULTATS

L'évaluation des post-traitements (Section III-H) a été réalisée sur la classification basée sur la série temporelle (configuration : stack). Pour rappel, cinq tirages aléatoires sont réalisés pour chaque configuration et donnent lieu à cinq classifications. La validation des résultats est réalisée sur la classification fusionnée des cinq modèles (*otbApplication FusionOfClassifications*) et la classification ayant obtenu les meilleurs scores (*Classif_Seed3.tif*). Les méthodes de simplification sont uniquement appliquées sur cette dernière.

1. SUPERFICIES REELLES

Le **Tableau 5** est extrait des statistiques fournies par l'Agreste, à l'échelle du département, au 1er décembre 2018. Le maïs grain humide, destiné à l'alimentation porcine, est récolté plus humide, avec des modes de conservation adaptés (ensilage). Aujourd'hui, la DRAAF n'a pas les moyens d'estimer ses superficies selon les pratiques d'irrigation et considère que les surfaces sont équitablement réparties entre le maïs grain humide irrigué et le non irrigué sur le département du Loiret. Le maïs fourrage et ensilage, quant à lui, est considéré comme non irrigué. On obtient donc une superficie pour l'ensemble du département **de 23 720 ha de maïs irrigué et 9 350 ha de maïs non irrigué**, en 2018 (*Statistiques Agreste du maïs en 2017, voir Annexe n°13*).

	Superficie (ha)	Rendement (q/ha)	Production (q)
Maïs grain (non humide)	28 770	104,3	3 001 320
<i>dont maïs grain irrigué</i>	<i>23 520</i>	<i>116,0</i>	<i>2 728 320</i>
<i>dont maïs grain non irrigué</i>	<i>5 250</i>	<i>52,0</i>	<i>273 000</i>
Maïs grain humide	400	68,0	27 200
Maïs semence	0	0,0	0
Total maïs (grain et semences)	29 170	103,8	3 028 520
Maïs fourrage et ensilage (plante entière)	3 900	80,0	312 000

Tableau 5 : Statistiques agricoles sur les productions du maïs en 2018 (Agreste)

2. SUPERFICIES OBTENUES

Le Tableau 6 présente les surfaces en hectares et en pourcentage entre le maïs irrigué et le non irrigué obtenus avec les classifications, ceux avec les différents post-traitements, celles publiées par l'Agreste et les données de références fournies par la DRAAF. Pour les données de référence, la méthode de régularisation proposée par Iota² sans utiliser le RPG (Section III-H) permet d'améliorer les résultats avec des superficies très proches des observations terrain. A l'échelle du Loiret, le vote majoritaire est la méthode la plus performante des trois, mais elle dégrade les résultats. En effet, les superficies obtenues sans simplification, pour la meilleure classification (Classif_Seed3.tif), sont les plus proches des données de l'Agreste. On note un décalage de près de 1 000 ha de la superficie totale du maïs entre les données de l'Agreste (33 070 ha) et celle du RPG (32 238 ha).

	Configuration : STACK	Données de référence (Erosion 10m)				Loiret			
		Superficie maïs non irrigué	% non irrigué*	Superficie maïs irrigué	% irrigué	Superficie maïs non- irrigué	% non irrigué	Superficie maïs irrigué	% irrigué
	Observation terrain	275	18	1 242	82	-	-	-	-
	Conjoncture Agreste déc.2018	-	-	-	-	9 350	28	23 720	72
	Meilleure classification Classif_Seed 3	250	17	1 267	83	8 607	27	23 631	73
	Classification fusionnée Fusion 5 modèles	269	18	1 248	82	10 212	32	22 026	68
1	Vote majoritaire Classif_Seed 3	492	32	1 026	68	7 415	23	24 823	77
2	Régularisation	280	18	1 237	82	11 191	35	21 042	65
3	Vote majoritaire Régularisation	738	49	779	51	11 563	36	20 670	64

* % non irrigué = (superficie maïs non irrigué / superficie maïs totale) X 100

Tableau 6 : Surfaces obtenues par culture et validation des résultats

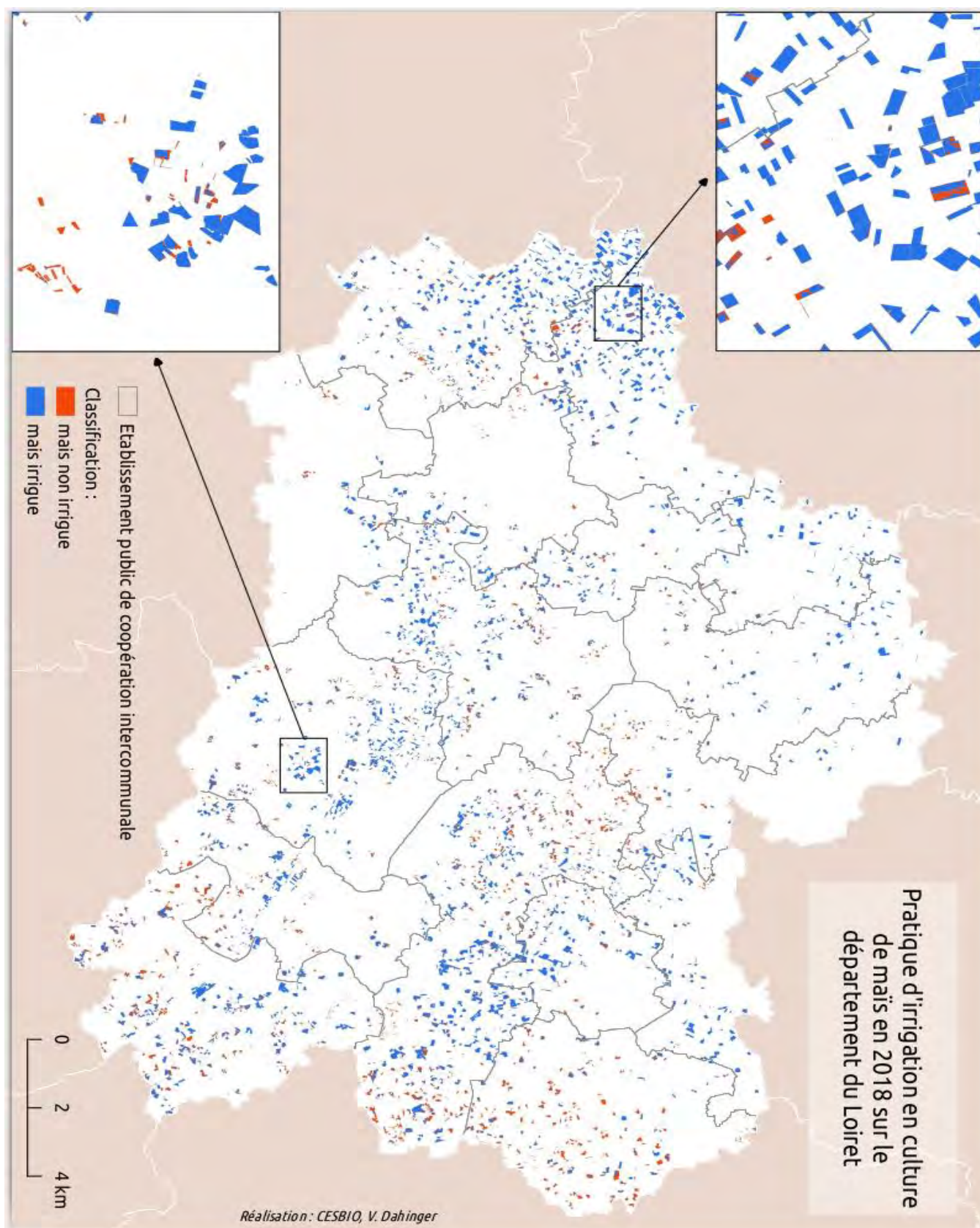


Figure 39 : Cartographie du maïs irrigué et non irrigué du Loiret en 2018 (configuration "stack")

CONCLUSION & PERSPECTIVES

Le libre accès à des données satellitaires prétraitées, à haute résolution spatiale et temporelle, est un levier de développement inédit pour la détection automatique des surfaces irriguées. En France, la distinction des pratiques culturales du maïs et plus particulièrement de l'apport artificiel en eau, est un besoin des services du Ministère de l'Agriculture pour estimer les niveaux de production, en fin et en cours de saison. Le stage a permis de tester, pour le compte de la DRAAF Centre-Val de Loire, la faisabilité et la pertinence d'une détection automatique sur le département du Loiret. L'étude de l'intérêt des images Sentinel-2 pour la distinction entre le maïs irrigué et non irrigué en 2017 et 2018 a été évaluée avec la chaîne de traitement de classification supervisée *iota*² configurée avec l'algorithme *Random Forest*. La vaste plaine du Loiret, traversée par la Loire, est composée principalement de terres agricoles et de forêts. Le climat tempéré de type océanique dégradé s'accompagne de périodes de sécheresse de plus en plus marquées. L'été pluvieux associé au printemps et à l'automne, plus sec, observé en 2017 s'oppose à l'extrême sécheresse de l'été et du début d'automne constatés en 2018. Les sécheresses ont des répercussions directes sur le développement du maïs, particulièrement efficient à l'utilisation de l'eau. Cette caractéristique est la clé de la distinction, à partir d'images optiques, des cultures de maïs irrigué et non irrigué. L'absence de sécheresse marquée pendant l'été 2017 associée à l'importante couverture nuageuse qui dégrade les données Sentinel-2, sont des freins à la répliquabilité de la méthode de classification (*Overall Accuracy* : 0,63 en 2017 / 0,85 en 2018). Les méthodes d'échantillonnage testées lors de l'étude ont montré qu'un jeu d'apprentissage déséquilibré entre les cultures (parcelles plus grandes et plus nombreuses pour le maïs irrigué) impacte les résultats. La sélection aléatoire d'un nombre de pixels constant (10 000 pixels) pour chaque culture améliore les résultats (OA : 0,63 contre 0,60 en 2017 / 0,83 contre 0,79 en 2018). De plus, l'utilisation des données correspondant à la période de culture du maïs (mai à octobre inclus) en 2018 donne de meilleurs résultats que le traitement des données de l'année entière. Les bandes *vegetation red edge* des satellites Sentinel-2 permettent d'évaluer la vigueur d'une plante et de décliner des indices spectraux qui facilitent la discrétisation (NDVI, NDWI, NDRE-1, NDRE-2). A partir de la série temporelle Sentinel-2 interpolée tous les 10 jours (produite par *iota*²), les indices spectraux traités en cumuls saisonniers (1 image par saison de culture de maïs) et en cumuls mensuels (1 image par mois observé) soulignent

les écarts de développement du maïs liés aux pratiques (irrigué ou non irrigué) et les dynamiques temporelles. En 2018, à partir des analyses en fin de saison, les résultats montrent que les indices traités en cumuls mensuels sur les 6 mois de culture et utilisés conjointement (OA : 0,88) sont plus pertinents que les cumuls saisonniers (OA : 0,85). Dans les deux cas, le NDWI est un indice pertinent pour identifier les cultures irriguées (OA : 0,87 et 0,80 respectivement). De plus, la méthode de traitement permet de se rapprocher des résultats obtenus à partir de la série temporelle Sentinel-2 interpolée tous les 10 jours (OA : 0,89). Or l'utilisation des cumuls mensuels diminue considérablement le nombre d'images à traiter puisque le nombre de bandes passent de 180 à 6 seulement. Dans ces conditions, les temps de traitement sont divisés par sept (2h contre 24 min passées sur le processeur) et les besoins de stockage sont cinq fois plus faibles (129 Go contre 26 Go). Pour autant, l'intégration dans *iota*² des indices traités en cumuls nécessitent un travail préliminaire important. Pour la même année, les analyses en cours de saison, réalisées à partir des cumuls mensuels des indices révèlent que la distinction entre les deux cultures est possible à partir de la fin du mois de juillet (OA : 0,87). L'ajout des données *GlobalSoilMap* relatives à la composition du sol n'améliore pas les résultats. Par ailleurs, la proximité des erreurs d'omission et de commission observées à partir des matrices de confusion montrent que les écarts importants de résultats entre le maïs non irrigué (F-score : 0,69) et le maïs irrigué (F-score : 0,93) conservent les proportions des données de référence. Malgré les confusions, la méthode d'estimation des superficies des deux cultures par télédétection présente en 2018 de bons résultats. Ainsi, à partir des images Sentinel-2 on estime la superficie totale du maïs non irrigué du Loiret à 8 607 ha (Agreste : 9 350 ha) et 23 631 ha de maïs irrigué (Agreste : 23 720 ha), soit 27 % et 73 % respectivement, ce qui représente un delta d'un point seulement avec les superficies réelles de l'Agreste. En revanche, les différents post-traitements de simplification réalisés sur les classifications dégradent les résultats.

Les résultats peu satisfaisants obtenus pour l'année 2017 soulignent le travail de recherche à poursuivre pour obtenir une méthode de distinction des pratiques culturales du maïs parfaitement transposable et reproductible. Le manque de données Sentinel-2 exploitables en raison de l'importante nébulosité de l'été 2017, n'est pas entièrement compensé par l'interpolation linéaire réalisée par *iota*². Le lissage des comportements des végétaux qui en résulte complexifie leur analyse et montre les limites de la méthode. Conçu pour capter les hyperfréquences moins sensibles à l'atmosphère, le radar est particulièrement pertinent pour le suivi des zones où la couverture nuageuse est fréquente ou importante. Les satellites Sentinel-1 fournissent gratuitement des images à haute résolution spatiale et temporelle. Dotés de radar à bande C à synthèse d'ouverture (SAR), Sentinel-1 est notamment pertinent pour l'observation du maïs (Fieuzal, 2013). La combinaison des images optiques et radar devrait faciliter l'élaboration d'une méthode plus générique dans le temps et l'espace. Ces différentes perspectives seront étudiées dans les prochains travaux.

BIBLIOGRAPHIE

- Balaghi R. (2006). *Wheat grain yield forecasting models for food security in Morocco*. Thèse en Science de l'Environnement. Université de Liège, septembre 2006.
- Battude M. (2017). *Estimation des rendements, des besoins et consommations en eau du maïs dans le Sud-Ouest de la France : apport de la télédétection à hautes résolutions spatiale et temporelle*. Thèse en Hydrologie. Université Toulouse III - Paul Sabatier, 3 février 2017.
- Cheret V., Denux J-P., Gacherieu C., Ortisset J-P (2011). "Utilisation de séries temporelles d'images satellitaires pour cartographier le dépérissement des boisements résineux du Sud Massif Central". In : *Rendez-vous Techniques de l'ONF* (n° 31). pp. 55-62. ISSN 1763-6442.
- Demarez V. (2018). *Fonctionnement hydrique des cultures à l'échelle du territoire*. Mémoire d'Habilitation à Diriger des Recherches. Université Toulouse III - Paul Sabatier, 25 janvier 2018.
- Demarez V., Helen F., Marais-Sicre C. and Baup F. (2019). "In-season mapping of irrigated crops using Landsat-8 and Sentinel-1 time series". In : *Remote Sensing 2019*, 11, 118. Published : 10 janvier 2019.
- Dheeravath, V.; Thenkabail, P.S.; Chandrakantha, G.; Noojipady, P.; Reddy, G.P.O.; Biradar, C.M.; Gumma, M.K.; Velpuri, M. "Irrigated areas of India derived using MODIS 500 m time series for the years 2001–2003". In : *Remote Sensing* 2010, 65, 42–59.
- Dobarco R., Bourennane H., Arrouays D., Saby N., Cousin I., Martin M. (2019). "Uncertainty assessment of GlobalSoilMap soil available water capacity products: A French case study". In : *Geoderma* 344 (2019) 14-30.
- Ferrant S., Selles A., Le Page M., Herrault P-A., Pelletier C., Al-Bitar A., Mermoz S., Gascoin S., Bouvet A., Saqalli M., Dewandel B., Caballero Y., Ahmed S., Maréchal J-C. and Kerr Y. (2017). "Detection of irrigated crops from Sentinel-1 and Sentinel-2 data to estimate seasonal groundwater use in south India". In : *Remote sensing* 2017, 9, 1119. Published : 3 novembre 2017.
- Fieuzal R. (2013). *Apports des données radar pour l'estimation des paramètres biophysiques des surfaces agricoles*. Thèse en Sciences de la Terre. Université Toulouse III - Paul Sabatier, 2013.
- Gaussen H. (1956), « L'étude des climats pour les courbes ombrothermiques : application à la cartographie », *L'information Géographique*, p. 191-193.
- Geffard S. (1994), "Utilisation d'une série temporelle de données satellite NOAA-AVHRR pour le suivi hydrologique des grands fleuves en Afrique de l'Ouest". *Xe journées hydrologiques - Orstom*, septembre 1994, p. 685-695.
- Genuer R., Poggi J-M (2016). *Arbres CART et Forêts aléatoires, Importance et sélection de variables*. 2016. hal-01387654v.

- Goetz, S.J.; Varlyguin, D.; Smith, A.J.; Wright, R.K.; Prince, S.D.; Mazzacato, M.E.; Tringe, J.; Jantz, C.; Melchoir, B. « Application of multitemporal Landsat data to map and monitor land cover and land use change in the Chesapeake Bay watershed ». In : *Analysis of Multi-temporal Remote Sensing Images*; World Scientific: Singapore, 2004.
- Inglada J., Vincent A., Arias M., Tardy B., Morin D. and Rodes I. (2017). "Operational high resolution land cover map production at the country scale using satellite image time series". In : *Remote Sensing* 2017, 9, 95. Published : 22 janvier 2017
- Liaw and Wiener (2002). *Classification and regression by random Forest*. R News, 2(3):18–22, 2002.
- Lonjou V., Desjardins C., Hagolle O., Petrucci B., Tremas T., Dejus M., Makarau A., and Auer S. "MACCS-ATCOR joint algorithm (MAJA)", Proc. SPIE 10001, Remote Sensing of Clouds and the Atmosphere XXI, 1000107 (19 October 2016).
- Mulder V.L., Lacoste M., Richer-de-Forges A.C., Arrouays D. (2016). "GlobalSoilMap France: High-resolution spatial modelling the soils of France up to two meter depth". In : *Elsevier, Science of the Total Environment* 573 (2016) 1352-1369.
- Pan Y., Li X., Gong P., He C., Shi P. & Pu R. (2003). *An integrative classification of vegetation in China based on NOAA AVHRR and vegetation-climate indices of the Holdridge life zone*, Intern
- Pelletier C. (2017). *Cartographie de l'occupation des sols à partir des séries temporelles d'images satellitaires à hautes résolutions. Identification et traitement des données mal étiquetées*. Thèse en Hydrologie. Université Toulouse III - Paul Sabatier, 11 décembre 2017.
- Vuolo F., Neuwirth M., Immitzer M., Atzberger C., Ng W-T. (2018). "How much does multi-temporel Sentinel-2 data improve crop type classification?". In : *Elsevier, International Journal of Applied Earth Observation and Geoinformation* 72 (2018) 122-130.
- Zheng B., Myint S., Thenkabail P., Aggarwal R. (2015). "A support vector machine to identify irrigated crop types using time-series Landsat NDVI data". In : *Elsevier, International Journal of Applied Earth Observation and Geoinformation* 34 (2015) 103-112.
- Zhou T., Pan J., Zhang P., Wei S. & Han T. (2017). "Mapping winter wheat with multi-temporal SAR and optical images in an urban agricultural region". In : *Sensors* 2017, 17, 1210. Published : 25 may 2017.